NAUKA-DYDAKTYKA-PRAKTYKA

28556+CD
nr 2

BIBLIOTEKI CYFROWE



WYDAWNICTWO SBP

MONOGRAFIA – BILIOTEKI CYFROWE

SBP STOWARZYSZENIE BIBLIOTEKARZY POLSKICH 95 lat!

Polish Librarians Association

SCIENCE – DIDACTICS – PRACTICE

Digital Libraries

Collective work edited by
Małgorzata Janiak, Monika Krakowska and Maria Próchnicka

WYDAWNICTWO SBP

Warsaw 2012

JAGIELLONIAN UNIVERSITY IN KRAKOW

Stowarzyszenie Bibliotekarzy Polskich

NAUKA – DYDAKTYKA – PRAKTYKA

Biblioteki cyfrowe

Praca zbiorowa pod redakcją
Małgorzaty Janiak, Moniki Krakowskiej i Marii Próchnickiej

WYDAWNICTWO SBP

Warszawa 2012

UNIWERSYTET JAGIELLOŃSKI W KRAKOWIE

Komitet Redakcyjny serii wydawniczej
«NAUKA – DYDAKTYKA – PRAKTYKA»

Jacek WOJCIECHOWSKI (przewodniczący), Stanisław CZAJKA,
Artur JAZDON, Danuta KONIECZNA, Dariusz KUŹMINA, Krzysztof MIGOŃ,
Mieczysław MURASZKIEWICZ, Janusz NOWICKI (sekretarz),
Joanna PAPUZIŃSKA-BEKSIAK, Wanda PINDEL, Maria PRÓCHNICKA,
Jadwiga SADOWSKA, Barbara SOSIŃSKA-KALATA, Barbara STEFANIAK,
Elżbieta STEFAŃCZYK, Hanna TADEUSIEWICZ

Publikacja dofinansowana przez Instytut Informacji Naukowej i Bibliotekoznawstwa Uniwersytetu Jagiellońskiego

Recenzja
dr hab. prof UJ Wanda PINDEL

Projekt okładki
Grzegorz HAMERA

Redakcja techniczna i korekta
Marta LACH

© Copyright by Stowarzyszenie Bibliotekarzy Polskich
© Copyright by Instytut Informacji Naukowej i Bibliotekoznawstwa Uniwersytet Jagielloński

CIP – Biblioteka Narodowa

Biblioteki cyfrowe : praca zbiorowa / pod red.
Małgorzaty Janiak, Moniki Krakowskiej i Marii
Próchnickiej ; Stowarzyszenie Bibliotekarzy
Polskich. - Warszawa : Wydawnictwo Stowarzyszenia
Bibliotekarzy Polskich, 2012. - (Nauka, Dydaktyka,
Praktyka ; 135)

Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich
00-335 Warszawa, ul. Konopczyńskiego 5/7 tel. 22 827 52 96
www.sbp.pl; wydawnictwo@sbp.pl, biuro@sbp.pl
Warszawa 2012. Wyd 1. Ark. wyd 35,5 Ark. druk. 34,25
Łamanie: Funky Worky Studio Składu Komputerowego
kontakt@funkyworky.pl
Druk i oprawa: Drukarnia i Introligatornia OPRAWA Sp. z o.o.
ul. Dowborczyków 17, 90-019 Łódź, tel. 42 676 42 82

Spis treści

Małgorzata Janiak, Monika Krakowska, Maria Próchnicka
Wstęp ... 9

Część 1
Część ogólna ... 13

Małgorzata Janiak
Biblioteka cyfrowa, biblioteka elektroniczna, biblioteka wirtualna ... 15

Władysław Marek Kolasa
Biblioteki cyfrowe na świecie – powstanie i rozwój ... 66

Monika Krakowska
Biblioteki cyfrowe w świetle międzynarodowych konferencji 1993-2011 ... 96

Mirosław Górny
Biblioteki cyfrowe w pracy naukowca humanisty ... 126

Małgorzata Janiak, Monika Krakowska
Ocena jakości bibliotek cyfrowych ... 135

Sybilla Stanisławska-Kloc
Biblioteki cyfrowe a regulacje prawne ... 179

Małgorzata Kowalska
Koszty i finansowanie digitalizacji ... 233

Małgorzata Janiak
Zagadnienia estetyczne, a biblioteki cyfrowe ... 252

Bożena Bednarek-Michalska
Standardy i dobre praktyki budowania kolekcji cyfrowych ... 269

Magdalena Wójcik
Literatura na temat bibliotek cyfrowych – analiza bibliometryczna ... 284

Część 2
Ludzie związani z bibliotekami cyfrowymi ... 295

Jolanta Mazurek
Użytkownicy polskich bibliotek cyfrowych ... 297

Monika Krakowska
Modele mentalne użytkowników bibliotek cyfrowych ... 309

Łukasz Mesek
Model kompetencji zawodowych bibliotekarza cyfrowego ... 330

Część 3
Zasoby i procesy 339

Cezary Mazurek, Tomasz Parkoła, Marcin Werla
Tworzenie i rozwój bibliotek cyfrowych 341

Marek Nahotko, Piotr Myszkowski, Łukasz Mesek, Danuta Patkaniowska
Struktury danych 362
Metadane w polskich bibliotekach cyfrowych 372
Odwzorowanie treści dokumentów elektronicznych 382
Opracowanie rzeczowe dokumentu 384
Korelacje między opisami w katalogach bibliotecznych i bibliotekach cyfrowych 396

Władysław Marek Kolasa
Formaty dokumentów w bibliotekach cyfrowych 403

Bartłomiej Siedlarz
DjVu 423

Tomasz Kalota, Marcin Szala
Organizacja i logistyka digitalizacji 437

Część 4
Przykładowe polskie biblioteki cyfrowe (case studies) 447

Cezary Mazurek, Marcin Werla
Federacja Bibliotek Cyfrowych – studium przypadku 449

Ewa Wichlińska
Wielkopolska Biblioteka Cyfrowa – analiza przypadku 468

Joanna Potęga
Cyfrowa Biblioteka Narodowa Polona – (cBN Polona) 475

Bożena Bednarek-Michalska
Kujawsko-Pomorska Biblioteka Cyfrowa – studium przypadku 486

Wojciech Kowalewski, Władysław Marek Kolasa
Małopolska Biblioteka Cyfrowa – regionalne repozytorium wiedzy 494

Remigiusz Lis
Śląska Biblioteka Cyfrowa – analiza przypadku 507

Łukasz Mesek
Jagiellońska Biblioteka Cyfrowa – opis przypadku 522

Summary 547

CD-Rom z ilustracjami

Contents

Małgorzata Janiak, Monika Krakowska, Maria Próchnicka
Introduction 9

Part 1
Overal issues 13

Malgorzata Janiak
Digital library, electronic library, virtual library 15

Władysław Marek Kolasa
Origins and development of digital libraries around the world 66

Monika Krakowska
International conferences on digital libraries 1993-2011 96

Mirosław Górny
The role of digital libraries in research work in humanities 126

Małgorzata Janiak, Monika Krakowska
Quality evaluation of digital libraries 135

Sybilla Stanisławska-Kloc
Digital libraries vs. legal requirements 179

Małgorzata Kowalska
Costs and financing of digitalisation 233

Małgorzata Janiak
Digital libraries vs. aesthetic issues 252

Bożena Bednarek-Michalska
Standards and good practices in developing digital collections 269

Magdalena Wójcik
Bibliometric analysis of literature on digital libraries 284

Part 2
People engaged in digital libraries' development and use 295

Jolanta Mazurek
The users of the Polish digital libraries 297

Monika Krakowska
Mental models of digital libraries' users 309

Łukasz Mesek
The model of digital librarian professional competencies 330

Part 3
Resources and processes 339

Cezary Mazurek, Tomasz Parkoła, Marcin Werla
Werla Digital libraries' formation and development 341

Marek Nahotko, Piotr Myszkowski, Łukasz Mesek, Danuta Patkaniowska
Data structures 362
Metadata in Polish digital libraries 372
Mapping content of electronic documents 382
Indexing documents 384
Correlations between library catalogues' and digital libraries' descriptions 396

Władysław Marek Kolasa
Documents' formats in digital libraries 403

Bartłomiej Siedlarz
DjVu 423

Tomasz Kalota, Marcin Szala
Organisation and logistics of digitalisation 437

Part 4
Polish digital libraries – case studies 447

Cezary Mazurek, Marcin Werla
Digital Libraries Federation 449

Ewa Wichlińska
Digital Library of Wielkopolska 468

Joanna Potęga
National Digital Library 475

Bożena Bednarek-Michalska
Kujawsko-Pomorska Digital Library 486

Wojciech Kowalewski, Władysław Marek Kolasa
Małopolska Digital Library – regional knowledge repository 494

Remigiusz Lis
Silesian Digital Library 507

Łukasz Mesek
Jagiellonian Digital Library 522

Summary 547

CD-Rom with illustrations

Wstęp

Małgorzata Janiak, Monika Krakowska, Maria Próchnicka

Przekazywana do rąk czytelnika monografia dotycząca bibliotek cyfrowych obejmuje całokształt zagadnień poczynając od genezy bibliotek cyfrowych, problemów terminologicznych, prawnych, poprzez opis procesów, metod, modeli, aż po omówienie istniejących bibliotek (*case-study*). W zamyśle autorów objęcie tak szerokiego zakresu tematycznego miało na celu przedstawienie czytelnikom zainteresowanym bibliotekami cyfrowymi całej problematyki ich dotyczącej w jednej książce. Do opracowania poszczególnych rozdziałów zostali zaproszeni teoretycy i praktycy zajmujący się tworzeniem bibliotek cyfrowych z całej Polski.

Odbiorcami książki mogą więc być zarówno pracownicy naukowi zainteresowani tworzeniem i wykorzystywaniem zasobów cyfrowych, praktycy budujący i planujący budować biblioteki cyfrowe, studenci oraz obecni i przyszli użytkownicy kolekcji cyfrowych.

*

Biblioteka cyfrowa to termin heterogeniczny, odnoszący się do różnorodnych systemów tworzonych na całym świecie, których głównym zadaniem jest przechowywanie i udostępnianie zasobów z zakresu nauki i kultury dla współczesnych użytkowników oraz przyszłych pokoleń. W licznych pracach bibliotekami cyfrowymi nazywane są zarówno repozytoria metadanych i cyfrowych obiektów, systemy linkujące, odsyłające do innych serwisów, archiwa, systemy zarządzania treścią oraz złożone systemy, które łączą w sobie zaawansowane usługi biblioteczne.

*

Niniejsza monografia lokuje się w cyklu książek wydanych w XXI wieku w Polsce, których tematyka związana była z bibliotekami cyfrowymi lub z procesami digitalizacyjnymi. Jedną z pierwszych była wydana w 2002 r. książka Juranda B. Czermińskiego *Cyfrowe środowisko współczesnej biblioteki*. W publikacji, poświęconej głównie samemu systemowi, na który składają się zasoby ludzkie, sprzętowe, programowe oraz wszelkiego rodzaju dane, nie występował termin „biblioteka cyfrowa", czy „biblioteka wirtualna".

Do zagadnień związanych z digitalizacją, stanem projektów digitalizacyjnych w wielu krajach oraz wykorzystywaniem zasobów tworzonych w tych procesach dla różnych dyscyplin, głównie w odniesieniu do aspektu komunikowania się, odwołuje się opublikowana w 2004 r. książka *Cyfryzacja w procesach komunikowania* pod redakcją Wandy Krzemińskiej i Piotra Nowaka. W książce tej poza opracowaniami zagadnień praktycznych znajdują się też analizy terminologiczne pojęć „cyfryzacja" i „digitalizacja".

Następne publikacje odnoszące się przede wszystkim do procesów digitalizacyjnych oraz tworzenia bibliotek cyfrowych to: praca Aleksandra Trembowieckiego *Digitalizacja zbiorów bibliotecznych: teoria i praktyka* oraz materiały z konferencji ogólnopolskiej (Warszawa, 2005) *Dygitalizacja zbiorów bibliotecznych* pod redakcją Elżbiety Stefańczyk oraz organizowanej przez Bibliotekę Politechniki Łódzkiej pod znaczącym tytułem *Biblioteki XXI wieku: czy przetrwamy?* (wszystkie prace wydane w 2006 r.).

W 2008 r. opublikowana została książka, *Standardy w procesie digitalizacji obiektów dziedzictwa kulturowego* pod red. Grzegorza Płoszajskiego poświęcona kwestiom technicznym, dotyczącym samego procesu digitalizacji, a zwłaszcza formatom danych i różnego typu metadanym, opisującym zarówno oryginalny obiekt digitalizowany, jak i proces digitalizacji oraz strukturę złożonego obiektu cyfrowego.

Ważny dla działań digitalizacyjnych jest także wydany w 2010 r. przez Bibliotekę Narodową, która spełnia funkcję Centrum Kompetencji w zakresie digitalizacji materiałów bibliotecznych, podręcznik *Digitalizacja piśmiennictwa* w opracowaniu i pod redakcją Dariusza Paradowskiego. Książka ta omawia proces digitalizacji, począwszy od planowania projektów, przez prace przygotowawcze, aż po opis i przedstawienie praktycznych wskazówek dotyczących działań digitalizacyjnych.

Z innych prac, analizujących biblioteki cyfrowe, a zwłaszcza procesy digitalizacji wymienić należy pracę zbiorową pod red. Jadwigi Woźniak-Kasperek i Jerzego Franke *Biblioteki cyfrowe: projekty, realizacje, technologie* (2007), książkę Małgorzaty Kowalskiej *Dygitalizacja zbiorów bibliotek polskich* (2008), czy Anny Wałek *Biblioteki cyfrowe na platformie dLibra* (2009).

Przypomnieć też należy materiały konferencyjne, *Informacja dla nauki a świat zasobów cyfrowych* pod red. Haliny Ganińskiej (2008), *Uniwersum piśmiennictwa wobec komunikacji elektronicznej* pod red. Krzysztofa Migonia i Marty Skalskiej-Zlat, czy *Cyfrowy świat dokumentu: wydawnictwa, biblioteki, muzea, archiwa* pod red. Henryka Hollendra, w których to zamieszczono artykuły związane z bibliotekami cyfrowymi. Publikacje konferencyjne w całości poświęcone praktyce i teorii bibliotek cyfrowych to materiały z odbywanych corocznie w Poznaniu ogólnopolskich i międzynarodowych konferencji *Polskie Biblioteki Cyfrowe* (2008, 2009, 2010, 2011), których organizatorami są Biblioteka Kórnicka PAN, Poznańska Fundacja Bibliotek Naukowych oraz Poznańskie Centrum Superkomputerowo-Sieciowe.

*

Monografia podzielona została na cztery części: 1) część ogólną, w której znalazły się artykuły związane z zagadnieniami teoretycznymi, historycznymi, prawnymi i fi-

nansowymi dotyczącymi bibliotek cyfrowych, 2) część poświęconą ludziom związanym z bibliotekami cyfrowymi (użytkownikom, twórcom i pracownikom bibliotek), 3) część omawiającą procesy i projekty tworzenia kolekcji cyfrowych oraz 4) część, w której przedstawione zostały analizy istniejących bibliotek cyfrowych.

W artykule autorstwa Małgorzaty Janiak rozpoczynającym część ogólną omówione zostały – na podstawie analizy piśmiennictwa polskiego i zagranicznego – problemy terminologiczne związane z zakresem znaczeniowym oraz wzajemnymi relacjami takich terminów jak 'biblioteka elektroniczna', 'biblioteka cyfrowa', 'biblioteka wirtualna'. W dwóch kolejnych artykułach: Władysława Marka Kolasy oraz Moniki Krakowskiej przedstawiona została historia tworzenia projektów i samych bibliotek cyfrowych na świecie i w Polsce oraz tematyka podejmowana na konferencjach poświęconych bibliotekom cyfrowym. Analiza bibliometryczna literatury na temat bibliotek cyfrowych przygotowana została przez Magdalenę Wójcik. Zagadnienia prawne, istotne z punktu widzenia tworzenia i wykorzystywania bibliotek cyfrowych omówione zostały przez Sybillę Stanisławską-Kloc. Problemy standaryzacji na podstawie przyjętego przez wiele instytucji *A Framework of Guidance for Building Good Digital Collections* sporządzonego przez National Information Standards Organization (NISO) Framework Working Group, z pomocą Institute of Museum and Library Services (IMLS) przedstawiła Bożena Bednarek-Michalska. Kryteria oceny jakości bibliotek cyfrowych przeanalizowane zostały przez Małgorzatę Janiak i Monikę Krakowską, a zagadnienia estetyczne przez Małgorzatę Janiak. Zagadnienia związane z finansowaniem bibliotek cyfrowych omówiła Małgorzata Kowalska. Mirosław Górny przedstawił zagadnienia wykorzystania zasobów kolekcji cyfrowych przez pracowników naukowych. Część druga monografii zawiera artykuły: Jolanty Mazurek oraz Moniki Krakowskiej poświęcone problematyce użytkowników bibliotek cyfrowych. Zagadnieniom kompetencji zawodowych bibliotekarza cyfrowego poświęcony jest artykuł Łukasza Meska.

Pierwszy artykuł części trzeciej odnoszącej się do zasobów bibliotek cyfrowych, procesów związanych z ich funkcjonowaniem, a także do tworzenia kolekcji (oprogramowanie, tworzenie BC: gromadzenie, opracowanie, wyszukiwanie, udostępnianie) przygotowany został przez Cezarego Mazurka, Tomasza Parkołę, Marcina Werlę. Tomasz Kalota i Marcin Szala przedstawili zagadnienia organizacyjne związane z digitalizacją. Opis formalny oraz rzeczowy omówiony został przez zespół autorów Łukasza Meska, Marka Nahotko, Piotra Myszkowskiego i Danutę Patkaniowską. Analiza formatów dokumentów przedstawiona została przez Władysława Marka Kolasę oraz Bartłomiej Siedlarza.

Ostatnia część monografii poświęcona została istniejącym bibliotekom cyfrowym. Część tą otwiera praca o Federacji Bibliotek Cyfrowych autorstwa Cezarego Mazurka i Marcina Werli. Następnie zamieszczono omówienia: najstarszej polskiej biblioteki cyfrowej stworzonej na platformie dLibra Wielkopolskiej Biblioteki Cyfrowej – Ewy Wichlińskiej oraz Ogólnopolskiej Narodowej Biblioteki Cyfrowej Polona – Joanny Potęgi. Biblioteki regionalne opracowane zostały: Kujawsko-Pomorska Biblioteka Cyfrowa przez Bożenę Bednarek-Michalską, Małopolska Biblioteka Cyfrowa przez

Wojciecha Kowalewskiego i Władysława Marka Kolasę, zaś Śląska Biblioteka Cyfrowa przez Remigiusza Lisa. Jagiellońska Biblioteka Cyfrowa, jako przykład biblioteki instytucjonalnej, przedstawiona została przez Łukasza Meska.

*

Redaktorki monografii mają nadzieję, iż wielość podejść do tematu bibliotek cyfrowych sprawi, że czytelnicy (pracownicy naukowi, praktycy, studenci oraz obecni i przyszli użytkownicy kolekcji cyfrowych) znajdą opracowania na interesujące ich tematy, a różna objętość opracowań, zależna wszakże od tematu i autora nie przeszkodzi w odbiorze treści. Jednocześnie dziękują wszystkim autorom za wzięcie udziału w tym przedsięwzięciu.

Część 1

Część ogólna

Biblioteka cyfrowa, biblioteka elektroniczna, biblioteka wirtualna

Małgorzata Janiak

1. Wstęp

Termin biblioteka cyfrowa jest terminem bardzo heterogenicznym, odnoszącym się do wielu systemów. Taka różnorodność znalazła swoje odzwierciedlenie także w wydanym ostatnio (2011) raporcie *The digital library reference model*, opracowanym przez DL.org: Coordination Action on Digital Library Interoperability, Best Practices and Modelling Foundations w ramach Network of Excellence on Digital Libraries. W pracy z 2002 r. Dietmara Wolframa i Hong (Iris) Xie *Traditional IR for web users: a context for general audience digital libraries* przedstawione zostały wcześniejsze badania uczonych: Mounir'a Khalil'a i Raji Jayatilleke (2000), w których użytkownicy końcowi opisali bibliotekę cyfrową na 35 różnych sposobów.

Zakres znaczeniowy terminu oscyluje od repozytoriów metadanych oraz cyfrowych obiektów, przez systemy linkujące, odsyłające do innych serwisów, archiwa, systemy zarządzania treścią, po złożone systemy, które łączą w sobie zaawansowane cyfrowe usługi biblioteki. Termin 'biblioteka cyfrowa' obejmuje zarówno istniejące obiekty, jak i same procesy tworzenia i dystrybucji wszystkich rodzajów informacji w sieciach, począwszy od konwersji materiałów historycznych aż do tego rodzaju informacji, które nie mają odpowiedników w świecie fizycznym. Do całego systemu zaliczeni zostali też ludzie związani z tworzeniem oraz wykorzystywaniem bibliotek cyfrowych – użytkownicy, klienci, gracze itp.

Sytuacja taka jest efektem braku zgody wśród specjalistów, czym tak naprawdę jest biblioteka cyfrowa, jaka powinna być jej budowa, funkcjonalność i misja [zob. np. wymieniony wcześniej *The Digital Library Reference Model*]. Współcześnie myśli się o bibliotekach cyfrowych jako o złożonych tworach, stąd obecnie przedstawiane definicje są na tyle ogólne, aby móc opisać różnorodne systemy.

W tej pracy starano się przedstawić znaczenia terminu 'biblioteka cyfrowa', jak i poszczególnych elementów budujących cały rozbudowany system, a więc jej zawartość, elementy technologiczne, budujące jego zawartość oraz użytkowników, którzy zarówno współtworzą, jak i wykorzystują biblioteki cyfrowe. Zakres pojęć omówiony został na podstawie piśmiennictwa polskiego oraz zagranicznego (głównie angielskojęzycznego).

W związku z tym, iż termin ewoluował i zmieniał swoje zakresy semantyczne oraz przewijał się na zmianę z pojęciami 'biblioteka elektroniczna', 'biblioteka wirtualna', 'biblioteka bioniczna', czy 'hiper-biblioteka' analiza ich zakresów przedstawiona została w artykule w podziale: dzieła encyklopedyczne, piśmiennictwo światowe, polskie oraz w układzie chronologicznym i tematycznym.

2. Terminy w encyklopediach, słownikach i leksykonach

W *Encyclopedia of library and information science* pod red. Aliena Kenta i Harolda Lancoura najwcześniej (w **1997** r. – t. 59) pojawiła się 'elektroniczna biblioteka przyszłości' – *Electronic libraries of the future*, autorstwa Philipa Barkera. Wirtualność została tutaj połączona z koncepcją książki, która jako ogólnoświatowe dziedzictwo kulturowe, zyskała nowe nośniki (CD-ROMy, sieć WWW itp.). Książki elektroniczna, przechowywane także w bibliotekach przyszłości, które muszą się na nie otworzyć i przygotować, podzielone zostały na statyczne (stałe strony, stała informacja), dynamiczne (rekonfigurowane strony, stała informacja) oraz 'żywe' (living) (rekonfigurowane strony, rekonfigurowana informacja). Natomiast wśród typów wymieniono książki: tekstowe, obrazkowe, mówione, z ruchomymi obrazami, multimedialne (tekst, obraz i dźwięk), polimedialne (wykorzystujące elektroniczne i nie-elektroniczne media), hypermedialne (zawierające nielinearną strukturę informacji), inteligentne (adaptujące się do zachowań użytkowników), telemedialne (wykorzystujące urządzenia telekomunikacji), cyberksiążki (wykorzystujące techniki wirtualnej rzeczywistości).

Biblioteki przyszłości natomiast określone zostały jako biblioteki polimedialne, elektroniczne, cyfrowe oraz wirtualne. 'Polimedialne' przechowują dane różnych mediów (papier, mikrofilm, dysk kompaktowy itp.). 'Elektroniczne' są skomplikowanym interdyscyplinarnym systemem z własnymi specyficznymi cechami, wymaganiami i problemami. W 'cyfrowych' informacja istnieje wyłącznie w formie elektronicznej. Natomiast 'wirtualne' związane są z rzeczywistością wirtualną. W analizach tych widać więc progres samych bibliotek, złączony ze zmiennością form i nośników przechowywanych oraz dostarczanych materiałów, a także z systemami umożliwiającymi wyszukiwanie danych.

W kolejnym artykule tej encyklopedii, którego autorkami są Virginia D. Sorkin oraz Judith Farley, *National Digital Library* (t. 62 – **1998**) pojawił się termin 'biblioteki cyfrowej' (digital library). Jej definicja odniesiona została do tradycyjnej biblioteki, której podstawowymi zadaniami jest selekcja, katalogowanie, udostępnianie oraz

zabezpieczanie kolekcji. Podstawową różnicą jest fakt posiadania przez bibliotekę cyfrową danych czytelnych dla komputerów (machine-readable). Cechy tworzące taką bibliotekę, zidentyfikowane przez Association of Research Libraries (**1995**), to:

a) nie bycie pojedynczą jednostką,
b) dostęp do technologii sieciowych umożliwiających połączenie się wielu jednostkom,
c) dostępność (transparentność) linków dla końcowych użytkowników,
d) ustalenie celu: uniwersalny dostęp do cyfrowych zasobów i informacji,
e) nie ograniczanie się do elektronicznych 'kopii' rzeczywistych obiektów, ale też posiadanie cyfrowych artefaktów, które nie mają swoich odpowiedników w tradycyjnych formatach.

Koncepcja biblioteki cyfrowej zakotwiczona została w marzeniu stworzenia wirtualnej biblioteki, ogromnej, idealnej, uniwersalnej kolekcji informacji, do której istnieje stały dostęp, bez względu na to, gdzie fizycznie się one znajdują. Historycznymi protoplastami idei stali się twórcy Biblioteki Aleksandryjskiej, Paul Otlet, Vannevar Bush i Ted Nelson.

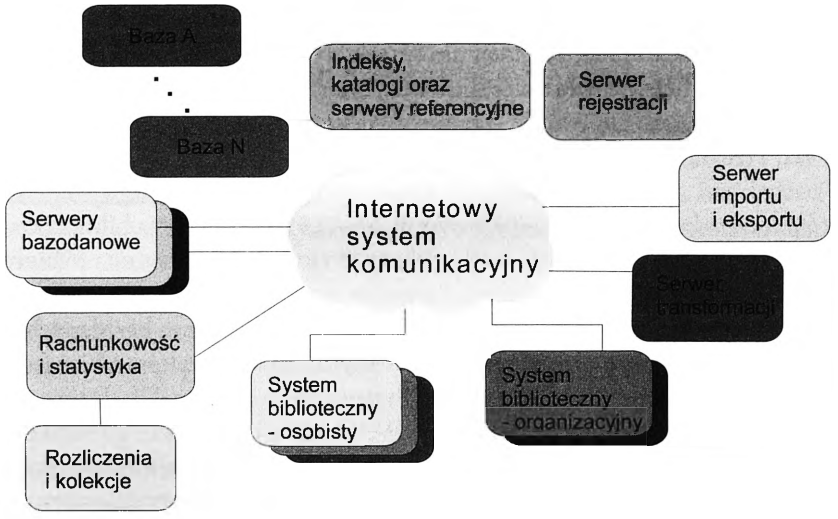
W następnym artykule tej samej encyklopedii autorstwa Gary'ego Marchionini *Research and development in digital libraries* (t. 63 – **1998**) koncepcja biblioteki cyfrowej została przypisana do różnych środowisk, z zaznaczeniem, że jest ona różnorodnie przez nie rozumiana. Dla środowiska związanego z komputerami (informatyków, inżynierów itp.) jest metaforą dla nowego sposobu dystrybuowania baz danych, które pozwalają zarządzać także danymi multimedialnymi. Dla polityków i biznesmenów jest nowym rynkiem dla źródeł i serwisów informacji światowej. Dla futurystów reprezentuje manifestację światowego umysłu Wells'a. Dla bibliotekarzy i informatologów jest logicznym rozszerzeniem i fizycznym powiększeniem bibliotek w społeczeństwie informacyjnym. I jako taka wymaga przy jej fizycznym tworzeniu opracowania problematyki związanej z jej zawartością (np. typami danych) oraz zarządzaniem (selekcja, gromadzenie, indeksowanie, przechowywanie, dostęp), utrzymaniem, patronami (zarządzający, jak i stali klienci), serwisami (typy, jakość, przeszukiwalność, selekcja, instrukcje) oraz przygotowania analiz ekonomicznych oraz kulturowych. Dodatkowe zalecenie przekazane przez autorów mówiło o tym, że za każdym razem, poza stosowaniem się do wyników ogólnych analiz, czy propozycji, twórcy powinni także opracować zagadnienia specyficzne dla konkretnych kolekcji.

*

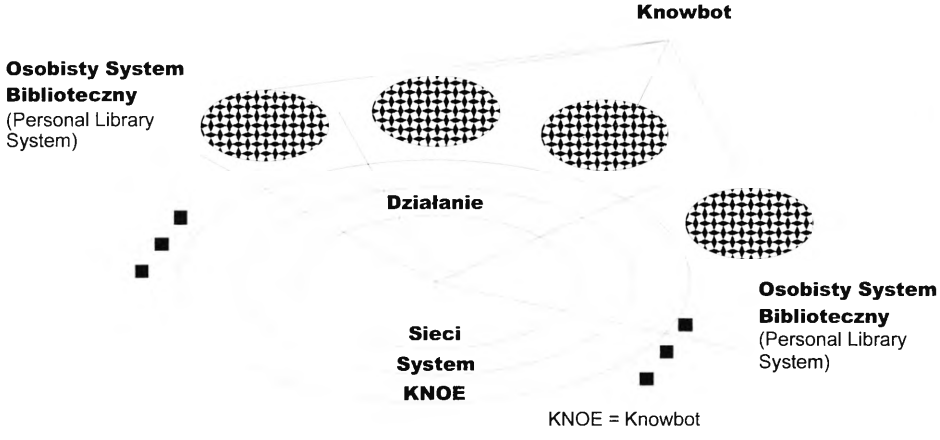
Inna encyklopedia, najpopularniejsza wśród społeczności internautów *Wikipedia* prezentuje hasło 'biblioteka cyfrowa (digital library)'. Jest to biblioteka, w której zbiory przechowywane lokalnie lub zdalnie dostępne są w formatach cyfrowych (w przeciwieństwie do druku, mikroform lub innych mediów) poprzez komputery, podłączone do sieci, właściwie z każdego miejsca świata [http://en.wikipedia.org/wiki/Digital_library]. Biblioteka cyfrowa jest więc rodzajem systemu wyszukiwania informacji, jej przechowywania oraz zarządzania. Definicja ta oparta jest na pracy *The Digital Library: a biography* autorstwa Daniela Greensteina i Suzanne E. Thorin, przygotowanej pod auspicjami Digital Library Federation, przedstawionej w ostatecznej

wersji w **2002** r. W tekście tym przedstawiono rozwój bibliotek cyfrowych jako naturalny proces rozwoju od systemów umożliwiających migrację katalogów do formy dostępnej online, przez serwisy usługowe: pomocne w odszukiwaniu i wykorzystaniu materiałów archiwalnych, nagranego dźwięku, grafiki i fotografii oraz innych kolekcji (model hybrydowy) po rozbudowane systemy zawierające materiały cyfrowe.

W *Wikipedii* podano też datę **1988** jako rok pierwszego użycia terminu 'biblioteka cyfrowa' – w sprawozdaniu Corporation Inicjatyw National Research, przygotowanym przez Roberta E. Kahna i Vintona G. Cerfa. Tekst ten skupiony był wokół 'systemu biblioteki cyfrowej' (Digital Library System), którego struktura oraz sekcje przedstawione zostały na rys. 1 i 2.

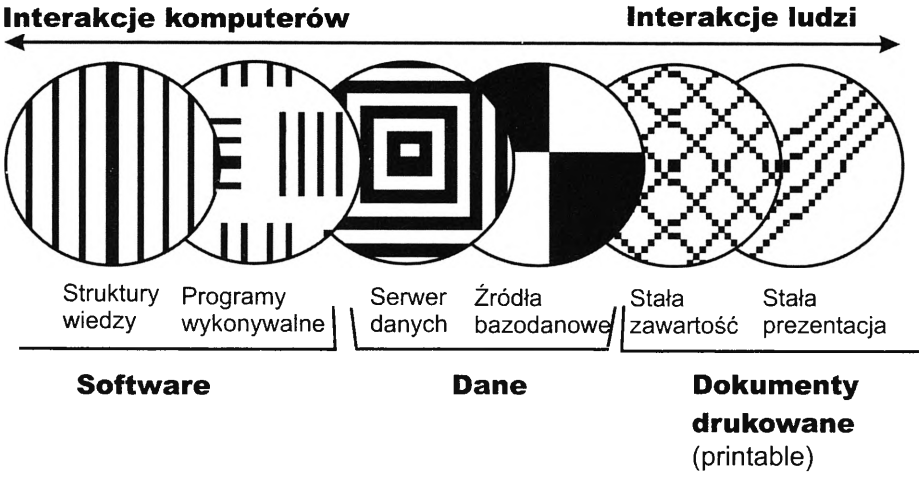


Rys. 1 – struktura systemu biblioteki cyfrowej – za Kahn; Cerf (1988), s. 17.



Rys. 2 – sekcje systemu biblioteki cyfrowej – za Kahn; Cerf (1988), s. 36.

Natomiast całe spektrum biblioteki zwizualizowano następująco:



Rys. 3 – spektrum zawartości biblioteki – za Kahn, Cerf (1988), s. 15.

Jak widać cały system jest bardzo szeroko rozumiany. Włączone zostały do niego poza zawartością (dane komputerowe, wersje drukowane itp.) czy softwarem także interakcje człowiek-maszyna. Używany w sprawozdaniu termin 'knowbot' odnosił się do aktywnych, inteligentnych programów zdolnych do wymiany danych między różnymi systemami, działających zgodnie z wolą użytkownika, mogącego określić plany poszukiwań, kryteria, formaty wyjściowe itp. 'Ingerencja' użytkownika ostatecznego w cały system przez 'biblioteczny system osobisty' była więc jednym z ważnych elementów tworzących cały twór.

W *Wikipedii* zaznaczono także, iż spopularyzowanie pojęcia 'biblioteki cyfrowej' dokonało się przez Inicjatywę na rzecz Bibliotek Cyfrowych (Digital Libraries Initiative) podjętą przez NSF/DARPA/NASA w 1994 r. [zob. Edward A. Fox *The Digital Libraries Initiative – Update and Discussion,* 1999]. Jako protoplastę biblioteki cyfrowej wymieniono Vannevara Busha z jego artykułem *As We May Think* (1945 r.), w którym wizja nowego systemu nie jest oparta tylko na podstawach technologicznych, ale także na doświadczeniach użytkownika [za: Candela, Castelli, Pagano 2011, s. 1-30]. Wskazano też, iż termin biblioteka cyfrowa był początkowo zamiennie używany z pojęciem biblioteki wirtualnej, choć ta druga rozumiana była przede wszystkim jako system rozproszony. Czasami też starano się wskazywać różnice między zasobami zdigitalizowanymi, a od początku cyfrowymi – born-digital. Pojawiał się też w piśmiennictwie termin biblioteki hybrydowej jako biblioteki, która posiada zarówno fizyczne (drukowane, analogowe), jak i cyfrowe zbiory (np. American Memory w Bibliotece Kongresu).

Podkreślono również, iż ważne biblioteki cyfrowe służą również jako długoterminowe archiwa: np. Eprint arXiv i Internet Archive.

*

W słowniku *Dictionary for library and information science* (**2004**) opracowanym przez Joan M. Reitz pojawiły się dwa terminy: 'biblioteka cyfrowa' (digital library – d-lib) oraz 'wirtualna' (virtual library) z zaznaczeniem, iż pierwszy jest bardziej adekwatny, gdyż wirtualność sugeruje inność od realności, a przecież zawartość zawsze odnosi się do tego, co realne. Opisowo przedstawiając biblioteki autorka zaznaczyła, iż w cyfrowej (bibliotece) zawartość w znaczącym procencie jest czytelna dla komputerów – ma odpowiednie formaty. Sam proces digitalizacji natomiast zaczyna się od katalogów, serwisów abstraktowych, czasopism, aż po kolejne źródła informacji. Biblioteka wirtualna zaś jest 'biblioteką bez ścian', w której zawartość zawsze dostępna jest przez sieć.

*

W późniejszym, z **2007** r., leksykonie *Lexikon Buch – Bibliothek – neue Medien* autorów: Dietmar Strauch, Margarete Rehm pojawiły się trzy terminy: 'biblioteka cyfrowa' ('digitale Bibliothek'), 'hybrydowa' ('hybride Bibliothek') oraz 'wirtualna' ('virtuelle Bibliothek'). Biblioteka cyfrowa jest obiektem/systemem informacyjnym, zawierającym treści będące elektronicznymi odpowiednikami źródeł materialnych, jak i unikatowymi (istniejącymi tylko w formie cyfrowej). Biblioteka hybrydowa łączy oczywiście cechy biblioteki tradycyjnej z cyfrową. Natomiast wirtualna daje użytkownikowi możliwość dotarcia za pośrednictwem komputera do zasobów elektronicznych, także multimedialnych.

*

Polski słownik z **2002** r. opracowany przez Bożennę Bojar *Słownik encyklopedyczny informacji, języków i systemów informacyjno-wyszukiwawczych* [s. 34] złączył terminy biblioteki cyfrowej i wirtualnej oraz ulokował je w indeksie rzeczowym w dziale system informacyjno-wyszukiwawczy (I.2. typologia). Termin przyjęty – biblioteka wirtualna – przedstawiony został w dwóch znaczeniach: 1) zbioru dokumentów elektronicznych gromadzonych, opracowanych i udostępnianych użytkownikom zgodnie z regułami stosowanymi w bibliotekarstwie – a więc obiektu oraz 2) instytucji zajmującej się obsługą biblioteki wirtualnej, która ewentualnie może zajmować się również ochroną takiego zbioru, chociaż powinna to robić wyspecjalizowana jednostka.

*

Na koniec tego podrozdziału dodać jeszcze należy, iż w niektórych encyklopediach w ogóle nie występują terminy 'biblioteka cyfrowa', 'biblioteka wirtualna'. Encyklopedia PWN dostępna online po wprowadzeniu hasła 'biblioteka cyfrowa' odsyła do opisu Internetu [http://encyklopedia.pwn.pl/haslo/3915155/internet.html], natomiast Portal Wiedzy ONET informuje jedynie, iż nie znaleziono takich pojęć.

3. Biblioteka elektroniczna, cyfrowa, wirtualna

a. piśmiennictwo światowe

Wcześniejsze od wymienionych w encyklopediach publikacji Daniela Greensteina i Suzanne E. Thorin prace, takie jak np. *Analytical Review of the Library of the Future* Karen Drabenstott (**1993**) przyniosły definicje oraz różne przedstawienia terminów związanych z pojęciem biblioteki cyfrowej ('digital library'). Praca Karen Drabenstott jest istotna, gdyż autorka przeprowadziła analizę piśmiennictwa oraz wyodrębniła cytaty z wielu artykułów, stąd ta publikacja omówiona zostanie szerzej. Termin 'biblioteka cyfrowa' został przez Drabenstott zaprezentowany jako jeden z najnowszych w długiej serii nazw dla koncepcji systemów informacyjnych, które powstawały w umysłach twórców, prawie tak długo, jak długo trwa rozwój maszyn cyfrowych. Jako protoplaści przedstawieni zostali: Vannevar Bush (1945) i jego 'memex', Joseph Carl Robnett Licklider (1965) z frazą 'biblioteki przyszłości' w odniesieniu do bibliotek w pełni skomputeryzowanych, F. Wilfrid Lancaster (1978) z 'elektroniczną biblioteką' oraz Ted Nelson (1974), twórca hipertekstu i hiperprzestrzeni. Inne pojęcia, pokazujące rozwój bibliotek, a może nawet 'tworów, które je zastąpią', to 'wirtualna biblioteka', 'biblioteka bez ścian', 'biblioteka bioniczna (bionic)'.

'Biblioteka przyszłości' opisana została jako sieć systemów wiedzy, w której ludzie i maszyny będą ze sobą współpracować [za: Feigenbaum 1989, s. 122]. Miała też być siecią bibliotekarzy – ogólnych i specjalistów, którzy byliby dostępni online, przy czym każdy z nich był specjalistą w wyszukiwaniu informacji. Wszyscy (także użytkownicy) mieli być połączeni z bazami danych, obejmującymi: pliki multimedialne, pełne teksty, cyfrowe obrazy, pliki audio, audio-wizualne, animacje itp. Usługi dostępne sieciowo opierać się miały na pośrednictwie kamer, syntezatorów mowy oraz specjalnego interfejsu (za: Eagle 1992, 99).

'Elektroniczna biblioteka' [za Kibby, Evans 1989] miała nie być jednolitą jednostką, w której są przechowywanie źródła informacji, lecz 'miejscem', w którym zbiory poprzez różnego rodzaju usługi udostępniane są w sieci, wykraczającej poza sferę poszczególnych uczelni, ośrodków badawczych itp. Była więc traktowana jako nowa strategia dostarczania informacji, w której teksty dokumentów są w pełni dostępne w Internecie. Biblioteką można byłoby sterować zdalnie z dowolnego miejsca przy użyciu osobistych lub przenośnych komputerów. Idea powszechnej dostępności reprezentuje w tym myśleniu najbardziej rewolucyjny aspekt elektronicznej biblioteki. Biblioteka sama ma nieskończoną przestrzeń i nie ma ograniczeń co do liczby dokumentów, które może zawierać, zwłaszcza, że poszczególne kolekcje mogą być ze sobą połączone [za: Landoni, Catenazzi i Gibb 1993, s. 176].

'Biblioteka wirtualna (virtual library)' miała być natomiast systemem, dzięki któremu użytkownik może łączyć się z biblioteką, jej bazami danych, katalogiem online oraz wszelkimi innymi źródłami danych – nie tylko odpowiednikami źródeł drukowanych [za: Saunders 1992b]. Przez ten system, ciągle rozwijany przez biblioteki, użytkownik mógł docierać do materiałów przechowywanych, ale i zdobywanych dla

użytkownika [za: Lynch 1991, s. 77]. Tak więc użytkownik miał możliwość wpływu na bibliotekę, która stawała się efektem synergii stworzonej przez technologiczne połączenie zasobów bardzo wielu bibliotek i ośrodków informacji [za: Gapen 1993, s. 1].

'Biblioteka cyfrowa' przedstawiona została przez cztery podstawowe cechy: 1) zmniejszenie ograniczeń co do czasu i miejsca, 2) umożliwienie tworzenia i stosowania nowych, bardziej dynamicznych, zintegrowanych formatów przedstawiania danych, informacji i wiedzy, 3) wspierania nowych form pracy grupowej w zakresie tworzenia i wykorzystania informacji, integracji praktyków, 4) przystosowania i personalizacji informacji, w tym pomocy w zarządzaniu informacjami, a zwłaszcza jej nadmiarem [za: Atkins 1993, s. 2].

Najmniej znany w Polsce termin 'biblioteka bioniczna (bionic library)' został zaproponowany przez Harolda Billingsa jako połączenie tradycyjnego organizmu biblioteki z elektronicznym, przez co następowało wzmocnienie instytucji [Billings 1991b, s. 4].

Poza zebraniem cytatów dotyczących terminologii, przyszłości, modeli bibliotek cyfrowych, jak i motywacji użytkowników, bibliotekarzy, wydawców Karen Drabenstott przedstawiła także w swojej książce pokrótce projekty bibliotek cyfrowych: ADONIS, American Memory Project, Ariel, CORE, DOCDEL, DocuTech, Gateway to Information, IO+ Extended OPAC, Mann Library Gateway, MemRI, MiteyBook, NCSU DDTP, Network Fax Project, OJCCT, Primis, Project Athena, Project Janus, Project Mercury, Project QUARTET, Red Sage, RightPages™, SuperBook, TULIP.

Publikacja ta oparta na badaniach literatury tematu miała swoje liczne odniesienia w kolejnych pracach. Chociaż pojawiały się także prace proponujące nowe pojęcia jak np. ta, przygotowana przez Monicę Landoni, Nadię Catenazzi i Forbesa Gibba *Hyper-books and visual-books in an electronic library* (**1993**). Poza hiper-książkami i książkami wizualnymi autorzy przedstawili wirtualno-elektroniczną bibliotekę (Virtual Electronic Library).

Odniesienie do pracy Karen Drabenstott oraz terminy 'biblioteka elektroniczna', 'biblioteka cyfrowa', czy 'biblioteka wirtualna' ponownie pojawiły się np. w artykule Philipa Barkera *Electronic Libraries – visions of the future* (**1994**), skupionym przede wszystkim na przedstawieniu nowych technologii. Omówiono w nim podstawowe funkcje systemów bibliotecznych oraz ich roli jako elementów transferu wiedzy w nowoczesnych społeczeństwach. Systemy podzielone zostały na: polimedialne, elektroniczne, cyfrowe i wirtualne, przy czym kolejność ich wymienienia wskazywała na stopień ich rozwoju.

*

Nowy termin został także zaproponowany przez Nadię Catenazzi oraz Lorenzzo Sommaruga pt. *Hyper-lib: a formal model for an electronic library based on hyper-books* (**1995**). Termin hiper-biblioteki opiera się na idei przeniesienia przestrzennego środowiska rzeczywistej biblioteki do systemu elektronicznego. Biblioteka, jej sale, półki, książki, czy usługi są odtwarzane poprzez symulację w trójwymiarowym środowisku, w którym użytkownik może komunikować się z kolekcją dokumentów, tak jakby

przebywał w fizycznej bibliotece. Oprócz cech fizycznych biblioteki, hiper-biblioteka wprowadza kilka nowych funkcji tzn. szybszy i łatwiejszy dostęp do źródeł, oraz umożliwienie użytkownikowi zmiany stanu systemu.

*

Podobne analizy zmian funkcji systemów bibliotecznych oraz transformacji biblioteki tradycyjnej w wirtualną podjęli także Edward A. Fox, Robert M. Akscyn, Richard K.Furuta, oraz John J. Leggett w pracy *Digital Libraries* (**1995**). To co ważne w tej pracy to argumentacja za stosowaniem terminu biblioteka cyfrowa, a nie elektroniczna, gdyż ten termin bardziej wskazuje na zasoby cyfrowe – digitalne (digital), które budują taką kolekcję. Poza tym autorzy ci zaakcentowali różne role 'głównych graczy' bibliotek cyfrowych: bibliotekarzy, informatyków i użytkowników. Do bibliotekarzy należy wykonywanie funkcji przynależnych do tradycyjnych bibliotek w nowy sposób, do informatyków stworzenie systemu informatycznego z zasobem rozproszonych różnorodnych w typie informacji oraz usług, do użytkowników końcowych wykorzystywanie bibliotek cyfrowych oraz 'wymuszanie' poprawy wydajności, organizacji, funkcjonalności i użyteczności. Zadania te powodują, iż 'główni gracze' nie koniecznie w ten sam sposób widzą schematy informacyjne, a co za tym idzie mają różne 'modele zewnętrzne' bibliotek cyfrowych.

*

W kolejnej publikacji z **1995** r. Petera J. Nürnberga, Richarda Furuty, Johna J. Leggetta, Catherine C. Marshall i Franka M. Shipmana *Digital libraries: issues and architectures* autorzy starali się rozpatrzyć poszczególne elementy oraz architekturę bibliotek cyfrowych. Ich praca została rozbudowana w analizie przygotowanej dla IFLA, stąd szczegółowo omówiona została później. W tym miejscu zostaną zaprezentowane jedynie najważniejsze elementy, budujące bibliotekę cyfrową, czyli jej taksonomiczne przedstawienie oraz konieczne kwestie do poruszenia przy przygotowywaniu danych, metadanych oraz procesów (tab. 1 i 2).

Tabela 1. Taksonomia elementów biblioteki cyfrowej

	Dane	***Metadane***	***Procesy***
Transpozycje jednostek tradycyjnej biblioteki	Książki Czasopisma Filmy	Statyczny indeks Klasyfikacje Układ przestrzenny	Gromadzenie danych Polecanie (sugerowanie) źródeł Pomoc przy źródłach lokalnych
Jednostki biblioteki cyfrowej	Hiperpowieści Wizualizacje naukowe Programy komputerowe	Dynamiczny indeks Personalizowana struktura Adnotacje	Przeszukiwanie pełno tekstowe Personalizowane prezentowanie danych Pobieranie danych

Tabela 2. Kwestie do poruszenia przy taksonomii elementów biblioteki cyfrowej

	Dane	*Metadane*	*Procesy*
Transpozycje jednostek tradycyjnej biblioteki	Jak przenosić?	Jak przenosić metadane zależne fizycznie od informacji?	Jak dostarczyć narzędzia do zaangażowania ludzi w ten proces?
Jednostki biblioteki cyfrowej	Jak dostosowywać się do szybkiej i ciągłej zmiany nowych typów danych?	Jak zabezpieczać spójność oddzielnie utrzymywanych metadanych?	Jak dystrybuować?

*

W kolejnej pracy *What is a Digital Library? Definitions, Content, and Issues,* zaprezentowanej na konferencji *KOLISS DL ,96: International Conference on Digital Libraries and Information Services for the 21st Century, September 10-13,* ***1996****, Seoul, Korea* Stephen P. Harter także odniósł się do pojęć przedstawionych przez Karen Drabenstott oraz do prac Francisa L. Miksa i Philipa Doty (1994) oraz Davida M. Levy i Catherine C. Marshall (1994). Podobnie ulokował powstanie bibliotek cyfrowych w ewolucji systemów komputerowych.

Przedstawił także typologię źródeł informacji dostępnych w bibliotekach cyfrowych:

1) 'tradycyjne' zasoby biblioteczne:
   a) metadane: indeksy, katalogi itp.
   b) katalogi katalogów, biblioteki bibliotek
   c) pełne teksty:
      a. książek, artykułów, poezji, raportów technicznych, materiałów dydydaktycznych itp.
   d) informacje odsyłające (słowniki, tezaurusy itp.)
   e) kolekcje multimedialne: zdjęć, sztuki, muzyki, archiwa nagrań dźwiękowych, pliki wideo, zapis śpiewu ptaków itp.
   f) dane liczbowe: spisy, dane na temat dotyczące pogody, środowiska
2) 'nietradycyjne' zasoby biblioteczne dla komunikacji naukowej
   g) archiwa pre-printów
   h) grupy dyskusyjne, listy, komentarze, zapisy MUD-ów itp.
   i) kolekcje-mailowe
3) zasoby, które nie są materiałem informacyjnym (dla tradycyjnej biblioteki)
   j) gry
   k) komercyjne witryny
   l) materiały dla tzw. grup interesu
   m) dokumenty osobiste, zbiory, eseje, strony domowe itp.
   n) tzw. reszta Internetu
4) serwisy biblioteczne
   o) serwisy typu „pytanie – odpowiedź"
   p) analizy statystyczne
   q) utworzone niestandardowe mapy

r) usługi pomocowe przy narzędziach wyszukiwawczych
s) serwis doradczy dla czytelników
t) serwis zapewniający dostęp i pomoc przy korzystaniu z usług komercyjnych.

Natomiast cechy biblioteki cyfrowej zostały przedstawione z trzech punktów widzenia, co w formie tabeli prezentuje się następująco:

Tabela 3. Trzy punkty widzenia na bibliotekę cyfrową

Wąskie pole widzenia (oparte na wizji tradycyjnej biblioteki)	**Szersze spojrzenie (wypośrodkowane między skrajnościami)**	**Najszersze spojrzenie (luźno oparte na Internecie)**
obiektami są zasoby informacyjne	większość obiektów to zasoby informacyjne	obiektem może być cokolwiek (wszystko)
obiekty są wybierane na podstawie ich jakości	niektóre z obiektów są wybierane na podstawie ich jakości	bez barier jakości, bez barier dostępu
obiekty znajdują się w jakimś fizycznym miejsce	obiekty znajdują się w tzw. logicznej przestrzeni (mogą być dystrybuowane)	obiekty nie są zlokalizowane ani w fizycznym, ani w logicznym miejscu
obiekty są zorganizowane		brak organizacji
obiekty poddawane są jakimś organom kontroli	niektóre obiekty poddawane są jakimś organom kontroli	bez kontroli
obiekty są stałe (nie należy ich zmieniać)	obiekty zmieniają się w jakiś zorganizowany, kontrolowany sposób	obiekty mogą się zmieniać ciągle
obiekty są stałe (nie znikają z zasobów)	zanikanie źródeł jest procesem kontrolowanym	obiekty mogą zniknąć w każdym momencie
autorstwo jest ważnym pojęciem	pojęcie autorstwa jest słabsze	brak pojęcia autorstwa
dostęp do obiektów ogranicza się do określonych kategorii użytkowników	dostęp do niektórych obiektów ogranicza się do określonych kategorii użytkowników	dostęp do wszystkiego dla każdego
istnieją usługi, które są oferowane użytkownikom, takie jak np. pomoc		usługi są wykonywane wyłącznie przez program komputerowy (AI)
można dotrzeć do człowieka specjalisty (called librarians)		nie ma bibliotekarzy
istnieją dobrze zdefiniowane grupy użytkowników	dla pewnych klas obiektów istnieją dobrze zdefiniowane grupy użytkowników	nie ma zdefiniowanych klas użytkowników lub alternatywnie jest ich zbyt wiele

Stephen P. Harter zadał też ciąg pytań, których część znalazła już swoje odpowiedzi (np. co robią bibliotekarze cyfrowi), ale niektóre do dziś są aktualne np. w kontekście zapewniania jakości, praw dostępu do materiałów itp. Pytania te, określające problematykę bibliotek cyfrowych, ich tworzenia oraz utrzymywania, brzmiały:

1) Jak możemy ustalić i kontrolować aktualność, dokładność, integralność źródeł informacji (problem jakości)?

2) Jak możemy zachować dane i ich integralność intelektualną (problem kontroli autorytatywności)?
3) Co można zrobić, aby zapewnić dostęp do źródeł informacyjnych (problem organizacyjny)?
4) Jak możemy rozpoznać różne wersje tego samego źródła (problem płynności, ale i redundancji informacji)?
5) Jak można rozwiązać problem nieustalonych źródeł informacyjnych (problem zachowania)?
6) Jak możemy zachować pojęcie autorstwa?
7) Jak należy przestrzegać praw autorskich (problem prawny)?
8) Czy dostęp do niektórych źródeł informacji powinien być ograniczony dla niektórych kategorii użytkowników (problem polityki)?
9) Jakie usługi i serwisy powinny być oferowane przez biblioteki cyfrowe?
10) Czy biblioteki cyfrowe powinny być włączane w struktury tradycyjnych bibliotek, a jeśli tak, to jak można to osiągnąć?
11) Czy biblioteki cyfrowej mają bibliotekarzy, a jeśli tak, to co oni robią?
12) Czy biblioteki cyfrowe mają dobrze zdefiniowane klasy użytkowników?

*

Inne prace, jak np. *The digital library: myths and challenges* autorstwa Terry'ego Kuny i Gary'ego Clevelanda (**1998**) w dość prowokacyjny sposób przedstawiały, czym są, a właściwie, czym nie są biblioteki cyfrowe, po to by wskazać, jakim wyzwaniem, ale z drugiej strony i dobrym wykorzystaniem pieniędzy są projekty tworzenia takich systemów. 'Biblioteki bez ścian' mają bowiem nowe ograniczenia, nie tylko architektoniczne i budowlane, ale także technologiczne, prawne, administracyjne itp. Mity, z którymi 'walczył' wręcz ten tekst, próbując zakreślić ramy pojęciowe to mity: 1) internet jest biblioteką cyfrową, 2) będzie istnieć dostępna dla wszystkich biblioteka cyfrowa, otwarta, bezpłatna, 3) biblioteki cyfrowe szybko zapewnią bardziej sprawiedliwy dostęp z każdego miejsca i o każdej porze, 4) biblioteki cyfrowe będą tańsze niż realne biblioteki. Skupiono się więc na takich istotnych cechach biblioteki cyfrowej jak: organizacja, celowość tworzenia kolekcji, przeszukiwalność, jakość zasobów, także w sensie ich wiarygodności, funkcjonalność, ale i różnorodność, czy praco- i czasochłonność projektów wymagających sporych nakładów finansowych.

*

Wymieniony wcześniej Gary Cleveland przygotował również opracowanie dla IFLA – International Federation of Library Associations and Institutions *Digital libraries: definitions, issues and challenges* (**1998**). Po raz kolejny zostało przedstawione zamieszanie terminologiczne: wokół pojęć biblioteki elektronicznej, wirtualnej biblioteki, biblioteki bez ścian, biblioteki cyfrowej. Cleveland w wielu miejscach odniósł się do pracy Petera J. Nürnberga, Richarda Furuty, Johna J. Leggetta, Catherine C. Marshall i Franka M. Shipmana *Digital libraries: issues and architectures*, o której wspomniano wcześniej.

Autor napisał, iż biblioteka cyfrowa z punktu widzenia: 1) wyszukiwania informacji jest dużą bazą danych, przygotowaną dla osób pracujących na hipertekstowej technologii, 2) dla osób pracujących w szerokim obszarze dostarczania informacji np. dla badaczy z zakresu bibliotekoznawstwa i informacji naukowej jest kolejnym krokiem w automatyzacji bibliotek, która rozpoczęła się w latach siedemdziesiątych, 3) dla informatyków i programistów są to także zbiory algorytmów lub elementy programów komputerowych, 4) dla dostawców baz danych lub dostawców dokumentów są usługami, 5) dla dużych korporacji są systemami zarządzania dokumentami, 6) dla wydawców mogą być wersjami online katalogu, 7) dla wielu ludzi to po prostu World Wide Web, gdyż to sieć gromadzi tysiące dokumentów.

W związku z takim wielorakim rozumieniem pojęcia, przyjęto, iż punktem wyjścia będzie założenie, że funkcje bibliotek cyfrowych są analogiczne do misji tradycyjnych bibliotek: gromadzenie i zarządzanie informacjami, opracowanie danych, analizowanie tematów, tworzenie indeksów, zapewnienie dostępu użytkownikom, przechowywanie i zabezpieczanie materiałów itp. Elementem zmiany są: nowe nośniki, cyfrowe formaty, oprogramowanie itp. Natomiast nadal celem ich istnienia jest służenie społeczeństwu.

Biblioteki cyfrowe stały się więc organizacjami, które dostarczają zasobów, mają własnych wyspecjalizowanych pracowników, m.in. selekcjonujących zasoby; posiadającymi swoją strukturę oraz skupionymi na otwieraniu nowych możliwości dostępu do wiedzy, rozpowszechnianiu jej, przy jednoczesnym zachowaniu integralności i zapewnianiu trwałości w czasie zbiorów dzieł cyfrowych, tak aby były łatwo i ekonomicznie wykorzystywane.

Architektura takich systemów musi więc zawierać takie elementy, jak: szybkie sieci lokalne oraz szybkie połączenia z Internetem, relacyjne bazy danych, które obsługują różne formaty, interfejsy zapewniające wyszukiwanie każdego tekstu oraz dostęp do indeksów, serwery WWW i serwery FTP, elektroniczne funkcje zarządzania dokumentami, które pomogą w całościowym zarządzaniu zasobami cyfrowymi. Gary Cleveland podkreślił przy tym fakt, iż nie będzie to monolityczny system, lecz zbiór odrębnych systemów i zasobów połączonych przez sieć, zintegrowanych w jednym interfejsie. Źródła informacji zbierane w bibliotece cyfrowej obejmować mają zarówno informacje faktograficzne, pełnotekstowe (w rozumieniu dane źródłowe – a więc także multimedialne), jak i odsyłające. Gromadzone są zarówno dzieła zdigitalizowane, digital-born, jak i różne formy dostępu do zewnętrznych materiałów (np. innych baz danych). Opisywane powinny być one meta-danymi np. w standardzie Dublin Core. Co ważne: zaznaczone powinny być nie tylko same identyfikatory obiektów, które pozwolą na zidentyfikowanie ich, zarządzanie prawami autorskimi oraz w jakiś sposób wpłyną na ich trwałość (istnienie w sieci), ale także określone być powinny połączenia między obiektami.

*

Także inna instytucja – amerykańska federacja: Digital Library Federation (DLF), która związana jest z organizacją non-profit: Council on Library and Information Resources – CLIR (założone w 1997, jako kontynuator The Council on Library Resour-

ces – CLR – istniejącego od 1956 r.) przedstawiła w **1998** r. swoją definicję biblioteki cyfrowej. Mowa jest w niej o organizacji dysponującej zasobami i wyspecjalizowaną grupą pracowników, do których należy wybór, tworzenie struktury, intelektualnych narzędzi dostępu i reprezentacji cyfrowych obiektów, a ponadto ich rozpowszechnianie, przy równoczesnym dbaniu o zachowanie integralności i trwałości kolekcji, które mają być łatwo i po niewielkich kosztach dostępne dla określonej społeczności czy grup społecznych. Definicji ta traktuje biblioteki cyfrowe jako kompleksowy system, składający się zasobów informacyjnych oraz ludzi (twórców, użytkowników), a także takich wyszczególnionych elementów, jak: języki informacyjne, sposoby prezentacji danych, czy narzędzia dostępu.

Podkreślić jeszcze należy, iż komisja CLIR-u – Commission on Preservation and Access już pod koniec lat osiemdziesiątych, aż do połowy lat dziewięćdziesiątych rozpoznawała technologię cyfrową, współpracując z Library of Congress i National Archives. Na przykład w 1994 r. wezwano do tworzenia strategii planowania rozwoju bibliotek cyfrowych, a także starano się angażować władze lokalne, przedstawiając zalety tworzenia lokalnych bibliotek cyfrowych.

*

W tym samym roku **1998** Stephen M. Griffin, Gail McMillan oraz Nan Seamans przygotowali *Digital Library Definition for DLI2,* tymczasową definicję bibliotek cyfrowych dla National Digital Library Program – USA (NDLP). Zapisano w niej, iż 'biblioteka cyfrowa' nie jest jedynie odpowiednikiem cyfrowej kolekcji, do której mamy dostęp za pomocą narzędzi zarządzania informacją. Jest to także szereg działań, które umożliwiają połączenie zbiorów z usługami przy pomocy różnych osób. Mamy więc pełny cykl życia systemu informacyjnego: tworzenie, upowszechnianie, wykorzystanie oraz ochronę danych, informacji i wiedzy. Dodatkowo, podkreślono fakt, iż biblioteki cyfrowe tworzą wartość dodaną, pozwalają zaoszczędzić czas dostępu do danych, dają poczucie bliskości zasobów informacyjnych, zapewniając jakość tych zasobów. System może być indywidualnie dostosowany i ostatecznie łatwy w użyciu. Wspomoże też nauczanie w klasach szkolnych oraz e-learning.

*

O mitach związanych z terminem 'biblioteka cyfrowa' napisał poza Terry'm Kuny i Gary'm Cleveland'em także Donald J. Waters – w artykule *What are digital liraries?* (**1998**), walcząc z tezą, iż cały Internet jest biblioteką. I znowu biblioteka cyfrowa była organizacją, dostarczającą zasobów, przez wyspecjalizowanych pracowników, ze strukturą, ofertą dostępu do informacji oraz z celami, przedstawionymi już przez Gary'ego Clevelanda (*Digital libraries: definitions, issues and challenges*). Przy czym Waters zapisał, iż biblioteki cyfrowe nie muszą być zorganizowane na wzór tradycyjnych bibliotek (lub nawet w ramach tradycyjnych bibliotek). Choć zasoby bibliotek cyfrowych mają pełnić funkcje podobne do tych z tradycyjnych bibliotek, są one inne: wymagają bowiem umiejętności komputerowych: posługiwania się siecią i specjalistycznym oprogramowaniem.

*

Ramy pojęciowe zostały także zakreślone przez Iana Rowlandsa i Davida Bawdena w pracy przygotowanej w oparciu o raport przygotowany dla British Library Research and Innovation Centre *Digital libraries: a conceptual framework* (**1999**). To co ważne w tej pracy, to podkreślenie znaczenia bibliotek cyfrowych dla społeczeństwa, czy organizacji polityki informacyjnej.

*

W **1999** r. pojawił sie także tekst dwóch autorów: Gobindy G. Chowdhury, Sudatta Chowdhury *Progress in documentation: digital library research: major issues and trends*, często cytowany w późniejszych publikacjach. Jego popularność spowodowała, iż rozbudowana wersja opisu wielu istniejących bibliotek cyfrowych wraz z teorią ich tworzenia przedstawiona została w formie książki *Introduction to digital libraries* w 2003 r. (wersja elektroniczna dostępna od 2010 r.). Biblioteka cyfrowa rozumiana była przez tych badaczy jako kolekcja cyfrowa, zorganizowana, dostępna z jednego miejsca. Omówione przez nich zostały takie zagadnienia, jak: charakter i rozwój bibliotek cyfrowych, infrastruktura techniczna, cyfrowe zbiory biblioteczne, digitalizacja i zarządzanie dokumentami, organizacja cyfrowych informacji, wyszukiwanie informacji, interfejsy, użytkownicy, nastawienie user-centered, scalanie danych, archiwizacja i konserwacja, usługi proponowane w systemie, a także ocena bibliotek cyfrowych i ich wpływ na sam zawód pracownika informacji.

*

W przełomowym dla wieków **2000** r. pojawił się artykuł Colleen Cool i Marthy E. Williams *A typology of digital libraries and their user communities,* w którym przedstawiono różne projekty, inicjatywy oraz usługi opatrzone 'metką' biblioteka cyfrowa. Pokazano ich cele oraz zaplanowane przez twórców wizje społeczności użytkowników końcowych. Dokonano także porównania z typami tradycyjnych bibliotek. Najważniejsze było jednak wskazanie, iż najszersze zastosowanie biblioteki cyfrowe będą mieć w odniesieniu do zbiorów specjalnych.

*

W latach dziewięćdziesiątych XX w. pojawił się także cykl artykułów (chociaż nie było ich dużo), dotyczący użytkowników bibliotek cyfrowych, a zwłaszcza sposobów wyszukiwania informacji oraz tworzenia funkcjonalnych interfejsów. Część z nich odnosiła się do właśnie tworzonych bibliotek cyfrowych. W **1995** r. Ian H. Witten, Craig G. Nevill-Manning i Sally Jo Cunningham w raporcie *Building a public digital library based on full-text retrieval* opisali architekturę, indeksowanie, proces zbierania i eksploatacji zbiorów, a także interfejs wyszukiwania prototypu publicznie dostępnej biblioteki cyfrowej. Nie przedstawili ścisłej definicji biblioteki cyfrowej, ale opisali wyzwania, przed jakimi stają jej twórcy: 1) 'materiał podstawowy', który ma stanowić trzon kolekcji, dostępnej w domenie publicznej, 2) selekcję (prawa, jakość), 3) formaty danych, 4) katalogowanie zbiorów, 5) wyszukiwanie informacji. Zagadnienia te są aktualne do dziś.

W **1998** r. Steve Jones, Sally Jo Cunningham i Rodger McNab przeanalizowali dzienniki transakcji dużej kolekcji pełnotekstowych dokumentów oraz omówili spo-

soby akceptacji domyślnych ustawień systemów, przygotowanych zapytań złożonych oraz możliwości przeglądania wyników. Jako jedni z pierwszych spojrzeli więc na system biblioteki cyfrowej od strony użytkownika.

O projektowaniu interfejsów i narzędzi dla użytkowników dla Biblioteki Kongresu pisali Gary Marchionini, Catherine Plaisant i Anita Komlodi (**1998**), a o New Zealand Digital Library ponownie Ian H. Witten, Craig G. Nevill-Manning, Rodger J. McNab i Sally Jo Cunningham (**1998**). Interakcje użytkowników z bibliotekami cyfrowymi, w celu zidentyfikowania kluczowych zadań oraz opracowania zasad projektowania nowej generacji bibliotek cyfrowych, analizowali także Sandra D. Payette i Oya Y. Rieger (**1998**).

W cyklu prac tematycznie związanych z klientami bibliotek cyfrowych wymienić też można np. pracę Steve'a Mitchella (**1999**), z apelem o definiowanie użytkowników i ich potrzeb, celów, zachowań, oczekiwań oraz określanie tradycji, z jakimi są oni związani.

Kolejne artykuły ponownie odnosiły się do konkretnych projektów. I tak np. cykl projektowania i wdrażania Alexandria Digital Library, wraz z analizą odbiorców końcowych opisany został przez grupę autorów: Linda L. Hill, Larry Carver, Mary Larsgaard, Ron Dolin, Terence R. Smith, James Frew, Mary-Anna Rae (**2000**). Do analiz interfejsów powróciła Michelle Q. Wang Baldonado (**2000**), który przedstawiła 'zorientowany na użytkownika interfejs' (user-centered interface). Kolejna autorka: Soyeon Park (**2000**) analizowała interakcje użytkowników z heterogenicznymi i rozproszonymi źródłami informacji w bibliotekach cyfrowych.

Trzeba jednak zaznaczyć, iż nasilenie analiz grup użytkowników, typologia i próba zakreślenia zakresu terminu 'użytkownik' to początek wieku XXI – zob. artykuły Dietmara Wolframa, Hong (Iris) Xie czy Ying Zang, przedstawione później oraz pracę o użytkowniku autorstwa Jolanty Mazurek, zamieszczoną w tej książce.

b. Polska

Jednymi z pierwszych polskich publikacji poświęconych 'bibliotece cyfrowej' były prace Ewy Chmielewskiej-Gorczycy *Biblioteka wirtualna – wizja, czy rzeczywistość* (**1996**a) oraz *Ku bibliotece wirtualnej* (1996b). Autorka przedstawiła 'stan zamieszania definicyjnego', w którym jako wyrażenia synonimiczne dla biblioteki wirtualnej podawane są takie określenia jak 'biblioteka bez ścian', 'biblioteka logiczna' lub 'biblioteka cyfrowa'. W pracach tychj omówiono różnice pomiędzy tymi pojęciami, wyznaczając ich stopniowanie. Najważniejsza była biblioteka wirtualna. Jej koncepcja opierała się *na trzech wyraźnych elementach: biblioteki elektronicznej* [lub cyfrowej], *sieci telekomunikacyjnej oraz odpowiedniego interfejsu użytkownika.* Biblioteka cyfrowa lub elektroniczna była więc *niezbędnym warunkiem zaistnienia biblioteki wirtualnej*, ale nie była z nią tożsama, gdyż te terminy tożsame były z 'biblioteką zautomatyzowaną', posiadającą katalog OPAC oraz różne moduły funkcjonalne (gromadzenie, opracowanie, wypożyczanie itp.). Biblioteka określana mianem cyfrowej (digital library) *wymagała już pośrednictwa sieci telekomunikacyjnej oraz zbiorów w postaci maszynowej,*

oraz usług *elektronicznego dostarczania dokumentów, w tym i multimedialnych.* Biblioteka wirtualna natomiast umożliwiała korzystanie z zewnętrznych serwisów i źródeł informacji, otwartych i komercyjnych, za pośrednictwem jednego systemu (interfejsu) [s. 123]. Jeśli chodzi o rozwój bibliotek, to autorka wyodrębniła pięć stopni: 1) biblioteka tradycyjna, 2) polimedialna, 3) elektroniczna, 4) cyfrowa, 5) wirtualna.

*

Podobny rozwój bibliotek przedstawiony został przez Juranda B. Czermińskiego (**1997**) w referacie *Od biblioteki elektronicznej do biblioteki wirtualnej* przedstawionym na konferencji *Elektroniczna przyszłość bibliotek akademickich.* Trzy poziomy automatyzacji biblioteki to: biblioteka zautomatyzowana, elektroniczna i cyfrowa. Ta ostatnia została przedstawiona jako jedyna struktura, *która może być w pełni zintegrowana w jeden system z bibliotekami tego samego typu.* System, który może być konfigurowany i zarządzany z dowolnego miejsca. Biblioteka cyfrowa jest *też szczytowym etapem rozwoju każdej biblioteki tradycyjnej.* Natomiast biblioteka wirtualna nie jest kolejnym krokiem w rozwoju, *lecz sposobem na wyszukiwanie informacji z systemu takich bibliotek odpowiednio uformowanego pod względem zastosowanych norm* [s. 161]. Jest także spełnieniem marzeń o ogólnym *nieograniczonym* dostępie do informacji [s. 163].

*

Kolejny tekst *Nasza przyszłość: biblioteka elektroniczna czy biblioteka cyfrowa?* (**1997**) był autorstwa Anny Machalskiej-Garbacz, która opisała rozwój koncepcji bibliotek cyfrowych (1993 r. – rząd Stanów Zjednoczonych – nadanie tematowi rangi programu rządowego; październik 1994 r. Biblioteka Kongresu – National Digital Library Program; Biblioteka Kongresu – pomysł na stworzenie Światowej biblioteki Cyfrowej – World Digital Library)[1]. Przedstawiła także przygotowaną w ramach programu NDLP strategię tworzenia bibliotek cyfrowych, składającej się z 3 etapów: *1) automatyzacji bibliotek w sieci lokalnej (uniwersyteckiej), 2) zbudowania narodowej i globalnej biblioteki cyfrowej z odpowiednią infrastrukturą, łączącą istniejące i nowopowstające w sieci biblioteki, 3) integracji bibliotek, muzeów, odbiorców lokalnych i globalnych baz danych* [s. 204].

Autorka rozdzieliła także terminy biblioteka elektroniczna i biblioteka cyfrowa. Podstawą rozróżnienia była postać gromadzonych, przechowywanych i udostępnianych dokumentów oraz sposób ich pozyskiwania i udostępniania. Biblioteka elektroniczna miała gromadzić głównie materiały drukowane, przy czym czynności bibliotekarskie przeprowadzać przy pomocy odpowiednich programów komputerowych (gromadzenie zbiorów: zamawianie, rejestracja zakupów, rachunkowość, opracowanie formalne i rzeczowe: tworzenie rekordów bibliograficznych, katalogów komputerowych, udostępnianie: komputerowa rejestracja wypożyczeń, rezerwacja, monitorowanie, magazynowanie i zabezpieczanie zbiorów: dyskietki, CD-ROM-y, serwery). Biblioteka taka zapewnia także dostęp do systemów i sieci komputerowych (katalogi online, zasoby WWW, poczta elektroniczna). Natomiast biblioteka cyfrowa: *groma-*

[1] Historia – zob. rozdział W.M. Kolasy. Obecna odsłona World Digital Library, Biblioteka Kongresu, UNESCO – zob. http://www.wdl.org/en/about/background.html

dzi, przechowuje i zarządza informacją w postaci cyfrowej i udostępnia ją do przeszukiwania, wynajdywania i kopiowania za pomocą sieci komputerowych, w możliwie najbardziej przyjazny dla użytkownika sposób [s. 205]. Byłby to więc zbiór dokumentów *gromadzonych i przechowywanych wyłącznie na nośnikach, nie odpowiadający tradycyjnie rozumianej bibliotece* [s. 205]. Anna Machalska-Garbacz zaznaczyła przy tym, iż biblioteka cyfrowa nie musi mieć nie tylko swojego budynku (serwery znajdować się mogą w różnych miejscach), czy tradycyjnie pojętych zbiorów, przechowywanych na półkach bibliotecznych, ale także swojego personelu.

Zasoby biblioteki cyfrowej podzielone zostały na: bazy tekstowe (knowledge databases), multimedialne: obrazów (image libraries), dźwiękowe (audio libraries) i animacji (video libraries). I właśnie te multimedialne stać się miały w przyszłości trzonem dla bibliotek cyfrowych. Dodatkowo autorka zakreśliła przyszłe prace nad *opanowaniem metod zarządzania* informacją [s. 206] z włączeniem metod socjologicznych badających relacje człowiek-komputer, przy analizie praw autorskich i przygotowaniu współpracy z wydawcami, ochronie zbiorów, ujednoliceniu systemów (formatów), a także stworzeniu analiz ekonomicznych i finansowych.

*

W **1998** r. w Poznaniu odbyła się konferencja *Świat biblioteki elektronicznej w klasycznej bibliotece naukowej*, w czasie której dyskutowane były także sprawy związane z bibliotekami cyfrowymi. I tak Wanda Pindlowa w artykule *Biblioteka elektroniczna i wirtualna – co to znaczy dla bibliotekarzy i użytkowników* odniosła się do terminów biblioteka zautomatyzowana, elektroniczna, wirtualna i cyfrowa. Termin odnoszący się do największej liczby bibliotek to oczywiście biblioteka zautomatyzowana, w której do pewnych lub wszystkich procesów bibliotecznych wykorzystywane są komputery. W bibliotece elektronicznej działają elektroniczne urządzenia i nośniki, w cyfrowej wszystkie dane są elektroniczne. Biblioteka wirtualna zaś to zasób informacji i dokumentów istniejących w sieci. Autorka zacytowała także A. Browna oraz parafrazę Fredericka Lancastera, iż biblioteki, *jako instytucje umierającego, rozpadającego się społeczeństwa upadają* [s. 19], stąd ratować należy raczej zawód bibliotekarza, gdyż nowe biblioteki będą już innymi instytucjami, na które największy wpływ mieć będą rozwój techniki oraz warunki ekonomiczne.

Mirosław Górny w referacie *Biblioteki elektroniczne – próba kategoryzacji* przedstawił typologię tych *zróżnicowanych form techniczno-organizacyjnych* oraz ich zasobów [s. 25]. W pierwszym podziale odniósł się do cech tradycyjnej biblioteki, wyodrębniając katalogi biblioteczne i bibliografie na nośnikach elektronicznych, czasopisma na nośnikach elektronicznych, książki elektroniczne, zbiory specjalne na nośnikach elektronicznych oraz biblioteki jako kolekcje różnych typów źródeł dotyczących jednej tematyki. Kolejne kryterium to bycie kopią dzieła drukowanego lub zasobem elektronicznym od razu tworzonym na tym nośniku. Biblioteki elektroniczne jako nowy typ systemu informacyjnego zastały przez autora opatrzone cechami: multimedialnością, zastosowaniem hipertekstu, zintegrowanym sposobem dostępu, zasobami nie spotykanymi w świecie druku oraz inną ich organizacją (np. artykuły nie muszą tworzyć całych numerów czasopism, funkcjonują też samodzielnie). Dla nich też

podstawę kategoryzacji stanowić mogą: formaty przechowywania danych, właściwości kolekcji (zwłaszcza jej zmienność), czy rodzaj interfejsu użytkownika. Natomiast kategoryzacja w ramach ogólnego modelu systemu informacyjnego przedstawiona została przez kryteria: 1) nośnik informacji (drukowane, mikrofilmy, elektroniczne), 2) funkcji informacji (metainformacja, informacja faktograficzna, informacja w postaci wypowiedzi naukowych), 3) sposób dostępu (lokalny, zdalny), 4) typy własności (zasoby własne, obce).

W trakcie tej samej konferencji Aleksander Radwański w artykule *Drogi rozwoju biblioteki cyfrowej* przedstawił 3 ścieżki – warianty rozwoju bibliotek cyfrowych: 1) klasyczną, rozpoczynającą się od klasycznej komputeryzacji, a kończącą na technologiach internetowych (komputeryzacja katalogów, czy tworzenie własnych stron WWW), 2) internetową – skupioną wokół wykorzystania serwisów internetowych oraz 3) całościową – wykorzystującą równoczesne implementacje technologii. Jednocześnie zaznaczył, iż pomimo tego, że biblioteka elektroniczna może sugerować poprzez swoją nazwę niezależność od tradycyjnych instytucji, w praktyce często są to twory pozostające w relacji z tradycyjnymi bibliotekami lub archiwami. Dodał też, iż przymiotnik 'elektroniczna' staje się anachroniczny, gdyż wiele bytów nasyconych jest elektroniką. Stąd wybrał termin 'biblioteka cyfrowa', przedstawiając jej cechy: zbiory mają postać cyfrową, informacje o zbiorach także są cyfrowe, obsługa odbywa się przez sieć, a działalność wydawnicza i badawcza biblioteki jest skomputeryzowana. W trakcie dyskusji autor dodał jeszcze, iż biblioteka cyfrowa nie musi oznaczać biblioteki 'zrealizowanej w technologii wirtualnej' [s. 65]. Może to być jednostka dostępna przez telnet i pocztę elektroniczną.

W referacie Anny Łozowskiej i Jolanty Stylińskiej *Biblioteka wirtualna – alternatywą czy koniecznością badacza końca XX wieku* wizja biblioteki przyszłości połączona została z terminem biblioteki wirtualnej mającej dostęp do całości wiedzy ludzkiej [s. 43-44].

Marek M. Górski w referacie *Biblioteka elektroniczna – oczekiwane kierunki rozwoju (na podstawie przeprowadzonych badań)* umiejscowił taką bibliotekę w społeczeństwie informacyjnym, tworząc z niej element 'globalnej infostrady' [s. 48]. Przytoczył też definicję biblioteki wirtualnej [z konferencji w Grazu w 1993 r.], iż nie jest ona ograniczona ścianami budynku, a na dodatek skupia w sobie kilka lub wszystkie biblioteki. Jednocześnie wyznaczył kierunek rozwoju od biblioteki cyfrowej do biblioteki wirtualnej. Przedstawił też model biblioteki cyfrowej wytworzony w trakcie badań studentów III roku bibliotekoznawstwa i informacji naukowej, dla których jest to biblioteka całkowicie skomputeryzowana, ze zbiorami cyfrowymi oraz tradycyjnymi, dostępna przez sieć, w dużym stopniu samoobsługowa.

*

Podobne wieloaspektowe podejście do terminu 'biblioteka cyfrowa' odnaleźć można w innych polskich pracach. Artykuł wymienionego już Aleksandra Radwańskiego, który ukazał się w Elektronicznym Biuletynie Informacyjnym Bibliotekarzy w **1999** r. miał jednak tytuł związany ponownie z 'biblioteką wirtualną' [*Biblioteka wirtualna – problemy definicyjne*]. Określił on termin jako niejednoznaczny, przedsta-

wiając jego pola semantyczne jako: *połączenie spisu zasobów sieciowych z typowo bibliotekarską formą, czyli sposobem porządkowania według określonej klasyfikacji, lub każdy uporządkowany według dowolnych zasad spis połączeń do dowolnych zasobów w Internecie lub bibliotekę cyfrową jutra.* Biblioteka cyfrowa przedstawiona została natomiast jako *cyfrowe zasoby, cyfrowe opracowanie i cyfrowe udostępnianie,* których podstawową cechą wyróżniającą bibliotekę cyfrową od innego dowolnego zbioru jest świadomość jego twórców, że ich zasoby będą częścią pewnej całości.

Felieton wstępny do całego numeru 8 z **1999** r. p.t. *Biblioteki wirtualne, biblioteki cyfrowe...* autorstwa Anny Filipowicz, w którym przedstawiono pewną definicję biblioteki wirtualnej – konstrukcji logicznej, *która jest doraźnie tworzona przez użytkownika z wykorzystaniem niezbędnych narzędzi* w celu *wygodnego dostępu do danych znajdujących się na różnych, oddalonych od siebie komputerach.* Elementy takiej biblioteki to: uporządkowane zbiory danych zapisanych w postaci elektronicznej, istniejące sieci telekomunikacyjne oraz użytkownicy, dysponujący odpowiednimi narzędziami. Przytoczonym najprostszym przykładem mechanizmu działania biblioteki wirtualnej byłby dokument w języku HTML zawierający adresy stron (dokumentów, serwisów, baz danych), zaś najbardziej zaawansowanym – 'world wide library', który to termin zaczerpnięty został z referatu Henryka Holendra *World Wide Lib? O konsorcjach bibliotecznych teraz i w przyszłości*, wygłoszonego na konferencji „Elektroniczne rozpowszechnianie informacji naukowej", Łódź 22-23 VI 1998 r.

W tym też biuletynie pojawiło się sprawozdanie z Konferencji COLIS3 (Conceptions of Library and Information Science) – Dubrownik 23-26.05.**1999** r. Roberta Szymuli. Autor zacytował definicje bibliotek cyfrowych, które zostały przedstawione na tej konferencji, podkreślając, iż biblioteki cyfrowe nie są związane tylko z jedną dziedziną nauki. Informatycy koncentrują się na problemach i rozwiązaniach technicznych; teoretycy bibliotekoznawstwa i informacji naukowej zainteresowani są określeniem podstawowych pojęć i podstaw teoretycznych zagadnienia, natomiast dla bibliotekarzy najważniejsza kwestią jest tworzenie bibliotek cyfrowych i ich udostępnianie.

Pierwszą zaprezentowaną definicją była ta z **1997** r., autorstwa Michaela E. Leska, w której biblioteka cyfrowa to kolekcja – *zbiór informacji po pierwsze cyfrowej, po drugie uporządkowanej.* Druga Richarda M. Hayesa, była bardzo podobna, gdyż kolekcja to *zbiór publikacji rozprowadzanych lub udostępnianych w formie cyfrowej (tj. których symbole są zapisane w bitach i bajtach magnetycznych, elektronicznych lub optycznych).*

Trzecia, stworzona przez Christine L. Borgman (**1999**), często cytowana w innych artykułach, podkreślała dwoistość pojęcia, gdyż z jednej strony: 1) *biblioteki cyfrowe są zbiorem źródeł elektronicznych i połączonych możliwości technicznych tworzenia, wyszukiwania i używania informacji. (...) Biblioteki cyfrowe zawierają dane, metadane opisujące różne aspekty danych (np., reprezentację, twórcę, właściciela, prawa do odtwarzania) i metadane składające się z połączeń lub relacji z innymi danymi czy metadanymi bez względu na to, czy należą one do biblioteki cyfrowej czy do jej środowiska zewnętrznego,*

a z drugiej:

2) *biblioteki cyfrowe są tworzone – a ich zbiory gromadzone i klasyfikowane – przez i dla społeczności użytkowników, a ich możliwości funkcjonalne wspierają potrzeby infor-*

macyjne i cele tej społeczności. Są elementem społeczeństwa, w którym jednostka i grupa komunikują się i współpracują, wykorzystując dane, źródła wiedzy i systemy. W tym sensie rozwijają się, podnoszą jakość i integrują różne instytucje informacyjne jako miejsca, w których źródła informacji są wybierane, gromadzone, klasyfikowane, chronione i udostępniane użytkownikom.

Kolejna definicja Donalda J. Waters'a (przyjęta przez Digital Library Federation): przedstawiała termin z punktu widzenia bibliotekarza: *biblioteki cyfrowe są organizacjami, które dostarczają środków, włączając wyspecjalizowany personel, dobór, klasyfikację, zapewnienie dostępu, dystrybucji, ochrony danych przed ich modyfikacją i ochrony w sensie fizycznym, tworzenie kopii zapasowych i przechowywanie ich w bezpiecznym miejscu, aby mogły być użytkowane przez określoną społeczność lub grupy społeczności.*

R. Szymula przedstawił także problem zamiennego stosowania pojęć 'biblioteka cyfrowa' i 'biblioteka elektroniczna', z zaznaczeniem, iż np. wymieniona już Ch.L. Borgman, przedstawiła następstwo terminów w czasie, pisząc o bibliotekach cyfrowych: coś, *co było* wcześniej *nazywane bibliotekami elektronicznymi.*

Na tej samej konferencji COLIS3 w **1999** r. Maria Kocójowa i Wanda Pindlowa przedstawiły propozycję, a właściwie potrzebę stworzenia biblioteki cyfrowej, zawierającej informacje potrzebne badaczom z zakresu informacji naukowej i bibliotekoznawstwa. Wstępny projekt, nazywany biblioteką wirtualną, zgodnie zresztą z jej zawartością, na którą złożyły się głównie informacje odsyłające, prowadzony był w Instytucie Informacji Naukowej i Bibliotekoznawstwa Uniwersytetu Jagiellońskiego i służył zarówno naukowcom, jak i studentom [http://www.miks.uj.edu.pl/ibin/].

Dodać w tym miejscu jeszcze należy, iż pełnotekstowa Bibliologiczna Biblioteka Cyfrowa istnieje od 2009 r., a stworzona została przez Instytut Informacji Naukowej i Studiów Bibliologicznych Uniwersytetu Warszawskiego.

*

W pracy, co prawda już z **2001** r., ale jeszcze zaliczonej do tego podrozdziału ze względu na jej treść, Mirosławy Mocydlarz *Udostępnianie informacji naukowej na nośnikach elektronicznych* pojawiły się 4 terminy dotyczące ogólnie bibliotek, a wskazujące stopień ich rozwoju. Pierwszy z nich: biblioteka zautomatyzowana, czyli taka w pracy której wykorzystywane są komputery tylko w bardzo ogólny sposób odnosi się do rozważań podjętych w tej publikacji. Natomiast trzy kolejne: biblioteka elektroniczna, cyfrowa i wirtualna nie zostały potraktowane jako synonimiczne terminy. Biblioteka elektroniczna miała być placówką zautomatyzowaną, wykorzystującą w swojej działalności system wyszukiwawczy, działający w formie elektronicznej. Biblioteka cyfrowa była kolejnym stopniem rozwoju 'realnej' biblioteki, różniącym się od biblioteki elektronicznej tym, iż także zbiory miały postać cyfrową (digital born oraz zdigitalizowaną). Opisy tych dwóch bibliotek zbieżne są z przestawieniami biblioteki hybrydowej. Natomiast termin biblioteka wirtualna, określony został jako niejednoznaczny, odnoszący się do systemu gromadzącego różnorodne dokumenty cyfrowo zapisane, *dostępne poprzez elektroniczne media.* Byłby to więc inaczej *zestaw połączeń, czyli linków*, a jedyną różnicą w stosunku do biblioteki cyfrowej było fizyczne rozproszenie zbioru. Za Ole Harbo M. Mocylarz powtórzyła, iż biblioteka cyfrowa

jest *bramą do wszystkich dokumentów całego świata,* a *wszystkie cyfrowe zasoby mogą być zebrane razem z wielu źródeł, w sposób określony głównie przez użytkownika, dla zaspokojenia jego potrzeb naukowych, kulturalnych i poznawczych.*

Poza wymienionymi terminami pojawiło się jeszcze pojęcie: biblioteka pełnotekstowa jako odpowiednik biblioteki klasycznej, *w której można znaleźć książki elektroniczne będące przeniesieniem dzieła drukowanego na nośnik elektroniczny.*

Rozważania terminologiczne dotyczące znaczeń terminów były więc zbieżne z tymi przedstawionymi przez Ewę Chmielewską-Gorczycę oraz Aleksandra Radwańskiego.

4. Biblioteka cyfrowa – XXI wiek: DL, DLS, DLMS, zasób, użytkownik, funkcjonalność, jakość, polityka, architektura/model

a. DELOS

Najważniejsze prace początku XXI wieku (jak na razie), to opisy modeli biblioteki cyfrowej przygotowywane dla Sieci Doskonałości DELOS – Network of Excellence on Digital Libraries, częściowo finansowanej przez European Commission w ramach programu Information Society Technologies Programme (IST). Obecnie nadal funkcjonuje organizacja non-profit DELOS an Association for Digital Libraries [zob. http://www.delos.info], zaś część prac przygotowywana jest w ramach europejskiego projektu DL.org.

W **2001** r. powstał raport *Digital libraries: future directions for a European Research Programme,* w którym biblioteki cyfrowe zaprezentowane zostały jako nowa infrastruktura i wręcz środowisko, które zostały na globalną skalę przygotowane przez integrację usług komputerowych, komunikacyjnych oraz zasobów cyfrowych. Stanowią one istotny składnik systemów informacyjnych XXI w., który ma udostępniać skarby kultury różnych krajów, głównie w celach edukacyjnych i badawczych. Funkcjonalność bibliotek cyfrowych opierać się ma na jej podstawowych funkcjach: 1) spełnianiu różnorodnych potrzeb informacyjnych, 2-5) charakteryzowaniu się wielorakimi źródłami powiązanych informacji, heterogenicznymi informacjami, bogatymi w zasoby źródłami, multimedialnymi danymi, 6) zdefiniowaniu grupy docelowej, 7) posiadaniu zmotywowanych użytkowników, 8) orientacji na zadania, 9) orientacji na domenę, 10) dostępowi ponad ograniczeniami językowymi, 11) współpracy. Na koniec podkreślono jeszcze jedną jej ważną cechę: trwałość istnienia zasobów.

W manifeście z **2007** r. definicja została trochę uściślona – biblioteka cyfrowa stała się narzędziem aktywności jego twórców i odbiorców – różnego typu użytkowników – nieposiadającym logicznych, fizycznych, pojęciowych, czasowych, personalnych granic lub barier informacyjnych. W raporcie z **2008** r., a zwłaszcza w kolejnym manifeście z **2011** r. [*The Digital Library Reference Model*, DL.org: Coordination Action on Digital Library Interoperability, Best Practices and Modelling Foundations w ramach Network of Excellence on Digital Libraries], określono bibliotekę cyfrową jako:

potencjalną wirtualną organizację, która w sposób kompleksowy gromadzi, zarządza i zachowuje na długi czas bogate treści cyfrowe oraz oferuje swoim docelowym społecznościom użytkowników wyspecjalizowane funkcje (umożliwiające także personalizację systemu), określoną jakość oraz zgodność z polityką tworzenia całego systemu. Lokuje się on w centrum działalności intelektualnej człowieka, spełniając różne jego potrzeby: poza intelektualnymi i edukacyjnymi także kulturowe, estetyczne, czy rozrywkowe. System nie może posiadać/tworzyć logicznych, koncepcyjnych (pojęciowych), fizycznych, czy czasowych barier w dostępie do informacji oraz powinien wspomagać ich przełamywanie. Celem istnienia bibliotek cyfrowych jest więc tworzenie kanału komunikacyjnego, do którego każdy ma nie tylko dostęp, ale może też uczestniczyć w dyskusji, ocenie i w podnoszeniu poziomu informowania o zasobach. Na bibliotekę cyfrową składają się 3 systemy: biblioteka cyfrowa (Digital Library – DL), system biblioteki cyfrowej (Digital Library System – DLS) oraz system zarządzania biblioteką cyfrową (Digital Library Management System – DLMS).

Digital Library (DL) – to organizacja, która może być wirtualna, a która kompleksowo gromadzi, zarządza i zabezpiecza dla przyszłych pokoleń bogate treści cyfrowe oraz oferuje społecznościom swoich użytkowników wyspecjalizowane usługi działające na tych treściach, mierzalną jakość oraz ujednoliconą politykę.

Digital Library System (DLS) – to system, który opiera się na określonej architekturze i zawiera wszystkie funkcje wymagane przez konkretną DL. Użytkownicy współpracują z biblioteką cyfrową właśnie przez DLS.

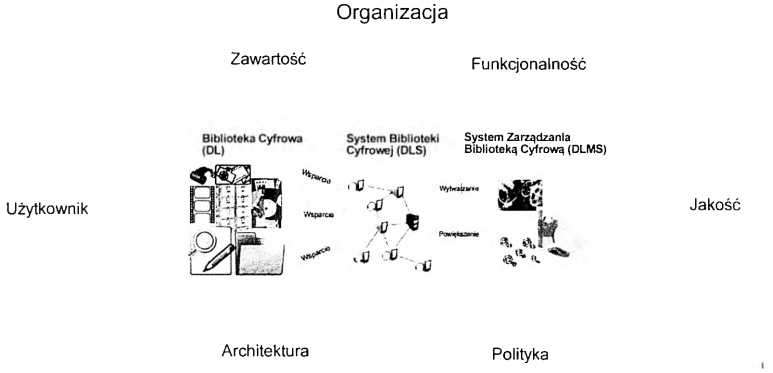
Digital Library Management System (DLMS) to uniwersalny system oprogramowania, który zapewnia odpowiednią infrastrukturę zarówno do tworzenia, jak i zarządzania. Posiada zestaw funkcji uważanych za podstawowe dla biblioteki cyfrowej oraz pozwala na do integrację dodatkowego oprogramowania oferującego bardziej wyrafinowane, wyspecjalizowane i zaawansowane usługi. W związku z tym, że DLMS przynależy do klasy 'oprogramowania' zapewnia on mechanizmy wykorzystywane jako platforma dla stworzenia DSL. Może przynależeć do jednego z trzech następujących typów:

- system rozszerzalny system biblioteki cyfrowej (Extensible Digital Library) – kompletny system w pełni operacyjny w odniesieniu do określonego zestawu podstawowych funkcji z otwartą architekturą oprogramowania, umożliwiający rozbudowywanie – przykładem może być DelosDLMS
- zbiór komponentów oprogramowania (Digital Library System Warehouse) – zbiór, zawierających zestaw podstawowych funkcji DL oraz zestaw narzędzi, które mogą być używane do łączenia tych elementów w różny sposób – oferuje dopasowane i integrację nowych funkcji – przykładem może być BRICK lub DILIGENT / D4Science
- generator systemu (Digital Library System Generator) – wysoko parametryzowany system, który obejmuje szablony zawierające szeroki zakres funkcjonalności, w tym określony zestaw podstawowych, jak i zaawansowanych usług, które zostały uznane za odpowiednie dla potrzeb konkretnych aplikacji. Dzięki sesji inicjalizacji, odpowiednie parametry są ustawiane i konfigurowane, a na

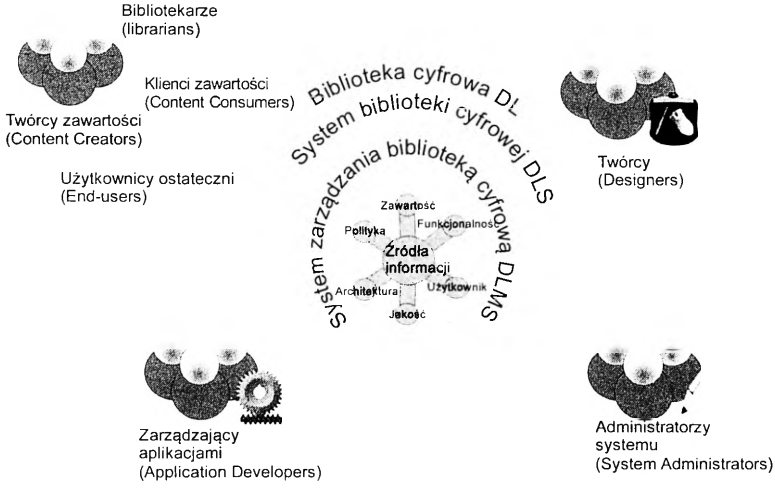
koniec sesji wniosek jest generowany automatycznie – przykładem są ramy (framework) MARIAN wyposażone w język 5SL.

Podstawowe elementy budujące bibliotekę cyfrową, które znalazły się w jej definicji to: zawartość (treść), użytkownik, funkcjonalność, jakość i polityka; szósty pojawia się w definicji Systemu Biblioteki Cyfrowej (DLS): architektura.

Wizualizacje komponentów w odniesieniu do biblioteki cyfrowej (DL), systemu biblioteki cyfrowej (DLS) oraz systemu zarządzania biblioteką cyfrową (DLMS) oraz ludzi związanych z wszystkimi elementami przedstawiają się następująco:



Rys. 4. Pojęcia tworzące definicję biblioteki cyfrowej.



Rys. 5. Elementy biblioteki cyfrowej oraz 'gracze'

Obydwa rysunki przetworzone za: Candela; Nardi 2011.

Dodatkowo przygotowana została także chmura tagów, która w sposób przejrzysty pokazuje terminy wykorzystywane w modelu referencyjnym oraz częstotliwość ich użycia, co może wskazywać na ich znaczenie.



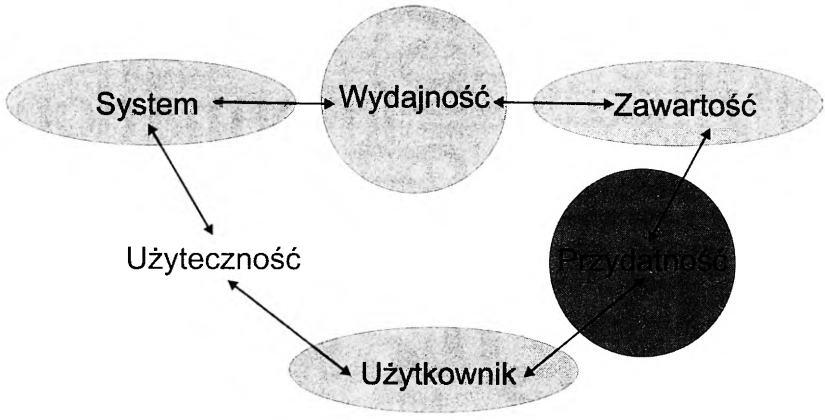
Rys. 6. Chmura tagów oprcowana przez Leonardo Candelę
http://www.dlorg.eu/index.php/outcomes/reference-model.

*

W trakcie przygotowywania raportów oraz ostatniego manifestu z 2011 r. dla sieci Delos powstało wiele prac publikowanych w różnych czasopismach. Część z nich poświęcona była ewaluacji bibliotek cyfrowych (co zostało wykorzystane także przy pracach NISO), jak np. te przygotowane przez Tefko Sarcevica w **2000** r. (*Digital library evaluation: challenges for digital library evaluation*). Biblioteki cyfrowe powiązane w nich zostały z rzeczywistymi bibliotekami oraz zaznaczone zostało pokrewieństwo ich funkcji. Stąd analiza modelowania i oceny zaczęła się od przeniesienia kryteriów, a zakończyła przedstawieniem różnorodności bibliotek cyfrowych z wykazaniem odmiennych ich funkcji. Ważnym zagadnieniem podjętym przez Saracevica było łączenie nastawienia systemo- i użytkowniko-centrycznego (system-center, user-center), a nie przeciwstawianie ich sobie. Takie spojrzenie dualistyczne pozwala na lepszą ocenę poszczególnych tworów.

Kolejny artykuł z **2007** r. *Evaluation of digital libraries* Norberta Fuhra, z którym współpracowali Giannis Tsakonas, Trond Adalberg, Maristella Agosti, Preben Hansen, Sarantos Kapidakis, Claus–Peter Klas, László Kovács, Monica Landowi, András Micsik, Christos Papatheodorou, Carol Peters, Ingeborg Sølvberg, także dotyczył oceny bibliotek. Poza odniesieniem się do 4 członowego systemu biblioteki cyfrowej: kolekcji (dane, o bardzo różnorodnej tematyce, jak i formie), komputerowego systemu informatycznego, oferującego przechowywanie, modyfikację oraz dostęp do danych (infrastruktura sprzętowo-programowa), ludzi (zarówno twórców, jak i odbiorców,

chociaż to użytkownik jest tu najważniejszym ogniwem) oraz projektów, a więc powodów, celów, całej infrastruktury dla jakiej te zasoby zostały stworzone, autorzy ci przygotowali model składający się z trzech elementów oraz łączących ich relacji, aby ułatwić samą ocenę. Wizualizacja przedstawia się następująco:



Rys. 7. Relacje łączone elementy tworzące bibliotekę cyfrową. Za: Fuhr 2007.

*

Inny artykuł, ale odnoszący się także przede wszystkim do 'komponentu': ludzie to praca z **2005** r. *Reference models for digital libraries: actors and roles* autorów Jose Borbinha, John Kunze, Angela Spinazzè, Peter Mutschke, Hans-Jörg Lieder, Michael Mabe, Larry Dixson, Howard Besser, Becky Dean, Warwick Cathro. Przedstawili oni poszczególnych 'aktorów' tworzących i wykorzystujących biblioteki cyfrowe, a także określili ich role.

*

Najlepiej jednak wszystkie osoby zaangażowane w tworzenie i użytkowanie bibliotek cyfrowych przedstawiła Ying Zhang, w pracach pod tymi samymi tytułami *Developing a holistic model for digital library evaluation* z **2007** i **2010** r. Podział wszystkich grup odbiorców oraz osób zaangażowanych w tworzenie bibliotek cyfrowych pokazuje 5 klas, niekoniecznie zawsze wykluczających się: 1) tzw. użytkownicy końcowi (klienci) (user), 2) badacze (researcher), 3) bibliotekarze (librarian), 4) zarządzający biblioteką cyfrową (developer) oraz 5) administratorzy (administrator). Natomiast cały system biblioteki cyfrowej podzielony został na 6 głównych grup: kontekst (z najważniejszymi cechami: zrównoważony rozwój, współpraca i wsparcie w zakresie zarządzania), użytkownik (cechy: wydajność systemu, wyszukiwanie zakończone sukcesem, satysfakcja), serwis (cechy: dostępność, niezawodność, szybkość reakcji, przydatność dla docelowych użytkowników, integralność informacji – łatwość wyszukiwania), interfejs (cechy: łatwość użycia, spójność, skuteczność), zawartość (cechy: dostępność, bezbłędność, przydatność dla użytkowników w celu osiągnięcia określonych celów), technologia (cechy: niezawodność i łatwość obsługi).

b. NISO – Framework of Guidance for Building Good Digital Collections

Początek XXI w. to także prace nad *A framework of guidance for building good digital Cellections*, którego pierwsze wydanie ukazało się w **2001** r., drugie w **2004**, a trzecie w **2007** r. *Framework* sporządzony został przez National Information Standards Organization (NISO) Framework Working Group, z pomocą Institute of Museum and Library Services (IMLS). Celem przygotowania tego opracowania było przedstawienie najważniejszych elementów budujących bibliotekę cyfrową, wraz z analizą ich jakości, przegląd działań podejmowanych w celu stworzenia dobrych zbiorów cyfrowych, identyfikacja istniejących zasobów, które wspierają praktykę tworzenia i zarządzania dobrymi zasobami cyfrowymi oraz zachęta do uczestniczenia w ciągłym rozwoju bibliotek cyfrowych. W podręczniku dotyczącym przygotowania tzw. *dobrej kolekcji cyfrowej*, w definicji biblioteki cyfrowej przedstawiono 4 elementy tworzące cały system: kolekcje (grupy obiektów), obiekty (materiały cyfrowe), metadane (informacje o obiektach i zbiorach) i inicjatywy (programy lub projekty do tworzenia i zarządzania zbiorami) (*Collections, Objects, Metadata, Initiatives*).

*

c. Model 5S

W modelu tym opisanym dla Digital Library Research Laboratory Virginia Tech (CS & IT) m.in. przez Marcosa André Gonçalvesa, Edwarda A. Foxa, Layne T. Watson, Neill A. Kipp oraz Barbarę L. Moreira czy Alberto H.F. Laendera, E. A. Lalova, Norbert Fuhr, Carlo Meghini (**1999, 2002, 2003, 2004, 2006, 2007, 2008, 2009**) popatrzono na bibliotekę cyfrową jako na system składający się z zasobu danych, metadanych umożliwiających dostęp do obiektów, różnego rodzaju rozwijalnych usług oraz społeczności użytkowników. 5S to pięć podstawowych elementów: 1) strumienie (streams), 2) struktury (structures), 3) przestrzenie (spaces), 4) scenariusze (scenarios), 5) społeczności (societies).

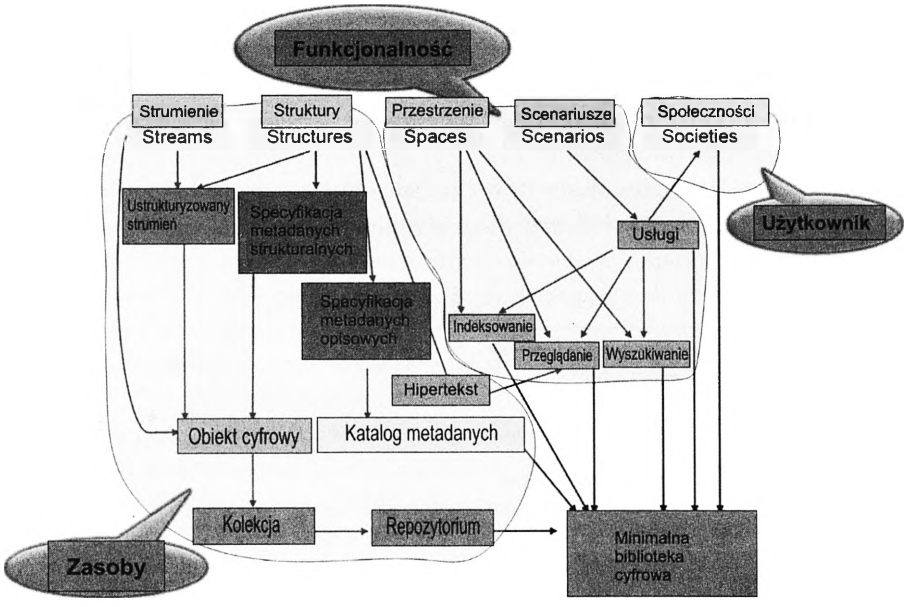
Opis poszczególnych komponentów za tabelą umieszczoną na stronie http://www.dlib.vt.edu/projects/5S-Model/ wygląda następująco:

Tabela 4. Komponenty modelu 5S.

Modele	*Materiały/elementy pierwotne*	*Formalizm*	*Cele*
Stream Model (strumienie)	tekst; wideo; audio; software (oprogramowanie)	porządek; typy	opisanie właściwości treści bibliotek cyfrowych, takich jak kodowanie i język dla materiałów tekstowych lub poszczególne formy danych multimedialnych

Structural Model (struktury)	kolekcja; katalog; hipertekst; dokument; meta dane; narzędzia organizujące dane	wykresy; węzły; linki; etykiety; hierarchie	określenie aspektów organizacyjnych treści bibliotek cyfrowych
Spatial Model (przestrzenie)	interfejs użytkownika; indeks; model wyszukiwania informacji	zestawy danych; operacje; przestrzenie wektorowe; przestrzeń miar; przestrzeń prawdopodobieństwa	definiowanie logicznych i prezentacyjnych sposobów przedstawiania składników bibliotek cyfrowych
Scenarios Model (scenariusze)	usługi; zdarzenia; warunki; działania	diagramy sekwencyjne; diagramy współpracy	ustalenie szczegółowych zachowań serwisów (usług) bibliotek cyfrowych
Societies Model (społeczności)	społeczność; menedżerowie; aktorzy; klasy; relacje; atrybuty; operacje	konstrukcje modelowane zorientowane na obiekt (object-oriented); wzory projektowania (design)	definiowanie menedżerów, odpowiedzialnych za działalność biblioteki cyfrowej; definiowanie aktorów, którzy korzystają z tych usług; ustalenie relacji między nimi

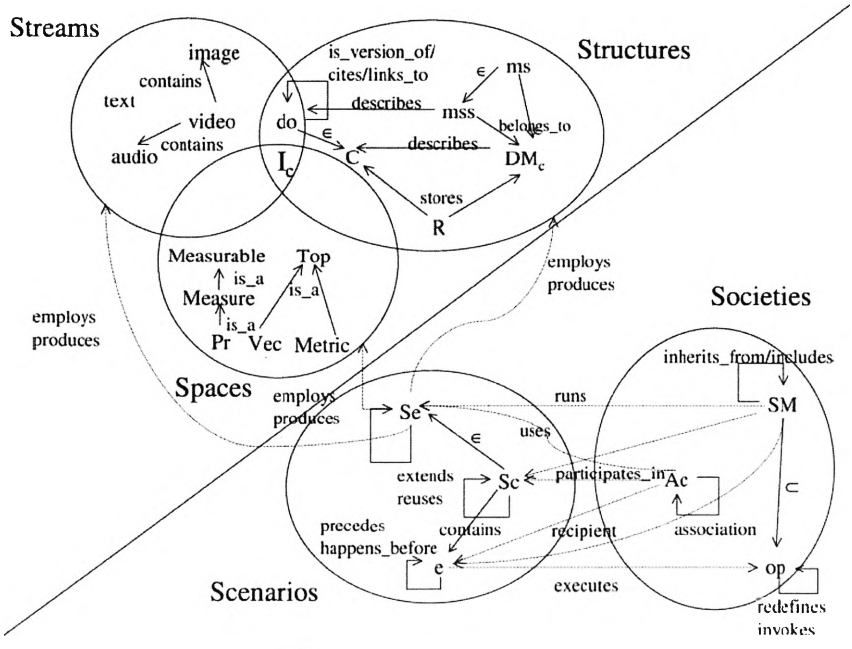
Wizualizację porównania modelu 5S do modelu referencyjnego zamieszczono np. w pracy pod red. Leonardo Candeli i A. Nardi, *The Digital Library Reference Model.* Przedstawia się ona następująco:



Rys 8. Model 5S i model referencyjny. Za: Candela; Nardi 2011.

*

W nawiązaniu do omówionego powyżej modelu 5S w **2008** r. przedstawiony został artykuł autorów Marcosa André Gonçalvesa, Edwarda A. Foxa i Layne T. Watson *Towards a digital library theory: a formal digital library ontology*, w którym zaproponowano konceptualizację, która *nie predestynowała do bycia teorią bibliotek cyfrowych*. W związku z tym, iż, jak już to wielokrotnie w tej pracy powtarzano, biblioteki cyfrowe nie ominęły dyskusje definicyjne oraz nie został ustalony ostateczny konsensus i porozumienie, autorzy zaproponowali formalną ontologię systemu biblioteki cyfrowej, składającą się z terminów dotyczących elementów składowych, jak i poszczególnych relacji oraz wyznaczyli elementy oceny dla każdego komponentu. Główna wizualizacja z zaznaczonymi relacjami zarówno wewnątrz, jak i pomiędzy elementami przedstawia się następująco:



Rys. 9. Model 5S bibliotek cyfrowych wizualizowanych za pomocą ontologii (s. 97)[2].

*

d. Piśmiennictwo zagraniczne

Jedną z pierwszych prac, gdzie pojawiły się szersze analizy użytkowników bibliotek cyfrowych oraz próba ich typologii, a nawet pojawił się termin 'ogólnej/powszechnej publiczności bibliotek cyfrowych' (generał audience digital libraries – GADL) jest ar-

[2] Terminy pozostawiono w wersji oryginalnej, nie przetłumaczono na język polski, gdyż przedstawienie wszystkich pojęć zajęłoby zbyt wiele miejsca – autorka odsyła wszystkich zainteresowanych do oryginalnego artykułu.

tykuł z **2002** r. *Traditional IR for web users: a context for general audience digital libraries* autorów: Dietmar Wolfram i Hong (Iris) Xie. Badania użytkowników końcowych miały przedstawić cechy określające populację użytkowników, określić jaka zawartość jest dla nich najbardziej interesująca oraz w jakim celu jest wykorzystywana, a także ustalić w jaki sposób użytkownicy wchodzą w interakcje z systemami informacyjnymi oraz z bibliotekami cyfrowymi. Autorzy badali wykorzystanie BadgerLink [http://www.badgerlink.net/], multiwyszukiwarkę naukową, która przeszukuje także zasoby pełnotekstowe, stąd w wynikach pojawiły się cele naukowe i edukacyjne, czy wykorzystywanie systemu dla potrzeb zawodowych.

Ci sami autorzy w **2002** r. opublikowali także artykuł *State digital library usability: contributing organizational factors,* w którym po raz kolejny pojawiły się terminy 'publiczności bibliotek cyfrowych' oraz 'graczy', do których zaliczeni zostali: programiści / projektanci, dostawcy treści, użytkownicy ostateczni, jak i organizacje z różnymi oczekiwaniami i potrzebami. Określone zostały także trzy tryby interakcji: wpływ, komunikacja, aktywność, z zaznaczeniem, iż nie funkcjonują jeszcze kanały komunikacji pomiędzy poszczególnymi graczami.

*

Kolejna publikacja, raport z **2002** r. przygotowany dla Digital Library Federation, Council on Library and Information Resources autorstwa Daniela Greensteina i Suzanne E. Thorin o intrygującym tytule *The digital library: a biography,* po raz kolejny wskazał na zamieszanie terminologiczne, a zwłaszcza opisał biblioteki cyfrowe jako różnorodne obiekty, istniejące w wielu formach, spełniające zupełnie inne funkcje, posiadające czasem wręcz odmienne priorytety i cele. Raport został przygotowany, gdyż bibliotekarze oraz dyrektorzy bibliotek stając często 'przed niewiadomym' spragnieni są informacji o różnych doświadczeniach, przekazu danych o tym, za co biblioteki cyfrowe są uważane i jakie inwestycje trzeba podejmować aby stworzyć kolekcje dobre, 'sensowne' i opłacalne. Stąd znalazły się tu przedstawienia 'case study' amerykańskich bibliotek cyfrowych: California Digital Library (University of California), Harvard University, Indiana University, New York University, University of Michigan, University of Virginia.

*

W **2007** r. ukazał się następny artykuł o ciekawym tytule *Defining a digital library* dwóch autorów Michaela Seadle'a i Elke Greifeneder. Autorzy przedstawili wiele definicji bibliotek cyfrowych. Pierwszą z nich jest opracowana przez National Science Foundation (NSF) z 1999 r., która przedstawiona została na stronie Digital Library Initiative: biblioteki cyfrowe w zasadzie przechowywują materiały w formie elektronicznej i 'sterują' (manipulate) skutecznie dużymi zbiorami. Badania bibliotek cyfrowych to badania sieciowych systemów informacyjnych, które koncentrować powinny się na tym, jak rozwijać niezbędną infrastrukturę, aby skutecznie zarządzać masową informacją. Niestety problem z tą definicją jest taki, iż głównie uwzględnia ona aspekty czysto techniczne. Brak włączenia ludzi w ten kontekst. Stąd omówiono propozycję Institute of Museum and Library Services (2005), w której odniesiono się

do 'rewolucji cyfrowej' (Digital Revolution). Rewolucja ta dotykając niemal każdego aspektu biblioteki, czy muzeum, od automatyzacji wewnętrznych systemów rejestracji nabytków poczynając, przez digitalizację zbiorów fizycznych, po nabywanie nowych 'born-digital' publikacji, czy dzieł sztuki wymaga wykorzystania nowych technologii oraz zaangażowania publiczności.

M. Seadle zacytował także sam siebie (**2006**), opisując bibliotekę cyfrową zasadniczo jako zasób, który *rekonstruuje intelektualnie substancję i usługi tradycyjnej biblioteki w formie cyfrowej*. Biblioteki cyfrowe składają się z treści cyfrowych, połączeń (linki metadanych, relacje) oraz oprogramowania (proste strony w HTML lub złożonych systemów zarządzania bazami danych). Nie są zamiennikami dla tradycyjnych bibliotek, raczej są ich przyszłością.

Autorzy przedstawili także próby stworzenia definicji podjęte przez ich studentów, a także odnieśli się do definicji z lat dziewięćdziesiątych, przedstawionych już w tym rozdziale. W konkluzji stwierdzili, iż biblioteki cyfrowe jako twór dość młody, są obiektem trudnym do określenia ich ostatecznej definicji, zwłaszcza, iż nie są w żaden sposób stałe, wręcz przeciwnie, można mówić o nich jak o 'dziełach / tworach' bardzo dynamicznych. I najprawdopodobniej przyszłe pokolenia popatrzą na nie w całkiem inny sposób, konceptualizując na nowo ich misję w świecie cyfrowym.

*

e. Polska

W XXI w. polskie piśmiennictwo przyniosło więcej tematów związanych z bibliotekami cyfrowymi, prezentowanych zarówno w książkach, materiałach konferencyjnych, czy artykułach w czasopismach i serwisach internetowych.

Jedną z pierwszych wydanych ***książek*** była praca Juranda B. Czermińskiego *Cyfrowe środowisko współczesnej biblioteki* (**2002**), w której nie ma co prawda wyjaśnionych terminów biblioteka cyfrowa, biblioteka wirtualna, ale jest właśnie odesłanie do cyfrowości bibliotek. Jednym z nielicznych odniesień do istniejących systemów jest wymienienie Wirtualnej Biblioteki Arcydzieł Literatury Polskiej (projekt UNESCO, Biblioteka Główna Uniwersytetu Gdańskiego: http://monika.univ.gda.pl/~literat/index.htm), jako zbioru pełnotekstowych dokumentów, jakościowo zweryfikowanego i opatrzonego notami biograficznymi. Natomiast analiza środowiska cyfrowego systemu nowoczesnej biblioteki przypomina definicje bibliotek cyfrowych. Na całość systemu składają się bowiem *zasoby ludzkie, sprzętowe, programowe oraz wszelkiego rodzaju dane. Zasobem systemu jest każdy jego element, który może być przydzielony określonej części systemu lub programowi użytkowemu* [s. 30]. Zarządza nimi system operacyjny. Zasoby podzielone zostały na sprzętowe (komputery, urządzenia peryferyjne, sprzęt sieciowy) oraz programowe (oprogramowanie systemowe, oprogramowanie użytkowe, dane przechowywane w pamięciach masowych). Natomiast architektura systemu przedstawiona została następująco: 1) stajnia serwerów (tu np. serwer produkcyjny, archiwizacyjny, weryfikacyjny, strefa prezentacyjna), 2) stanowiska klienta systemu (tu moduły wprowadzania i wyprowadzania informacji), 3) infrastruktura komunikacyjna.

W związku z tym, iż w systemie cyfrowym operuje się obiektami mającymi wyłącznie postać cyfrową, całe środowisko cyfrowe interpretowane zostało jako zespół czterech systemów: 1) bezpośredniego lub pośredniego (poprzez moduły rozpoznawania i konwersji) zapisu liczb, 2) ich przetwarzania, 3) transmisji i 4) prezentacji.

*

Do cyfryzacji odwoływała się także książka z **2004** r. *Cyfryzacja w procesach komunikowania* pod red. Wandy Krzemińskiej i Piotra Nowaka. W pracy tej poza artykułem Wandy Krzemińskiej *Co z tą cyfryzacją?*, w której autorka przedstawiła analizy na temat pojęć 'cyfryzacja', 'digitalizacja' Witold Król omówił problemy terminologiczne oraz wykorzystanie tych dwóch pojęć w piśmiennictwie polskim i zagranicznym (*Digitalizacja – problemy terminologiczne).* W tekście przedstawiono digitalizację głównie jako proces przekształcania dowolnej informacji do postaci cyfrowej. W tle omówiono stosowanie różnych ujęć terminów pochodnych: cyfryzacja, dyskretyzacja, dygitalizacja itp. [s. 25]. Jednocześnie autor wyznaczył *jednolity kierunek dla stosowania pojęcia digitalizacja.* Poza tym w książce tej przedstawiono zagadnienia związane z digitalizacją, wykorzystywaną w różnych krajach, czy dyscyplinach, a zwłaszcza przeanalizowano jej wpływ na komunikowanie się.

*

Następne publikacje powstałe z punktu widzenia procesów digitalizacyjnych, tworzenia bibliotek cyfrowych to np. książka Aleksandera Trembowieckiego *Digitalizacja zbiorów bibliotecznych: teoria i praktyka*, czy materiały z ogólnopolskiej konferencji *Dygitalizacja zbiorów bibliotecznych* (Warszawa, 2005) pod red. Elżbiety Stefańczyk – obie z **2006** r. W tym samym roku odbyła się także konferencja Biblioteki Politechniki Łódzkiej pod znaczącym tytułem: *Biblioteki XXI wieku: czy przetrwamy?* W trakcie obrad dyskutowano nad kierunkami rozwoju bibliotek, zwłaszcza akademickich, podkreślając już ich istniejącą hybrydowość (Małgorzata Rożniakowska, Marcin Margas, Iwona Kitlińska, Piotr Bógdoł *Biblioteki dawniej i dziś. Hybrydowe, cyfrowe...? Jakie będą i co może wpłynąć na ich kształt w przyszłości*), czy proponując nazywanie ich 'cyberary' – cyber-bibliotekami (Lidia Szczygłowska *„Cybrary" to wciąż biblioteka*).

*

W tym miejscu przedstawić jednak należy trochę bardziej szczegółowo kolejną pracę *Standardy w procesie digitalizacji obiektów dziedzictwa kulturowego* pod red. Grzegorza Płoszajskiego (**2008**). Publikacja opierała się na raporcie wewnętrznym Zespołu roboczego ds. standardów technicznych digitalizowanych obiektów, działającego w ramach Zespołu do spraw digitalizacji powołanego 24 kwietnia 2006 r. przez Ministra Kultury i Dziedzictwa Narodowego *Opracowanie standardów technicznych dla obiektów cyfrowych tworzonych przy digitalizacji dziedzictwa kulturowego.* Jej tematyka skupiła się więc na kwestiach technicznych dotyczących samego procesu digitalizacji oraz przedstawieniu różnego typu metadanych, opisujących zarówno oryginalny obiekt digitalizowany, jak i proces digitalizacji oraz strukturę złożonego obiektu cyfrowego, służąc zarówno bieżącemu zarządzaniu obiektem cyfrowym czy jego długotrwałemu przechowywaniu. Stąd brak w pracy definicji bibliotek cyfrowych jako

systemów. Odwołać się jednak do niej należy przy analizie metadanych, zwłaszcza technicznych i strukturalnych, czy całego procesu digitalizacji.

*

Z innych prace, analizujących biblioteki cyfrowe, a nie głównie procesy digitalizacji wymienić należy pracę zbiorową pod red. Jadwigi Woźniak-Kasperek i Jerzego Franke *Biblioteki cyfrowe: projekty, realizacje, technologie* (**2007**), gdzie we wstępie napisano, iż termin ten *obejmuje wszelkiego typu kolekcje, zbiory dokumentów elektronicznych dostępnych publicznie* oraz, że *wzniosła idea, która legła u podstaw tworzenia bibliotek publicznych, bezpłatnego dostępu do piśmiennictwa, dzięki bibliotekom cyfrowym nabiera dodatkowych znaczeń*. Redaktorzy dodali też uwagę, iż zdefiniowane wcześniej różnice między biblioteką elektroniczną, cyfrową czy wirtualną, *w konfrontacji z praktyką digitalizacyjną wydają się mało istotne*. Ważne jest tworzenie dobrych bibliotek cyfrowych, które niekoniecznie muszą być powiązane z zasobami jednej, czy wielu współpracujących ze sobą instytucji.

*

Jakby w odpowiedzi na tę uwagę kolejne publikacje skupiały się przede wszystkim na projektach, możliwościach technicznych, oprogramowaniu, standaryzacji itp. I tak w **2008** r. wydana została praca Małgorzaty Kowalskiej *Dygitalizacja zbiorów bibliotek polskich,* w której przedstawiono proces digitalizacji zbiorów, jego organizacyjne, ekonomiczne i prawne uwarunkowania oparte o badania przeprowadzone w bibliotekach polskich. Omówiono także działalność wybranych instytucji i organizacji międzynarodowych na rzecz cyfrowej ochrony światowego dziedzictwa kulturowego oraz przeanalizowano polskie plany i strategie rozwoju tworzone m.in. z myślą o zasobach cyfrowych.

Wtedy też (w 2008 r.) odbyła się kolejna konferencja: *Informacja dla nauki a świat zasobów cyfrowych*, z której materiały wydane pod red. Haliny Ganińskiej w dużym stopniu odnosiły się do istniejących oraz tworzonych właśnie bibliotek cyfrowych (zob. artykuły: Cezarego Mazurka, Elżbiety Skubały i Anny Kazan, Doroty Buzdygan i Doroty Lipińskiej, Barbary Kubiak i Anny Gogiel-Kuźmickiej, Doroty Gill-Tarnowskiej, Anny Łozowskiej i innych).

*

W tym też roku wystartowała pierwsza edycja, odbywanych odtąd corocznie w Poznaniu ogólnopolskich i międzynarodowych konferencji *Polskie Biblioteki Cyfrowe* (**2008, 2009, 2010, 2011**), której organizatorami są Biblioteka Kórnicka PAN, Poznańska Fundacja Bibliotek Naukowych oraz Poznańskie Centrum Superkomputerowo-Sieciowe. Integruje ona środowisko twórców oraz osób zainteresowanych tworzeniem oraz wykorzystywaniem bibliotek cyfrowych, a teksty, prezentacje oraz wystąpienia, jak przystało na osoby otwarte na dzielenie się wiedzą dostępne są na stronie: http://www.man.poznan.pl/PBC/ oraz oczywiście w formie materiałów drukowanych.

*

Kolejna publikacja książkowa także dotyczyła istniejących już systemów, a mianowicie bibliotek cyfrowych przygotowanych na platformie stworzonej przez PCSS – dLibra (Anna Wałek **2009**). Poza opisem technologicznych podstaw funkcjonowania bibliotek cyfrowych, przedstawieniem zarysu ich historii (w Polsce) i analizą modeli szerzej omówiono Dolnośląską Bibliotekę Cyfrową, Kujawsko-Pomorską Bibliotekę Cyfrową oraz Śląską Bibliotekę Cyfrową.

*

W **2009** r. ukazały się także materiały z konferencji *Uniwersum piśmiennictwa wobec komunikacji elektronicznej* pod red. Krzysztofa Migonia i Marty Skalskiej-Zlat, w których to znalazły się m.in. artykuły o funkcjonalności i dostępności bibliotek cyfrowych (Jadwigi Woźniak-Kasperek), digitalizacji zbiorów utraconych i rozproszonych (Ryszarda Nowakowskiego), czy wykorzystaniu zbiorów cyfrowych bibliotek na przykładzie Dolnośląskiej Biblioteki Cyfrowej (Anny Wałek).

*

Praca na podstawie zasobów Wielkopolskiej Biblioteki Cyfrowej, wspomagająca pracę filologów, autorstwa Mirosława Górnego i Piotra Wierzchonia *Polish digital libraries as a philologist's tool: based on 666 adjectives from the Digital Library of Wielkopolska*, ukazująca właśnie jak można wykorzystać bibliotekę cyfrową w pracy naukowej wyszła rok później – w **2010** r.

*

W tymże **2010** r. ukazał się także, wydany przez Bibliotekę Narodową, podręcznik, przedstawiający procesy digitalizacji, z analizą i przedstawieniem tzw. work-flows – *Digitalizacja piśmiennictwa* (oprac. i red. Dariusz Paradowski). Książka ta prezentuje wszelkie prace digitalizacyjne, począwszy od planowania projektów, przez prace przygotowawcze, aż po opis i przedstawienie *praktycznych wskazówek dotyczących działań cyfryzacyjnych* [http://www.bn.org.pl/aktualnosci/208-pierwszy-polski-podrecznik-digitalizacji-obiektow-pismienniczych.html]. Stąd znalazły się w niej modele metadanych, opis pracowni digitalizacyjnej, metody przetwarzania i przechowywania zasobów cyfrowych itp.

*

W **2010** r. Centrum Promocji Informatyki wydało książkę *Cyfrowy świat dokumentu: wydawnictwa, biblioteki, muzea, archiwa* pod red. Henryka Hollendra, na którą składają się ponownie rozprawy o digitalizacji dokumentów. Rozważono w niej także różne aspekty działań digitalizacyjnych: techniczne, użytkowe i społeczne (np. w artykułach *Polskie biblioteki cyfrowe A.D. 2010: kształtowanie paradygmatu* Aleksandar Trembowieckiego, *Do czego jest potrzebna informatyka w digitalizacji?* Marcina Werli, *Digitalizacja, cyfryzacja czy dyskretyzacja i jaka?* Kazimierza Schmidta, *Aspekty prawne digitalizacji* Rafała Golata, *Krajowa strategia cyfryzacji dziedzictwa narodowego* Tomasza Makowskiego i Katarzyny Ślaskiej, *Digitalizacja dokumentów piśmienniczych – strategie rozwoju* Jerzego Franke, *Jak posługiwać się biblioteką cyfrową?* Lidii Derfert-Wolf, czy *Pozyskiwanie finansowania na budowanie cyfrowych zasobów* Bożeny Bednarek-Michalskiej).

*

Z ***artykułami*** drukowanymi w czasopismach czy pracach zbiorowych sprawa wygląda podobnie, jak z pracami samodzielnymi wydawniczo. Głównie są to teksty poświęcone poszczególnym bibliotekom cyfrowym, nowym lub rozwijanym projektom, digitalizacji zbiorów.

Spośród nich wymienić należy referat Nancy Rajczak, wygłoszony w trakcie konferencji Instytutu Informacji Naukowej i Bibliotekoznawstwa UJ w **2001** r. [wydrukowany w 2002]. Autorka zaznaczyła w nim, iż biblioteka cyfrowa, to nie tylko odpowiednik dla cyfrowych zbiorów i narzędzi zarządzania informacją, ważny jest sam proces tworzenia informacji, kategoryzowania jej w różne zbiory oraz udostępniania on-line [s. 24]. Same zbiory, za propozycją przygotowaną przez Bibliotekarzy University of California, podzielone zostały na cztery poziomy: 1) materiał archiwalny (cyfrowy – zawsze dostępny), 2) dokumenty oferowane (czasowo), 3) dokumenty lustrzane (lokalne kopie dokumentów cyfrowych) oraz 4) dokumenty połączone (dostępne tylko poprzez hiperłącza z witryny biblioteki) [s. 26]. Jak widać różne zasoby wskazują na połączenie w jednym 'tworze' rozumienia biblioteki cyfrowej (dokumenty elektroniczne) z wirtualną (linki).

Jednym z nielicznych artykułów poświęconych terminologii jest artykuł Aliny Nowińskiej z **2003** r. *Niektóre problemy przyswajania angielskiej terminologii z zakresu bibliotekarstwa i informacji naukowej na przykładzie „Bibliografii Analitycznej Bibliotekoznawstwa i Informacji Naukowej".* Biblioteka elektroniczna została w nim określona m.in. jako nazwa dla programów np. krajowych w Holandii, Czechach, Wielkiej Brytanii (wspólne katalogowanie, gromadzenie, wypożyczanie oraz udostępnianie różnorodnych źródeł, w tym tekstów czasopism elektronicznych i dokumentów pozyskanych z www). Biblioteka wirtualna odniesiona została m.in. do wizji społeczeństwa informacyjnego, ale też do konkretnych projektów digitalizacyjnych. Natomiast biblioteka cyfrowa omówiona została jako zbiór dokumentów cyfrowych oraz ponownie jako programy krajowe, czy też system obsługujący określone grupy użytkowników, wykorzystujący nowoczesną technologię w celu zapewnienia szybkiego dostępu do zasobów.

*

Kolejny artykuł, w którym wiele miejsca zostało poświęcone bibliotekom cyfrowym to *Cyfrowa nauka – cyfrowe publikacje – cyfrowe biblioteki* Marka Nahotki z **2007** r. Tzw. globalna biblioteka cyfrowa przedstawiona została jako system składający się z następujących elementów: 1) zbiorów, które są rozproszone w sieci; prawa własności: Open Access, 2) bibliotekarzy cyfrowych, pracujących często w systemie telepracy; wolny zawód, 3) użytkowników potencjalnych, rozproszonych geograficznie, *z bardziej różnorodnymi potrzebami,* 4) pomieszczeń zlokalizowanych w wielu miejscach; biblioteka jako portal. Funkcje takiej biblioteki to: 1) gromadzenie *polegające na wyszukiwaniu i lokalizowaniu z sieci zasobów spełniających określone kryteria jakościowe,* 2) opracowanie, którego efekty w postaci rozproszonych metadanych wykorzystywane są przez różne narzędzia wyszukiwawcze, 3) przechowywanie – rozpro-

szone, 4) udostępnianie – także rozproszone, ale i zdalne, 5) informowanie głownie o zasobach wyselekcjonowanych z sieci [s. 23].

Według autora biblioteka taka znacznie bardziej będzie włączona w procesy badawcze i dydaktyczne i jako tako odpowiadać będzie na marzenia wielu naukowców o łatwym i szybkim dostępie do wszelakiej informacji.

*

W ostatniej już szerzej omówionej pracy z **2008** r. *eContent czyli o organizacji informacji i wyszukiwaniu w bibliotece cyfrowej* Jadwiga Woźniak-Kasperek na podstawie badań piśmiennictwa określiła terminy biblioteka wirtualna, cyfrowa i elektroniczna jako terminy traktowane synonimicznie lub jako pojęcia bliskoznaczne. Dodała też, iż w polskim piśmiennictwie daje się zauważyć tendencja do skupiania uwagi na problemach digitalizacji.

*

Z pozostałych polskich autorów zajmujących się bibliotekami cyfrowymi wymienić należy (w kolejności alfabetycznej): Bożenę Bednarek-Michalską (z artykułami o: Kujawsko-Pomorskiej Bibliotece Cyfrowej, digitalizacji zbiorów bibliotecznych, standardach itp.), Lucjana Bilińskiego (interpelacja poselska w sprawie bibliotek cyfrowych), Lidię Derfert-Wolf (nowy model komunikacji naukowej), Zdzisława Dobrowolskiego (nowe formy książki), Anetę Drabek (cyfrowy warsztat), Mirosława Górnego (Wielkopolska Biblioteka Cyfrowa, biblioteki cyfrowe), Henryka Hollendra (cykl esejów: Cyfrowe okolice), Bolesława Howorkę (prawo autorskie), Łukasza Jastrzębia (dLibra, prace Biblioteki Kórnickiej Polskiej Akademii Nauk, Wielkopolska Biblioteka Cyfrowa), Tomasza Kalotę (tworzenie bibliotek cyfrowych, digitalizacja), Władysława Marka Kolasę (tworzenie bibliotek cyfrowych i ich historia, digitalizacja), Wojciecha Kowalewskiego (Małopolska Biblioteka Cyfrowa), Małgorzatę Kowalską (projekty), Remigiusza Lisa (digitalizacja, Śląska Biblioteka Cyfrowa), Tomasza Makowskiego (Biblioteka Narodowa), Lidię Marcinkiewicz (Książnica Pomorska), Jolantę Mazurek (systemy bibliotek cyfrowych, użytkownicy, potrzeby informacyjne użytkowników), Agatę Muc (kościelne biblioteki wirtualne w Polsce), Marka Niezgódkę (Polska Biblioteka Wirtualna Nauki), Veslavę Osińską (wizualizacja w bibliotekach cyfrowych), Grażynę Piotrowicz (cyfrowa przyszłość), Joannę Potęgę (stan rozwoju polskich bibliotek cyfrowych, Polona), Sybillę Stanisławską-Kloc (zagadnienia prawne), Barbarę Szczepańską (zagadnienia prawne), Katarzynę Ślaską (polskie biblioteki cyfrowe, Polona), Annę Wałek (Dolnośląska Biblioteka Cyfrowa), Iwonę Zbroszczyk, Sylwię Ziółkowską (Bałtycka Biblioteka Cyfrowa) i innych. Dodać jeszcze należy pracowników PCSS-u, twórcy systemu dLibra oraz przedstawicieli Federacji Bibliotek Cyfrowych: Adama Dudczaka, Agnieszkę Lewandowską, Cezarego Mazurka, Tomasza Parkołę, czy Marcina Werlę.

*

Część materiałów dostępna jest oczywiście ***w sieci Internet*** np. na stronach konferencji Polskie Biblioteki Cyfrowe [http://www.man.poznan.pl/PBC/], Federacji Bibliotek Cyfrowych [http://fbc.pionier.net.pl/], Serwisu Elektroniczna Biblioteka

[http://www.nowyebib.info/] – w różnych działach: biuletyn, konferencje, forum itp., Biblioteka 2.0 [http://forum.biblioteka20.pl/], serwisie digitalizacja.pl redaktorów: Tomasz Kalota, Rafał Raczyński, Paweł Rękar, Bartłomiej Siedlarz, Marcin Szala [http://digitalizacja.pl] czy prywatnej stronie z 2008 r. [http://digitalizacja.weebly.com/].

Niektóre teksty, jak przystało na ich temat, dostępne są w istniejących bibliotekach cyfrowych np. prezentacja *Biblioteki cyfrowe i dygitalizacja tekstów* (2010/2011) autorstwa Janusza S. Bienia [http://bc.klf.uw.edu.pl/162/1/BC10anons.pdf] dostępna jest w Bibliotece Cyfrowej Katedry Lingwistyki Formalnej UW, *Status prawny materiałów cyfrowych w Kujawsko-Pomorskiej Bibliotece Cyfrowej* czy artykuł *Polskie Biblioteki Cyfrowe – sieć wspomagająca edukację* Bożeny Bednarek-Michalskiej w Kujawsko-Pomorskiej Bibliotece Cyfrowej [http://kpbc.umk.pl/dlibra/ doccontent?id=48880&from=FBC, /doccontent?id=52381&from=FBC, / doccontent?id=52380&from=FBC], *Biblioteka cyfrowa w środowisku akademickim: analiza oprogramowania pod kątem przydatności do celów edukacyjnych: implementacja wybranych procedur* Bartłomieja Burby w Wielkopolskiej Bibliotece Cyfrowej [http://www.wbc.poznan.pl/Content/2561/praca_magisterska.html], a *Długoterminowe przechowywanie zasobów cyfrowych* Adama Dudczaka w Małopolskiej Bibliotece Cyfrowej [http://mbc.malopolska.pl/dlibra/doccontent?id=13392&from=FBC], *Biblioteka cyfrowa w środowisku wirtualnym* Marty Grabowskiej w Bibliologicznej Bibliotece Cyfrowej [http://bbc.uw.edu.pl/Content/4/02.pdf], *Jak przygotować obiekty do włączenia ich do biblioteki cyfrowej – doświadczenia Biblioteki Uniwersyteckiej* we Wrocławiu Tomasza Kaloty w Księgozbiorze Wirtualnym Federacji Fides [http://digital.fides.org.pl/dlibra/doccontent?id=716&from=FBC], *Biblioteki cyfrowe – wstęp do problematyki* Władysława Marka Kolasy w Śląskiej Bibliotece Cyfrowej [http://fbc.pionier.net.pl/id/oai:www.sbc.org.pl:10975], a *Rola biblioteki cyfrowej w pracy naukowej przyrodnika* Jakuba Nodzyńskiego w Pedagogicznej Bibliotece Cyfrowej [http://dlibra.up.krakow.pl:8080/dlibra/dlibra/doccontent?id=2009&from=FBC]. Tam też, tzn. w bibliotekach cyfrowych i na stronach systemu dLibra dostępne są podręczniki systemu dLibra, autorstwa Marcina Werli i Tomasza Parkoły [np. http://dl.psnc.pl/community/display/DLI/Dokumentacja;jsessionid=E6C045CCA38A3D8A177C7178AA3507DF].

Na stronach systemu dLibra [http://dlibra.psnc.pl/] przedstawiono także informacje, iż oprogramowanie może *służyć do budowy profesjonalnych repozytoriów obiektów cyfrowych. Repozytoria oparte o oprogramowanie dLibra mogą pełnić rolę systemów takich jak: biblioteki cyfrowe, instytucjonalne repozytoria dokumentów, cyfrowe archiwa.* Można więc mówić o tym, iż główne funkcje serwisów pociągają za sobą nazewnictwo, chociaż wydaje się, iż biblioteki cyfrowe traktowane są jako termin najszerszy. Dodatkowo stworzone serwisy podzielone zostały na instytucjonalne i regionalne, zgodnie z tym, czy tworzone były na potrzeby jednej, czy wielu, współpracujących ze sobą w regionie instytucji.

Kolejnym przykładem odniesienia się do bibliotek cyfrowych na stronach internetowych jest strona domowa Centrum Zaawansowanych Technologii Informacyjnych, powołanego przez Ministerstwo Nauki i Informatyzacji w celu wdrażania najnowocześniejszych technologii i idei w ramach budowy Społeczeństwa Informacyjnego. Przedstawiono tam biblioteki cyfrowe jako *to systemy informatyczne, których podstawowym zadaniem jest udostępnienie publikacji cyfrowych pewnej grupie odbiorców. Mechanizmy zawarte w bibliotece cyfrowej umożliwiają prosty i równocześnie efektywny dostęp do zgromadzonych zasobów, dostarczają dodatkowych informacji na temat tych zasobów (metadane) oraz umożliwiają wygodne zarządzanie biblioteką, obejmujące zarówno umieszczanie nowych publikacji lub organizowanie zawartości biblioteki w tematyczne kolekcje, jak i kontrolę dostępu do biblioteki* [http://www.czti.pl/index.php?pp=dla_nauki].

Jak widać jest to jeszcze jedna analiza systemu, a właściwie jego funkcjonalności.

5. Zakończenie

Z przytoczonych definicji bibliotek cyfrowych, elektronicznych i wirtualnych wywnioskować można, iż w początkowej fazie tworzenia systemów rozróżnienia między poszczególnymi typami bibliotek miały swoje odzwierciedlenie w zakresach nazw. Tak więc w latach dziewięćdziesiątych pojawiało się więcej prac związanych z rozważaniami terminologicznymi. W fazie dalszego rozwoju systemów oraz konkretnych projektów w piśmiennictwie angielskojęzycznym poprzestano na pojęciu 'digital library', a w polskim: 'biblioteka cyfrowa'. Zaznaczyć przy tym należy, iż prac terminologicznych i teoretycznych przez cały analizowany okres pojawiało się mniej, niż publikacji praktycznych, opisujących tworzone oraz gotowe projekty, metody digitalizacji obiektów, czy funkcjonowanie bibliotek cyfrowych. Jest to jednak charakterystyczne właściwie dla wszystkich dziedzin wiedzy, a nie tylko dla zagadnień związanych z systemami informacyjnymi.

Na koniec już powtórzyć jeszcze trzeba za zaleceniem Komisji Europejskiej z dnia 26. sierpnia 2006 r. główne powody dla których zagadnienia związane z bibliotekami cyfrowymi są ważne: *udostępnienie w Internecie materiałów pochodzących z różnych kultur i obszarów językowych ułatwi obywatelom docenienie własnego dziedzictwa kulturowego, a także dorobku innych krajów.*

Bibliografia

1. Atkins, Daniel E. (1993). *The changing curriculum and research for the digital library.* [Maszynopis], 5 s.
2. Balcerowicz, Danuta (2007). *Wejherowska Biblioteka Cyfrowa.* „Bibliotekarz" [nr] 5, s. 16.
3. Baldonado, Michelle Q. Wang (2000). *A user centered interface for information exploration in a heterogeneous digital library.* "Journal of the American Society for Information Science" 51, s. 297-310.
4. Barker, Philip (1994). *Electronic libraries – visions of the future.* „Electronic Library" vol. 12 nr 4, s. 221-30.
5. Barker, Philip (1997). *Electronic libraries of the future.* [W:] Kent, Alien; Lansour, Harold red. (1997). *Encyclopedia of library and information science,* t. 59. New York: M. Dekker, s. 119-153.
6. Bednarek-Michalska, Bożena (2002). *Digitalizacja zbiorów bibliotecznych i nie tylko.* „Notes Wydawniczy" nr 7/8, s. 49-52.
7. Bednarek-Michalska, Bożena (2006). *Kujawsko Pomorska Biblioteka Cyfrowa a standardy* [dok. elektr.] „EBIB – Elektroniczny Biuletyn Bibliotekarzy" nr 4, http://www.ebib.info/2006/74/michalska.php [odczyt: 2.10.2011].
8. *Biblioteka w kryzysie czy kryzys w bibliotece? : IV Konferencja Biblioteki Politechniki Łódzkiej, Łódź, 15-17 czerwca 2010 r. : materiały konferencyjne* (2010). [Aut. Hanna Bort-Nowak i in.]. Łódź: Biblioteka Politechniki Łódzkiej, 221, [4] s.
9. *Biblioteki XXI wieku. Czy przetrwamy? II Konferencja Biblioteki Politechniki Łódzkiej,* Łódź (Poland), 19-21 June 2006. Politechnika Łódzka [dok. elektr.] http://ebipol.p.lodz.pl/Content/1249/Biblioteki+XXI+wieku+...czy+przetrwamy_.pdf [odczyt: 20.10.2011].
10. Biliński, Lucjan oprac. (2008). *Interpelacja poselska w sprawie bibliotek cyfrowych i rozszerzaniu ich oferty tytułowej.* „Poradnik Bibliotekarza" [nr] 10, s. 17-19.
11. Billings, Harold (1991a). *The bionic library.* „Library Journal" 116 nr 17, s. 38-42.
12. Billings, Harold (1991b). *Romancing the information flow: solving the information crisis.* [Maszynopis] 12 s.
13. Bojar, Bożenna oprac. (2002). *Słownik encyklopedyczny informacji, języków i systemów informacyjno-wyszukiwawczych.* Warszawa: Wydaw. Stowarzyszenia Bibliotekarzy Polskich, 363 s.
14. Borbinha, José; Kunze, John; Spinazzè, Angela; Mutschke, Peter; Lieder, Hans-Jörg; Mabe, Michael; Dixson, Larry; Besser, Howard; Dean, Becky; Cathro, Warwick (2005). *Reference models for digital libraries: actors and roles.* „International Journal on Digital Libraries" vol. 5 nr 4, s. 325-330.
15. Borgman, Christine L. (1999). *What are digital libraries, who is building them, and why?* [W:] *CoLIS3, Digital libraries: interdisciplinary concepts, challenges and opportunities. Proceedings of the Third International Conference on the Conceptions of the Library and Information Science.*[Aparac, Tatjana; Saracevic, Tefko; Ingwersen, Peter; Vakkari, Pertti red.]. Dubrovnik, Croatia: Zagreb: Lokve, s. 23-38.
16. Bożek, Agata (2009). *Europeana – cyfrowa kolekcja europejskiego dziedzictwa kultury.* „Zagadnienia Informacji Naukowej" nr 1, s. 130-139.
17. Brandt, D. Scott (2001). *Reference, mental models and teaching technology.* „The Reference Librarian" vol. 74, s. 37-48.

18. Butkiewicz, Marian (2003). *Biblioteka wirtualna: mit staje się rzeczywistością* [dok. elektr.] „EBIB – Elektroniczny Biuletyn Bibliotekarzy" nr 5, http://ebib.oss.wroc.pl/2003/45/butkiewicz.php [odczyt: 2.10.2011], Biblioteki wyższych szkół niepublicznych – problemy i nowe rozwiązania organizacyjne. Sprawozdania.
19. Buzdygan, Dorota; Lipińska, Dorota (2008). *Biblioteka Cyfrowa Politechniki Krakowskiej – repozytorium i forma udostępniania publikacji z zakresu nauk technicznych.* [W:] Ganińska, Halina red. (2008). *Informacja dla nauki a świat zasobów cyfrowych*. Poznań: Biblioteka Główna Politechniki Poznańskiej, s. 35-45.
20. Candela, Leonardo; Castelli, Donatella; Pagano, Pasquale (2011). *History, evolution and impact of digital libraries.*[W:] Iglezakis, Ioannis; Synodinou, Tatiana-Eleni; Kapidakis, Sarantos red. (2011). *E-Publishing and digital libraries: legal and organizational issues.* Hershey; New York: Information Science Reference, IGI Global, s. 1-30.
21. Candela, Leonardo; Nardi, A. red. (2011). *The digital library reference model* [aut.: L. Candela; G. Athanasopoulos; D. Castelli; K.El Raheb; P. Innocenti; Y. Ioannidis; A. Katifori; A. Nika; G. Vullo; S. Ross] [dok. elektr.] http://www.dlorg.eu/index.php/outcomes/reference-model; http://bscw.research-infrastructures.eu/pub/bscw.cgi/d222816/D3.2b%20Digital%20Library%20Reference%20Model.pdf [odczyt: 2.10.2011].
22. Casarosa, Vittore (2007). *DELOS reference model for digital libraries* [dok. elektr.] http://elag2007.upf.edu/papers/casarosa.pdf [odczyt: 2.06.2011].
23. Catenazzi, Nadia; Sommaruga, Lorenzo (1995). *Hyper-lib: a formal model for an electronic library based on hyper-books.* „Journal of Documentation" vol. 51 nr 3, s. 244-270.
24. Chmielewska-Gorczyca, Ewa (1996a). *Biblioteka wirtualna – wizja, czy rzeczywistość.* „Przegląd Biblioteczny" nr 2/3, s. 117-131.
25. Chmielewska-Gorczyca, Ewa (1996b). *Ku bibliotece wirtualnej.* „Zagadnienia Informacji Naukowej" nr 1, s. 3-13.
26. Cleveland, Gary (1998). *Digital fibraries: definitions, issues and challenges.* [dok. elektr.] http://archive.ifla.org/VI/5/op/udtop8/udtop8.htm [odczyt: 2.10.2011].
27. Cool, Colleen; Williams, Martha E. (2000). *A typology of digital libraries and their user communities.* „Proceedings of the National Online Meeting" vol. 21, s. 61-68.
28. Czermiński, Jurand B. (1997). *Od biblioteki elektronicznej do biblioteki wirtualnej.* [W:] Pindlowa, Wanda red. (1997). *Elektroniczna przyszłość bibliotek akademickich*. Kraków: Wyd. Uniwersytetu Jagiellońskiego, s. 158-167.
29. Czermiński, Jurand B. (2002). *Cyfrowe środowisko współczesnej biblioteki.* Gdańsk: Wydaw. Uniwersytetu Gdańskiego, 267, [3] s. [Wersja elektroniczna: http://panda.bg.univ.gda.pl/~jurand/cyfrowe_srodowisko/; odczyt: 20.09.2011].
30. Derfert-Wolf, Lidia (2010). *Nowy model komunikacji naukowej.* „Forum Akademickie" R. 17 nr 5, s. 62-64.
31. *Digital libraries: interdisciplinary concepts, challenges and opportunities.* Proceedings of the Third International Conference on the Conceptions of the Library and Information Science. Ed. Aparac, Tatjana; Saracevic, Tefko; Ingwersen, Peter; Vakkari, Pertti. Dubrovnik, Croatia: Zagreb: Lokve, 400 s.
32. Digital Library Federation (1998). *A working definition of digital library.* [dok.elektr.] http://www.diglib.org/about/dldefinition.htm [odczyt: 18.11.2010]
33. Domowicz, Iwona; Kalota, Tomasz [i in.] (2006). *Biblioteka Cyfrowa Uniwersytetu Wrocławskiego – cele, zadania, wytyczne i organizacja pracy* [dok. elektr.] „EBIB – Elektroniczny Biuletyn Bibliotekarzy" nr 4, http://www.ebib.info/2006/74/domowicz_in.php [odczyt: 2.10.2011].

34. Drabek, Aneta (2008). *Cyfrowy warsztat pracy bibliotekarza.* „Bibliotekarz" [nr] 7/8, s. 3-8.
35. Drabenstott, Karen (1993). *Analytical review of the library of the future.* Washington: Council on Library Resources. [Wersja elektroniczna: http://ylime.people.si.umich.edu/future/clrFutureRpt.pdf].
36. Drutar, Michael; Coleman, Charles; Fox, Edward (2006). *Creating a multi disciplinary digital library in the 5S framework. Poster for JCDL 2006,* June 11-16, Chapel Hill, NC, s. 351.
37. Eagle, Mark (1992). *The librarian of the future: image storage and transmission.* [W:] *Conference on integrated online library systems.* 7th ed. Medford: Learned Information, s. 99-103.
38. Feigenbaum, E.A. (1989). *Toward the library of the future.* „Long Range Planning" 22 1, s. 118-123.
39. Filipowicz, Anna (1999). *Biblioteka wirtualna, biblioteka cyfrowa...* [dok. elektr.] „EBIB – Elektroniczny Biuletyn Bibliotekarzy" nr 8, http://ebib.oss.wroc.pl/arc/e008-01.html [odczyt: 2.10.2011].
40. Fox, Edward A. (1999). *The 5S Framework for Digital Libraries and Two Case Studies: NDLTD and CSTC.* [W:] Proceedings of NIT99, Taipei, Taiwan, August 1999 [Wersja elektroniczna: http://snidely.dlib.vt.edu:8081/dspace/bitstream/2340/32/1/nit99fox.doc].
41. Fox, Edward A. (1999). *The Digital Libraries Initiative – update and discussion.* „Bulletin of the America Society of Information Science" vol. 26 nr 1 [Wersja elektroniczna: http://www.asis.org/Bulletin/Oct-99/index.html].
42. Fox, Edward A.; Akscyn, Robert M.; Furuta, Richard K.; Leggett, John J. (1995). *Digital libraries.* „Communications of the ACM" vol. 38 nr 4, s. 23-28.
43. Fox, Edward A.; Gonçalves, Marcos André; Kipp, Neill A. (2001). Digital libraries. [W:] Adelsberger, Heimo; Collis, Betty; Pawlowski, Jan red. (2001). *Handbook on information technologies for education and training.* Heidelberg: Springer, s. 623-641.
44. *A Framework of Guidance for Building Good Digital Collections – the 3rd edition* (2007) [dok. elektr.] Institute for Museum and Library Services http://framework.niso.org/ [odczyt: 2.02.2011].
45. Fuhr, Norbert [i in.], *Evaluation of digital libraries* [dok. elektr.] http://www.scribd.com/doc/185523/Evaluation–of–Digital–Libraries–Fulltext [odczyt: 2.10.2011].
46. Fuhr, Norbert; Hansen, Preben; Mabe, Michael; Micsik, Andras; Sølvbergh, Ingeborg (2001). *Digital libraries: a generic classification and evaluation scheme.* http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.23.6160&rep=rep1&type=pdf
47. Ganińska, Halina red. (1998). *Świat biblioteki elektronicznej w klasycznej bibliotece naukowej: możliwości rozwoju, uwarunkowania i ograniczenia: materiały konferencyjne, Poznań, 19-20 marca 1998.* Poznań: Biblioteka Główna Politechniki Poznańskiej, 270 s.
48. Ganińska, Halina red. (2008). *Informacja dla nauki a świat zasobów cyfrowych.* Poznań: Biblioteka Główna Politechniki Poznańskiej, [10], 189 s. [Materiały z konferencji, 11-13 września 2008 r., Świnoujście].
49. Gapen, D. Kaye (1993). *The virtual library: knowledge, society, and the librarian.* [W:] Saunders, Laverna M. red. (1993). *The virtual library: visions and realities.* Westport: Meckler Corporation, s. 1-14.
50. Gill-Tarnowska, Dorota (2008). *Pomorskie zasoby cyfrowe na terenie Vorpommern (Pomorza Przedniego).* [W:] Ganińska, Halina red. (2008). *Informacja dla nauki a świat zasobów cyfrowych.* Poznań: Biblioteka Główna Politechniki Poznańskiej, s. 55-63.

51. Gobinda G. Chowdhury, Sudatta Chowdhury (1999). *Progress in documentation: digital library research: major issues and trends*. „Journal of Documentation" vol. 55 nr 4, s. 409-448.
52. Gobinda G. Chowdhury, Sudatta Chowdhury (2003). *Introduction to digital libraries*. Library Assn Pub Ltd; Facet, 304 s.
53. Gonçalves, Marcos André, Fox, Edward A.; Watson, Layne T.; Kipp, Neill A. (2004). *Streams, Structures, Spaces, Scenarios, Societies (5S): A formal model for digital libraries* [dok. elektr.] http://www.dlib.vt.edu/projects/5S-Model/5s6.pdf [odczyt: 2.10.2011].
54. Gonçalves, Marcos André; Fox, Edward A. (2002). *5SL: a language for declarative specification and generation of digital libraries.* [W:] *Proceedings of the 2nd ACM/IEEE-CS Joint Conference on Digital Libraries* (Portland, Oregon, USA, July 14-18, 2002). JCDL '02. ACM, New York, s. 263-272.
55. Gonçalves, Marcos André; Fox, Edward A.; Watson, Layne T. (2008). *Towards a digital library theory: a formal digital library ontology*. „International Journal on Digital Libraries" vol. 8 nr 2, s. 91-114.
56. Gonçalves, Marcos André; Fox, Edward A.; Watson, Layne T.; Kipp, Neill A. (2004). *Streams, structures, spaces, scenarios, societies (5s): a formal model for digital libraries.* „ACM Trans. Inf. Syst." 22 nr 2, s. 270-312.
57. Gonçalves, Marcos André; Fox, Edward A.; Watson, Layne T.; Kipp, Neill A. (2003). *Streams, structures, spaces, scenarios, societies (5s): a formal model for digital libraries.* [dok. elektr.] „Technical Report TR-03-04. Computer Science Virginia tech." http://dl.acm.org/citation.cfm?doid=984321.984325 [odczyt: 2.10.2011]
58. Górny, Mirosław (1998). *Biblioteki elektroniczne – próba kategoryzacji.* [W:] Ganińska, Halina red. (1998). *Świat biblioteki elektronicznej w klasycznej bibliotece naukowej : możliwości rozwoju, uwarunkowania i ograniczenia : materiały konferencyjne, Poznań, 19-20 marca 1998*. Poznań: Biblioteka Główna Politechniki Poznańskiej, s. 47-53.
59. Górny, Mirosław; Niskich, Jan Andrzej (2003). *Wielkopolska Biblioteka Cyfrowa.* „Bibliotekarz" [nr] 7/8, s. 6-9.
60. Górny, Mirosław; Wierzchoń, Piotr (2010). *Polish digital libraries as a philologist's tool: based on 666 adjectives from the Digital Library of Wielkopolska*. Poznań: Institute of Linguistics. Adam Mickiewicz University, 261 s.
61. Górski, Marek M. (1998). *Biblioteka elektroniczna – oczekiwane kierunki rozwoju (na podstawie przeprowadzonych badań).* [W:] Ganińska, Halina red. (1998). *Świat biblioteki elektronicznej w klasycznej bibliotece naukowej : możliwości rozwoju, uwarunkowania i ograniczenia : materiały konferencyjne, Poznań, 19-20 marca 1998*. Poznań: Biblioteka Główna Politechniki Poznańskiej, s. 25-30.
62. Grabowska, Marta (2007). *Biblioteka cyfrowa w środowisku wirtualnym: nowe wyzwanie dla katalogów bibliotek w erze dokumentów elektronicznych* [dok. elektr.] http://bbc.uw.edu.pl/Content/4/02.pdf [odczyt: 2.10.2011].
63. Greenstein, Daniel I.; Thorin, Suzanne Elizabeth (2002). *The Digital Library: a biography*. Digital Library Federation [Wersja elektroniczna: http://www.clir.org/PUBS/reports/pub109/pub109.pdf].
64. Greenstein, Daniel; Thorin, Suzanne E. (2002). *The Digital Library: a biography* [dok. elektr.] http://www.clir.org/pubs/reports/pub109/pub109.pdf [odczyt: 2.10.2011].
65. Griffin, Stephen M.; McMillan, Gail; Seamans, Nan (1998). *Digital Library Definition for DLI2* [dok. elektr.] http://scholar.lib.vt.edu/DLI2/defineDL.html [odczyt: 2.10.2011].

66. Grygowski, Dariusz (2007). *Prawo autorskie w kontekście funkcjonowania bibliotek cyfrowych.* [W:] Zając, Michał; Zybert, Elżbieta Barbara red. (2007). *Biblioteki w systemie kultury jednoczącej się Europy: praca zbiorowa.* Warszawa: Stowarzyszenie Bibliotekarzy Polskich, s. 154-176.

67. Harter, Stephen P. (1996). *What is a digital library? Definitions, Content, and Issues, KOLISS DL ,96*: International Conference on Digital Libraries and Information Services for the 21st Century, September 10-13, 1996, Seoul, Korea [dok. elektr.] http://php.indiana.edu/~harter/korea-paper.htm [odczyt: 2.10.2011].

68. Hayes, Robert M. (1999). *The economics of digital libraries.* [W:] *CoLIS3, Digital libraries: interdisciplinary concepts, challenges and opportunities. Proceedings of the Third International Conference on the Conceptions of the Library and Information Science.* [Aparac, Tatjana; Saracevic, Tefko; Ingwersen, Peter; Vakkari, Pertti red.]. Dubrovnik, Croatia: Zagreb: Lokve [Wersja elektroniczna: http://www.ime.usp.br/~cesar/simposio99/hayes.htm].

69. Heliński, Marcin (2006). *Biblioteka cyfrowa jako otwarte, internetowe repozytorium publikacji.* [W:] *III konferencja: Internet w bibliotekach: zasoby elektroniczne: podaż i popyt. Wrocław, 12-14 grudnia 2005.* Wrocław: Stowarzyszenie Bibliotekarzy Polskich, K[omisja] W[ydawnictw] E[lektronicznych], Redakcja „Elektronicznej Biblioteki" [Wersja elektroniczna: http://www.ebib.info/publikacje/matkonf/iwb3/artykul.php?f]

70. Hill, Linda H.; Carver, Larry; Larsgaard, Mary; Dolin, Ron; Smith, Terence R.; Frew, James; Rae, Mary-Anna (2000). *Alexandria digital library: user evaluation studies and system design.* „Journal of the American Society for Information Science" vol. 51 nr 3, s. 246-259.

71. Hollender, Henryk (1999). *World Wide Lib ? O konsorcjach bibliotecznych teraz i w przyszłości.* „Biuletyn Głównej Biblioteki Lekarskiej" t. 45 nr 359, s. 29-38.

72. Hollender, Henryk (2004). *Biblioteki cyfrowe w Polsce i BWN: w poszukiwaniu drogi racjonalnego rozwoju* [dok. elektr.] http://bg.p.lodz.pl/konferencja2004/pelne_teksty/hollender.doc [odczyt: 2.10.2011].

73. Hollender, Henryk (2005). *Biblioteki cyfrowe w Polsce : warunki i zagrożenia* [dok. elektr.] „EBIB – Elektroniczny Biuletyn Bibliotekarzy" nr 4, http://ebib.oss.wroc.pl/2005/65/hollender.php [odczyt: 2.10.2011].

74. Hollender, Henryk (2005-2008). *Cyfrowe okolice.* „Poradnik Bibliotekarza" 2005 nr 9, s. 13-14; nr 10, s. 12; nr 11, s. 15-16; nr 12, s. 19-20; 2006 nr 1, s. 19-20; nr 2, s. 21; nr 3, s. 11; nr 4, s. 13-14; nr 5, s. 17; nr 6, s. 16-17; nr 7/8, s. 14; nr 9, s. 18; nr 10, s. 16-17; nr 11, s. 17-18; 2007, nr 1, s. 22; nr 3, s. 13-14; nr 4, s. 11; nr 6, s. 13-14; nr 7/8, s. 20; nr 9, s. 17; nr 11, s. 19-20; nr 12, s. 15; 2008: nr 1, s. 16-17; nr 3, s. 15-16; nr 4, s. 11-12; nr 5, s. 13-14; nr 7/8, s. 22-23; nr 10, s. 16-17; nr 12, s. 18; nr 30, s. 17-18.

75. Hollender, Henryk red. (2010). *Cyfrowy świat dokumentu: wydawnictwa, biblioteki, muzea, archiwa.* Warszawa: Centrum Promocji Informatyki, 274 s.

76. Howorka, Bolesław (2006). *Biblioteka cyfrowa i prawo autorskie.* „Bibliotekarz" [nr] 10, s. 7-11.

77. *i2010 Digital Library Initiative. High Level Expert Group on Digital Libraries* (2009). Final Report "Digital Libraries: Recommendations and Challenges for the Future". [dok. elektr.] http://ec.europa.eu/information_society/activities/digital_libraries/doc/hleg/reports/hlg_final_report09.pdf [odczyt: 2.10.2011].

78. Jagodzinski, Filip; Cassel, Lillian; Goncalves, Marcos A.; Fox, Edward A. (2003). *Digital library analysis: advantages of using the XML log standard versus the apache server logs. Poster for JCDL'2003*, May 27-31, 2003, Houston, TX.

79. Jastrząb, Łukasz (2007). *Praktyczne zastosowanie oprogramowania dLibra w bibliotekach cyfrowych.* „Zagadnienia Informacji Naukowej" nr 2, s. 95-[97].
80. Jastrząb, Łukasz (2007). *Wielkopolska Biblioteka Cyfrowa jako narzędzie edukacyjne.* „Dyrektor Szkoły" nr 3, s. 26, 29.
81. Jastrząb, Łukasz (2008). *Doświadczenia poznańskie w budowaniu biblioteki cyfrowej na przykładzie prac Biblioteki Kórnickiej Polskiej Akademii Nauk.* „Przegląd Biblioteczny" R. 76 z. 1, s. [79]-85.
82. Jones, Steve; Cunningham, Sally Jo; McNab, Rodger J. (1998). *Usage analysis of digital library.* [W:] *Proceedings of the third ACM conference on digital libraries.* New York: ACM, s. 293-294.
83. Kahn, Robert E.; Cerf, Vinton G. (1988). *The Digital Library Project.* Volume I: The World of Knowbots, (DRAFT): An Open Architecture For a Digital Library System and a Plan For Its Development. Reston: Corporation for National Research Initiatives [Wersja elektroniczna: http://www.cnri.reston.va.us/kahn-cerf-88.pdf].
84. Kalota, Tomasz (2006). *Marzenie o polskim systemie rozproszonych bibliotek cyfrowych.* [dok. elektr.] „EBIB – Elektroniczny Biuletyn Bibliotekarzy" nr 4, http://www.ebib.info/2006/74/kalota.php [odczyt: 2.10.2011].
85. Karwasińska, Emilia; Rychlik, Małgorzata (2008). *Model i etapy tworzenia repozytorium instytucjonalnego na podstawie badań własnych i doświadczeń bibliotek zagranicznych.* [W:] Ganińska, Halina red. (2008). *Informacja dla nauki a świat zasobów cyfrowych.* Poznań: Biblioteka Główna Politechniki Poznańskiej, s. 10-20.
86. Kent, Alien; Lansour, Harold red. (1968-2003). *Encyclopedia of library and information science.* 77 t. New York: M. Dekker.
87. Kibby, Mark; Evans, Nancy H. (1989). *The network is the library.* „EDUCOM Review" 24 nr 3, s. 15-20.
88. Kocikowski, Andrzej (2003). *Akademicka Biblioteka Internetowa* (ABI) [dok. elektr.] http://ebib.oss.wroc.pl/2003/51/kocikowski.php [odczyt: 2.10.2011].
89. Kocójowa, Maria (2005). *Znaczenie dla badań naukowych digitalizacji obiektów kultury i zamieszczania ich artefaktów w Internecie* [dok. elektr.] http://skryba.inib.uj.edu.pl/wydawnictwa/e01/kocojowa.pdf [odczyt: 2.10.2011].
90. Kocójowa, Maria; Pindlowa, Wanda (1999). *The need of digital library for LIS research in Poland.* [W:] *CoLIS3, Digital libraries: interdisciplinary concepts, challenges and opportunities. Proceedings of the Third International Conference on the Conceptions of the Library and Information Science.* [Aparac, Tatjana; Saracevic, Tefko; Ingwersen, Peter; Vakkari, Pertti red.]. Dubrovnik, Croatia: Zagreb: Lokve, s. 322-328.
91. *Komunikat Komisji do Parlamentu Europejskiego, Rady, Europejskiego Komitetu Ekonomiczno-Społecznego oraz Komitetu Regionów: Biblioteki cyfrowe - Komisja Wspólnot Europejskich, Bruksela, 30 września 2005 r.* [dok. elektr.] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2005:0465:FIN:PL:PDF [odczyt: 2.10.2011].
92. Kowalska, Małgorzata (2005). *Cyfrowa ochrona dorobku dokumentalnego ludzkości w programach i wytycznych Europejskiej Komisji Ochrony i Dostępu, IFLA i UNESCO.* „Przegląd Biblioteczny" R. 73 z. 2, s. [186]-200.
93. Kowalska, Małgorzata (2007). *Dygitalizacja zbiorów bibliotek polskich.* Warszawa: Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich, 293, [2] s.
94. Krzemińskia, Wanda; Nowak Piotr red. (2004) *Cyfryzacja w procesach komunikowania* Poznań: Sorus Wydawnictwo i Drukarnia, 244 s.

95. Kubiak, Barbara; Bogiel-Kuźmicka Anna (2008). *Podlaska Biblioteka Cyfrowa.* [W:] Ganińska, Halina red. (2008). *Informacja dla nauki a świat zasobów cyfrowych.* Poznań: Biblioteka Główna Politechniki Poznańskiej, s. 46-54.
96. Kuna, Terry; Cleveland, Gary (1998). *The digital library: myths and challenges.* [dok. elektr.] „IFLA Journal" 24: 107, http://ifl.sagepub.com/content/24/2/107.citation [odczyt: 2.10.2011].
97. Lancaster, F. Wilfrid (1978). *Toward paperless information systems.* New York: Academic Press, XII, 179 s.
98. Landoni, Monica; Catenazzi, Nadia; Gibb, Forbes (1993). *Hyper-books and visual-books in an electronic library.* „Electronic Library" vol. 11 nr 3, s. 75-186.
99. Levy, David M.; Marshall, Catherine C. (1994). *What Color was George Washington's White Horse? A Look at Assumptions Underlying Digital Libraries. Proceedings of the First Annual Conference on the Theory and Practice of Digital Libraries.* June 19-21, 1994, College Station, Texas. [Wersja elektroniczna: http://www.csdl.tamu.edu/DL94/paper/levy.html].
100. Licklider, J.C.R. (1965). *Libraries of the Future.* Cambridge: M.I.T. Press, XVII, 219 s.
101. Lis, Remigiusz (2007). *Śląska Biblioteka Cyfrowa.* [W:] Górska, Elżbieta red. (2007). *Praktyczne aspekty automatyzacji bibliotek publicznych: materiały z VII Ogólnopolskiej Konferencji pt: "Automatyzacja bibliotek publicznych: Warszawa 22-24 listopada 2006 r.* Warszawa: Stowarzyszenie Bibliotekarzy Polskich, s. 198-201.
102. Łozowska, Anna (2008). *Zachodniopomorska Biblioteka Cyfrowa – początek drogi.* [W:] Ganińska, Halina red. (2008). *Informacja dla nauki a świat zasobów cyfrowych.* Poznań: Biblioteka Główna Politechniki Poznańskiej, s. 64-66.
103. Łozowska, Anna; Stylińska, Jolanta (1998). *Biblioteka wirtualna – alternatywą czy koniecznością badacza końca XX wieku.* [W:] Ganińska, Halina red. (1998). Świat *biblioteki elektronicznej w klasycznej bibliotece naukowej : możliwości rozwoju, uwarunkowania i ograniczenia : materiały konferencyjne, Poznań, 19-20 marca 1998.* Poznań: Biblioteka Główna Politechniki Poznańskiej, s. 42-46.
104. Lynch, Clifford A. (1991). *"Visions of electronic libraries".* [W:] Simova, Filomena red. (1991). *The Bowker annual library and book trade almanac.* New Providence: Bowker, s. 75-82.
105. Ma, Yi ; Gonçalves, Marcos André; Fox, Edward A. (2007). *Personal digital library: pim through a 5s perspective.* [W:] *Proceedings of the First Ph.D. Workshop in CIKM, PIKM 2007, Sixteenth ACM Conference on Information and Knowledge Management, CIKM 2007, Lisbon, Portugal, November 9, 2007. ACM 2007,* s. 117-124.
106. Ma, Yi ; Gonçalves, Marcos André; Fox, Edward A. (2007). *PIM through a 5S perspective.* [W:] *Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital Libraries,* Vancouver, BC, Canada, June 18-23, 2007, 491 s.
107. Machalska-Garbacz, Anna (1997). *Nasza przyszłość: biblioteka elektroniczna czy biblioteka cyfrowa?* „Przegląd Biblioteczny" nr 2/3, s. 203-210.
108. Makowski, Tomasz (2008). *Biblioteka Narodowa w dobie technologii cyfrowej.* „Poradnik Bibliotekarza" [nr] 9, s. 13-17.
109. Marchionini, Gary (1998). *Research and development in digital libraries.* [W:] Kent, Alien; Lansour, Harold red. (1997). *Encyclopedia of library and information science,* t. 63. New York: M. Dekker, s. 259-279.
110. Marchionini, Gary; Plaisant, Catherine; Komlodi, Anita (1998). *Interfaces and tools for the Library of Congress National Digital Library Program.* „Information Processing and Management" 34(5), s. 535-555.

111. Marcinkiewicz, Lilia (2007). *Ku bibliotece cyfrowej – dwadzieścia lat doświadczeń w komputeryzacji Książnicy Pomorskiej.* „Bibliotekarz" [nr] 7/8, s. 7-11.

112. Matwiejczuk, Wanda (1996). *Biblioteka elektroniczna dzisiaj.* „Bibliotekarz" nr 9, s. 15-16.

113. Mazurek, Cezary [i in.] (2008). *Nowoczesne biblioteki cyfrowe w środowisku rozproszonych usług atomowych.* [W:] Ganińska, Halina red. (2008). *Informacja dla nauki a świat zasobów cyfrowych.* Poznań: Biblioteka Główna Politechniki Poznańskiej, s. 2-8.

114. Mazurek, Cezary; Stroiński, Maciej; Węglarz, Jan red. (2009). *Polskie Biblioteki Cyfrowe 2008 : materiały z konferencji zorganizowanej w dniach 24-25 listopada 2008 r. przez Bibliotekę Kórnicką PAN, Poznańską Fundację Bibliotek Naukowych, Poznańskie Centrum Superkomputerowo-Sieciowe.* Poznań: Ośrodek Wydawnictw Naukowych, 105 s.

115. Mazurek, Cezary; Stroiński, Maciej; Węglarz, Jan red. (2010). *Polskie Biblioteki Cyfrowe 2009 : materiały z konferencji zorganizowanej dnia 9 grudnia 2009 r. przez Bibliotekę Kórnicką PAN, Poznańską Fundację Bibliotek Naukowych, Poznańskie Centrum Superkomputerowo-Sieciowe.* Poznań: Ośrodek Wydawnictw Naukowych, 101 s.

116. Mazurek, Cezary; Stroiński, Maciej; Węglarz, Jan red. (2011). *Polskie Biblioteki Cyfrowe 2010 : materiały z konferencji zorganizowanej w dniach 20-21 października 2010 roku przez Bibliotekę Kórnicką PAN, Poznańską Fundację Bibliotek Naukowych, Poznańskie Centrum Superkomputerowo-Sieciowe.* Poznań: Ośrodek Wydawnictw Naukowych, s. 101 s.

117. Mazurek, Jolanta (2007). *Rozproszony system bibliotek cyfrowych w Polsce – doświadczenia i wyzwania na przyszłość.* „Bibliotekarz" [nr] 10, s. 2-5.

118. Merta, Tomasz (2008). *Odpowiedź podsekretarza stanu w Ministerstwie Kultury i Dziedzictwa Narodowego – z upoważnienia ministra – na interpelację nr 986 w sprawie rozwoju bibliotek cyfrowych w Polsce.* [dok. elektr.] http://orka2.sejm.gov.pl/IZ6.nsf/main/5A954AFA [odczyt: 2.10.2011].

119. Migoń, Krzysztof; Skalska-Zlat, Marta red. (2009). *Uniwersum piśmiennictwa wobec komunikacji elektronicznej.* Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego, 413 s.

120. Miksa, Francis L.; Doty, Philip (1994). *Intellectual Realities and the Digital Library. Proceedings of the First Annual Conference on the Theory and Practice of Digital Libraries. June 19-21, 1994, College Station, Texas.* [Wersja elektroniczna: http://www.csdl.tamu.edu/DL94/paper/miksa.html].

121. Ministerstwo Kultury i Dziedzictwa Narodowego. Departament Mecenatu Państwa (2010). *Biblioteki cyfrowe i książka elektroniczna (e-książka).* Raport Ministerstwa Kultury i Dziedzictwa Narodowego przygotowany dla sejmowej komisji kultury [dok. elektr.] http://docs.google.com/viewer?a=v&q=cache:pLVI8gIRWUIJ:www.nowoczesnapolska.org.pl/wordpress/wp-content/uploads/2010/03/E-book-raport-dla-Komisji.pdf+biblioteka+born+digital&hl=pl&gl=pl&pid=bl&srcid=ADGEESjcSfEhkzj_n3c_swKud5jH4UB7DEa8441aPtfcPQMDOxLNDR4cfUI4pDvP9qWcX0j30izTA9-xnuZBHTjD-3-izi1WbrHQ1c9vq55ANfxr1y9Nzc8cRwAdUKRxQVCyg8t0fP8Y&sig=AHIEtbQRIT9eKBWTgXGI_kCAI4RCmIM8Kw [odczyt: 2.10.2011].

122. Mitchell, Steve (1999). *Interface design considerations in libraries.* „Science and Technology Libraries" 17(3/4), s. 131-181.

123. Mocydlarz, Mirosława (2000). *Udostępnianie informacji naukowej na nośnikach elektronicznych* [dok. elektr.] Horyzonty Biuletyn Porozumienia „Biblioteka z Horyzontem" nr 2, http://www.pfsl.poznan.pl/horyzonty/nosniki/spis.html [odczyt: 2.10.2011].

124. Moreira, Barbara L.; Gonçalves, Marcos André; Fox, Edward A., Watson, Layne T. (2007). *What is a good digital library?: a quality model for digital libraries.* „Inf. Process. Manage". 43 (5), s. 1416-1437.

125. Moreira, Barbara L.; Gonçalves, Marcos André; Laender, Alberto H.F.; Fox, Edward A. (2007). *Evaluating Digital Libraries with 5SQual.* [W:] *Research and Advanced Technology for Digital Libraries, 11th European Conference, ECDL 2007, Budapest, Hungary, September 16-21, 2007.* Heidelberg: Springer, s. 466-470 [Wersja elektroniczna: http://www.springerlink.com/content/l1382154kr346816/fulltext.pdf].
126. Moreira, Barbara Lagoeiro; Goncalves, Marcos Andre; Laender, Alberto H.F.; Fox, Edward A. (2006). *5SQual: a quality assessment tool for digital libraries.* „Proc. WDL", 20 Oct. 2006, Florianopolis, SC, Brasil, 10 s.
127. Moreira, Barbara Lagoeiro; Goncalves, Marcos Andre; Laender, Alberto H.F.; Fox, Edward A. (2009). *Automatic Evaluation of Digital Libraries with 5SQual.* „Journal of Informetrics" 3 (2), s.102-123.
128. Murthy, Uma, Gorton, Douglas, Torres, Ricardo, Gonçalves, Marcos André; Fox, Edward A., Delcambre, Lois (2007). *Extending the 5S Digital Library (DL) Framework: From a Minimal DL towards a DL Reference Model: paper presented at the First Digital Library Foundations Workshop – held in conjunction with the 2007 Joint Conference on Digital Libraries (Vancouver, Canada, June 18-24, 2007).* [Wersja elektroniczna: http://www.dlib.vt.edu/projects/5S-Model/extending5S.pdf].
129. Murzynowska, Danuta; Potęga, Joanna (2006). *Nie tylko biblioteka cyfrowa... rzecz o bibliotecznych wystawach wirtualnych* [dok. elektr.] „EBIB – Elektroniczny Biuletyn Bibliotekarzy" nr 4, http://www.ebib.info/2006/81/a.php?murzynowska_potega [odczyt: 2.10.2011].
130. Nahotko, Marek (2004). *Ontologiczne modelowanie rzeczywistości w bibliotekach cyfrowych.* „Zeszyty Naukowe Wydziału ETI Politechniki Gdańskiej. Technologie Informacyjne" nr 5, s. 941-948.
131. Nahotko, Marek (2006). *Zasady tworzenia bibliotek cyfrowych* [dok. elektr.] „EBIB – Elektroniczny Biuletyn Bibliotekarzy" nr 4, http://www.ebib.info/2006/74/nahotko.php [odczyt: 2.10.2011].
132. Nahotko, Marek (2007). *Cyfrowa nauka – cyfrowe publikacje – cyfrowe biblioteki.* „Przegląd Biblioteczny" R. 75 z. 1, s. [7]-28.
133. Niezgódka, Marek (2004). *Polska Biblioteka Wirtualna Nauki : program i jego perspektywy.* „Archeion" t. 107, s. 354-362.
134. Nowakowski, Ryszard (2009). *Digitalizacja a zbiory utracone i rozproszone : analiza problemu na przykładzie projektu cyfrowego scalenia zbioru prasy polskiej (1801-1939) Zakładu Narodowego im. Ossolińskich, rozdzielonego po II wojnie światowej.* [W:] Migoń, Krzysztof; Skalska-Zlat, Marta red. (2009). *Uniwersum piśmiennictwa wobec komunikacji elektronicznej.* Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego, s. 343-360.
135. Nowińska, Alina (2003). *Niektóre problemy przyswajania angielskiej terminologii z zakresu bibliotekarstwa i informacji naukowej na przykładzie „Bibliografii Analitycznej Bibliotekoznawstwa i Informacji Naukowej".* „Przegląd Biblioteczny" R. 71 z. 1/2, s. 55-70.
136. Nürnberg, Peter J., Furuta, Richard K.; Leggett, John J.; Marshall, Catherine C.; Shipman III, Frank M. (1995). *Digital libraries: issues and architectures.* [W:] *Proceedings of the Second Annual Conference on the Theory and Practice of Digital Libraries.* Austin, Texas, June 11-13, s. 147-153 [Wersja elektroniczna: http://www.csdl.tamu.edu/DL95/papers/nuernberg/nuernberg.html].
137. Osińska, Veslava (2006). *Przybliżenie semantyczne w wizualizacji informacji w Internecie i bibliotekach cyfrowych.* [dok. elektr.] „EBIB – Elektroniczny Biuletyn Bibliotekarzy" nr 7 http://ebib.info/2006/77/osinska.php [odczyt: 2.10.2011].

138. Osińska, Veslava (2008). *Wizualizacja i mapowanie przestrzeni danych w bibliotekach cyfrowych.* „Toruńskie Studia Bibliologiczne" nr 1, s. 167-176.
139. Paradowski, Dariusz oprac. i red. (2010). *Digitalizacja piśmiennictwa.* [Aut. Maciej Hnydka i in.; red. merytoryczna Agnieszka Konopka]. Warszawa: Biblioteka Narodowa, 212, [1] s.
140. Park, Soyeon (2000). *Usability, user preferences, effectiveness, and user behaviors when searching individual and integrated full-text databases: implications for digital libraries.* „Journal of the American Society for Information Science" vol. 51 nr 5, s. 456-468.
141. Paszko, Artur (2006). *Impresje na temat hybrydyzacji usług bibliotek publicznych w Polsce.* „Bibliotekarz" [nr] 10, s. 2-6.
142. Payette, Sandra D.; Rieger, Oya Y. (1998). *Supporting scholarly inquiry: incorporating users in the design of the digital library.* „Journal of Academic Librarianship" vol. 24 nr 2, s. 121-130.
143. Pietruch-Reizes, Diana (2005). *Źródła informacji o ochronie dziedzictwa cyfrowego.* „Praktyka i Teoria Informacji Naukowej i Technicznej" nr 1, s. 8-12.
144. Pindlowa, Wanda (1998). *Biblioteka elektroniczna i wirtualna – co to znaczy dla bibliotekarzy i użytkowników.* [W:] Ganińska, Halina red. (1998). Świat biblioteki elektronicznej w klasycznej bibliotece naukowej : *możliwości rozwoju, uwarunkowania i ograniczenia : materiały konferencyjne, Poznań, 19-20 marca 1998.* Poznań: Biblioteka Główna Politechniki Poznańskiej, s. 17-24.
145. Płoszajski, Grzegorz red. (2008). *Standardy w procesie digitalizacji obiektów dziedzictwa kulturowego* [aut. Grzegorz Płoszajski i Kazimierz Schmidt i in.]. Warszawa: Biblioteka Główna Politechniki Warszawskiej; Oficyna Wydawnicza Politechniki Warszawskiej, 247 s. [Wersja elektroniczna: http://bcpw.bg.pw.edu.pl/Content/1262/BG_Stand_w_proc_digit.pdf].
146. *Porozumienie w sprawie utworzenia konsorcjum Polskie Biblioteki Cyfrowe.* (2008) [W:] Ganińska, Halina red. (2008). *Informacja dla nauki a świat zasobów cyfrowych.* Poznań: Biblioteka Główna Politechniki Poznańskiej, s. 185-189.
147. Potęga, Joanna (2006). *Stan rozwoju polskich bibliotek cyfrowych.* „Biuletyn Informacyjny Biblioteki Narodowej" [nr] 1, s. 27-28.
148. Potęga, Joanna; Ślaska, Katarzyna (2006). *Cyfrowa Biblioteka Narodowa i Biblioteka Europejska – droga do Europejskiej Biblioteki Cyfrowej.* „Biuletyn Informacyjny Biblioteki Narodowej" [nr] 4, s. 3-8.
149. Radwański, Aleksander (1998). *Drogi rozwoju biblioteki cyfrowej.* [W:] Ganińska, Halina red. (1998). Świat biblioteki elektronicznej w klasycznej bibliotece naukowej : *możliwości rozwoju, uwarunkowania i ograniczenia : materiały konferencyjne, Poznań, 19-20 marca 1998.* Poznań: Biblioteka Główna Politechniki Poznańskiej, s. 31-40.
150. Radwański, Aleksander (1999). *Biblioteka wirtualna – problemy definicyjne* [dok. elektr.] „EBIB – Elektroniczny Biuletyn Bibliotekarzy" nr 8, http://www.oss.wroc.pl/biuletyn/ebib08/radwan.html [odczyt: 2.10.2011].
151. Radwański, Aleksander (2003). *Biblioteki w ePolsce – trzecia odsłona.* „Bibliotekarz" nr 10, s. 5-11.
152. Rajczak, Nancy (2002). *Rola bibliotek cyfrowych w publikowaniu elektronicznym.* [W:] Kocójowa, Maria red. nauk. (2002). *Elektroniczne publikacje w bibliotekach.* Kraków: Wydaw. Uniw. Jagiel., s. 21-35.
153. Reitz, Joan M. (2004). *Dictionary for library and information science.* Westport; London: Libraries Unlimited, 788, [1] s.

154. Rowlands, Ian; Bawden, David (1999). *Digital libraries: a conceptual framework.* „Libri: International Journal of Libraries & Information Services" vol. 49 nr 4, s. 192-202.
155. Rożniakowska, Małgorzata; Margas, Marcin (2006). *„eBiPol" – Biblioteka Cyfrowa Politechniki Łódzkiej na tle innych inicjatyw bibliotek cyfrowych w kraju od strony technicznej, formalnej i projektowej* [dok. elektr.] „EBIB – Elektroniczny Biuletyn Bibliotekarzy" nr 4, http://www.ebib.info/2006/74/rozniakowska_margas.php [odczyt: 2.10.2011].
156. Rożniakowska, Małgorzata; Margas, Marcin; Kitlińska, Iwona; Bógdoł, Piotr (2006). *Biblioteki dawniej i dziś. Hybrydowe, cyfrowe...? Jakie będą i co może wpłynąć na ich kształt w przyszłości.* [W:] *II Konferencja Biblioteki Politechniki Łódzkiej, Łódź (Poland),19-21 June 2006.* Łódź: Politechnika Łódzka, s. 267-294 [Wersja elektroniczna: http://eprints.rclis.org/handle/10760/9187#.TtTFtvKoEco].
157. Sandusky, Robert J. (2002). *Digital Library Attributes: Framing Usability Research* [dok. elektr.] http://www.uclic.ucl.ac.uk/annb/docs/Sandusky35.pdf [odczyt: 2.06.2011].
158. Saracevic, Tefko (2000). *Digital Library Evaluation: Toward an Evolution of Concepts - 1 - evaluation criteria for design and management of digital libraries.* Library Trends, [dok. elektr.] http://findarticles.com/p/articles/mi_m1387/is_2_49/ai_72274400/pg_18/?tag=content;col1 [odczyt: 2.06.2011].
159. Saracevic, Tefko; Covi, Lisa (2000). *Challenges for digital library evaluation.* [W:] Kraft D.H. red. (2000). *Knowledge Innovations: Celebrating Our Heritage, Designing Our Future. Proceedings of the 63rd Annual Meeting, November 11-16, 2000, Chicago, IL.* Washington: American Society for Information Science, s. 341-350 [Wersja elektroniczna: http://www.is.informatik.uni-duisburg.de/courses/dl_ss04/folien/saracevic00.pdf].
160. Saunders, Laverna M. (1992a). *The virtual library revisited.* „Computers in Libraries" 12 nr 10, s. 51-54.
161. Saunders, Laverna M. (1992b). *The virtual library today.* „Library Administration and Management" 6 nr 2, s. 66-70.
162. Seadle, Michael; Greifeneder, Elke (2007). *Defining a digital library.* „Library Hi Tech" vol. 25 nr 2, s. 169-173.
163. Shen, Rao; Srinivas Vemuri, Naga; Fan, Weiguo; Fox, Edward A. (2006). *What is a Successful Digital Library?* [W:] *Proc. ECDL 2006, Alicante, Spain, Sept. 17-21, 2006, Research and Advanced Technology for Digital Libraries, Lecture Notes in Computer Science,* Vol. 4172. Berlin; Heidelberg: Springer, s. 208-219.
164. Skubała, Elżbieta; Kazan, Anna (2008). *Polskie biblioteki cyfrowe na platformie dLibra – zasób w kontekście tworzenia nowoczesnych kolekcji źródeł informacji dla nauk technicznych.* [W:] Ganińska, Halina red. (2008). *Informacja dla nauki a świat zasobów cyfrowych.* Poznań: Biblioteka Główna Politechniki Poznańskiej, s. 21-33.
165. Ślaska, Katarzyna; Potęga, Joanna (2006). *Cyfrowa Biblioteka Narodowa Polona.* [dok. elektr.] „EBIB – Elektroniczny Biuletyn Bibliotekarzy" nr 11, http://www.ebib.info/2006/81/a.php?slaska_potega [odczyt: 2.10.2011].
166. Sorkin, Virginia; Farley, Judith (1998). *National Digital Library.* [W:] Kent, Alien; Lansour, Harold red. (1997). *Encyclopedia of library and information science,* t. 62. New York: M. Dekker, s. 215-228.
167. Stanisławska-Kloc, Sybilla (2005). *Prawo autorskie a biblioteka cyfrowa - opinia prawna wirtualnych* [dok. elektr.] „EBIB – Elektroniczny Biuletyn Bibliotekarzy" nr 9, http://www.ebib.info/publikacje/matkonf/iwb3/artykul.php?e [odczyt: 2.10.2011].
168. Stefańczyk, Elżbieta red. (2006). *Dygitalizacja zbiorów bibliotecznych: materiały z ogólnopolskiej konferencji pt. Digitalizacja zbiorów bibliotecznych. Warszawa 3-4 czerwca 2005 r.* Warszawa: Wydawnictwo SBP, 120, [3] s.

169. Strauch, Dietmar; Rehm, Margarete (2007). *Lexikon Buch - Bibliothek - neue Medien.* 2 aktualisierte und erw. Ausg. München: K. G. Saur, [4], 472 s.

170. Szczepańska, Barbara (2006). *Europejska Biblioteka Cyfrowa.* „Poradnik Bibliotekarza" [nr] 9, s. 16-17.

171. Szczepańska, Barbara (2007). *Zasady publikowania obiektów cyfrowych w repozytoriach europejskich : zagadnienia prawnoautorskie.* „Przegląd Biblioteczny" R. 75 z. 2, s. [193]-220.

172. Szczygłowska, Lidia (2006). *„Cybrary" to wciąż biblioteka.* [W:] *II Konferencja Biblioteki Politechniki Łódzkiej, Łódź (Poland),19-21 June 2006.* Łódź: Politechnika Łódzka, s. 51-70.

173. Szczygłowska, Lidia (2006). *Biblioteka wirtualna jako kolejny etap komputeryzacji biblioteki.* [W:] *Internet w bibliotekach II : łączność, współpraca, digitalizacja : Wrocław, 23-26 września 2003 roku.* Warszawa Stowarzyszenie Bibliotekarzy Polskich, K[omisja] W[ydawnictw] E[lektronicznych], Redakcja „Elektronicznej Biblioteki", 2003. [Wersja elektroniczna: http://ebib.oss.wroc.pl/matkonf/iwb2/szczyglowska.php].

174. Szymula, Robert (1999). *Biblioteki cyfrowe – zagadnienia ogólne i problemy terminologiczne* (Konferencja COLIS3, Dubrownik 23-26.05.1999.) [dok. elektr.] „EBIB – Elektroniczny Biuletyn Bibliotekarzy" nr 8 http://www.oss.wroc.pl/biuletyn/ebib08/szymula.html [odczyt: 2.10.2011].

175. Trembowiecki, Aleksander (2006b). *Digitalizacja w bibliotekach: problemy, szanse, zagrożenia.* „Poradnik Bibliotekarza" nr 5, s. 3-7.

176. Trembowiecki, Aleksander (2006a). *Digitalizacja zbiorów bibliotecznych : teoria i praktyka.* Warszawa: Wydawnictwo Centrum Edukacji Bibliotekarskiej, Informacyjnej i Dokumentacyjnej, 137 s.

177. Trembowiecki, Aleksander (2006c). *Digitalizacja zbiorów bibliotecznych – ujęcie teoretyczne.* „Poradnik Bibliotekarza" nr 2, s. 3-9.

178. Wałek, Anna (2008). *Dolnośląska Biblioteka Cyfrowa – kolekcje, technika i organizacja pracy.* „Zagadnienia Informacji Naukowej" nr 2, s. 16-28.

179. Wałek, Anna (2009). *Biblioteki cyfrowe na platformie dLibra.* Warszawa: Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich, 143, [1] s.

180. Wałek, Anna (2009). *Wykorzystanie zbiorów cyfrowych bibliotek na przykładzie Dolnośląskiej Biblioteki Cyfrowej.* [W:] Migoń, Krzysztof; Skalska-Zlat, Marta red. (2009). *Uniwersum piśmiennictwa wobec komunikacji elektronicznej.* Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego, s. 361-369.

181. Waters, Donald J. (1998). *What are digital liraries?* [dok. elektr.] http://www.clir.org/pubs/issues/issues04.html#dlf, CLIR – Council on Library and Information Resources, issues Number 4 July/August 1998 [odczyt: 2.10.2011].

182. *Wikipedia* (2001-) [dok. elektr.] http://pl.wikipedia.org/; http://en.wikipedia.org/ [odczyt: 2.01.2011].

183. Witten, Ian H., Nevill-Manning, Craig G., McNab, Rodger; Cunningham, Sally Jo (1998). *A public library based on full text retrieval.* „Communications of the ACM" vol. 41 nr 4, s. 71-75.

184. Witten, Ian H.; Nevill-Manning, Craig G.; Cunningham, Sally Jo (1995). *Building a public digital library based on full-text retrieval.* (Working paper 95/24). Hamilton, New Zealand: University of Waikato, Department of Computer Science [dok. elektr.] http://researchcommons.waikato.ac.nz/bitstream/10289/1101/1/uow-cs-wp-1995-24.pdf [odczyt: 2.10.2011].

185. Wójtowicz, Renata (2009). *Biblioteki cyfrowe – prezentacja przykładowa, ale nie wzorcowa.* „Biblioteka – Szkolne Centrum Informacji" nr 3, s. 11-13.
186. Wolfram, Dietmar; Xie, Hong (Iris) (2002). *State digital library usability: contributing organizational factors.* „Journal of the American Society for Information Science and Technology" vol. 53 nr 13, s. 1085-1097 [Wersja elektroniczna: http://www.sois.uwm.edu/xie/IrisArticles/Articles/Statedigitallibrary02.pdf].
187. Wolfram, Dietmar; Xie, Hong (Iris) (2002). *Traditional IR for web users: a context for general audience digital libraries.* „Information Processing and Management" 38, s. 627-648.
188. Woźniak-Kasperek, Jadwiga (2008). *eContent czyli o organizacji informacji i wyszukiwaniu w bibliotece cyfrowej.* „Zagadnienia Informacji Naukowej" nr 2, s. 50-58.
189. Woźniak-Kasperek, Jadwiga (2009). *Funkcjonalność i dostępność bibliotek cyfrowych.* [W:] Migoń, Krzysztof; Skalska-Zlat, Marta red. (2009). *Uniwersum piśmiennictwa wobec komunikacji elektronicznej.* Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego, s. 393-406.
190. Woźniak-Kasperek, Jadwiga; Franke, Jerzy red. (2007). *Biblioteki cyfrowe: projekty, realizacje, technologie: praca zbiorowa.* Warszawa: Stowarzyszenie Bibliotekarzy Polskich, 212 s. [wersja elektr. http://bbc.uw.edu.pl/dlibra/docmetadata?id=4&from=publication&].
191. Żabicki, Sławomir (2009). *Bałtycka Biblioteka Cyfrowa – sprzęt i oprogramowanie, jego techniczne plusy i minusy.* „Bibliotekarz" [nr] 7/8, s. 30-32.
192. Zbroszczyk, Iwona (2008). *Europeana – Europejska Biblioteka Cyfrowa.* „Bibliotekarz". [nr] 7/8, s. 11-13.
193. Zbroszczyk, Iwona (2009). *Ku bibliotece wirtualnej.* „Bibliotekarz" [nr] 3, s. 7-10.
194. Zhang, Yan (2010). *Dimensions and elements of people's mental models of an information-rich web space.* „Journal of the American Society for Information Science and Technology" vol. 61 nr 11, s. 2206-2218.
195. Zhang, Ying (2007). *Developing A Holistic Model For Digital Library Evaluation.* A Dissertation Submitted to the Graduate School-New Brunswick Rutgers, The State University of New Jersey in partial fulfillment of the requirements for the degree of Doctor of Philosophy Graduate Program in Communication, Information, and Library Studies written under the direction of Dr. Tefko Saracevic. New Brunswick, New Jersey, 248 s.
196. Zhang, Ying (2010). *Developing a Holistic Model for Digital Library Evaluation* [dok. elektr.] http://onlinelibrary.wiley.com/doi/10.1002/asi.21220/pdf [odczyt: 2.10.2011].
197. Ziółkowska, Sylwia (2008). *Projekt i serwis Biblioteki Europejskiej (The European Library – TEL).* „Zagadnienia Informacji Naukowej" nr 2, s. 3-15.

Biblioteki cyfrowe na świecie – powstanie i rozwój

Władysław Marek Kolasa

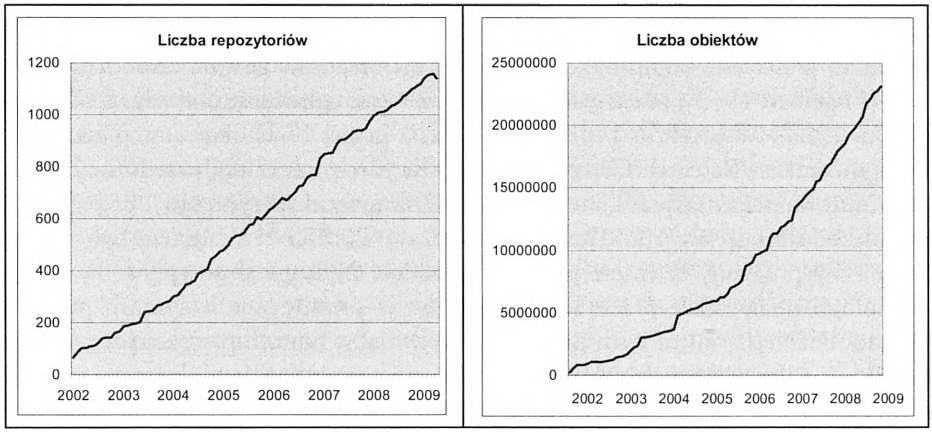
W opracowaniu przedstawiono zwięzły zarys dziejów bibliotek cyfrowych na świecie oraz wskazano główne determinanty ich rozwoju. W części pierwszej, która obejmuje okres do 1993 r. przedstawiono rozwój idei biblioteki cyfrowej i pierwsze eksperymenty z tego zakresu; w drugiej (dotyczącej lat 1994-2001) omówiono refleksję teoretyczną oraz rozwój bibliotek cyfrowych tzw. pierwszej generacji; w trzeciej natomiast skoncentrowano się na rozwoju systemów po roku 2001, czyli na bibliotekach tzw. drugiej generacji oraz systemach rozproszonych. Całość zamykają krótkie rozważania dotyczące perspektyw rozwoju bibliotek cyfrowych, gdzie wskazano aktualne kierunki ich ewolucji.

Wstęp

Biblioteki cyfrowe, jakkolwiek są zjawiskiem stosunkowo młodym, mają za sobą długą i rozbudowaną historię. Wprawdzie ich rozwój prowadzący do systemów informacyjnych istniejących w obecnym kształcie liczy sobie dopiero niespełna dekadę, lecz wnikliwa analiza przekonuje, że poprzedziło ją szereg doświadczeń i eksperymentów, bez których współczesna forma nie mogłaby zaistnieć. W sensie porządkującym można wskazać kilka głównych etapów ich rozwoju, w których skrystalizowały się kluczowe wizje i wdrożono udane projekty. Pierwszy okres, który rozpoczął się od literackiej wizji Herberta George'a Wellsa z 1937 r. trwał do początku lat 90. XX wieku; w czasie tym powstawały rozmaite koncepcje, budowano prototypy urządzeń i pojawiły się kolekcje eksperymentalne. Drugi okres trwał całą dziesiątą dekadę XX wieku i przesunął się o kilka lat na wiek XXI. W czasie tym powstały już pierwsze biblioteki cyfrowe, które szybko zaistniały w rodzącej się sieci Internet oraz podjęto badania podstawowe nad ich modelem i funkcją; w efekcie podjętych prac wypraco-

wano pierwsze standardy i skrystalizowały się grupy wsparcia. Trzeci okres związany immanentnie z ideą Web 2.0 rozpoczął się u progu pierwszej dekady XXI wieku i trwa do dziś. Trudno wprawdzie wskazać konkretną datę, ale analiza danych największego światowego serwisu OAISTER (rys. 1) przekonuje, że proces ów rozpoczął się ok. 2001 r. i uległ przyspieszeniu w 2005 r.

Ostatni okres można określić jako przejście do fazy dojrzałej; w sensie praktycznym charakteryzuje się on silnym trendem w kierunku budowy serwisów rozproszonych i orientacją w kierunku mediów społecznościowych.



Rys. 1. Biblioteki cyfrowe w serwisie OAISTER w latach 2002-2009. Przeliczenia własne na podstawie OAIster Statistics – http://www.lib.umich.edu/digital-library-production-service-dlps/oaister-statistics [odczyt: 19.05.2011].

Pierwsze próby i doświadczenia

Ideę biblioteki cyfrowej znaleźć można już w opublikowanych przed wojną esejach Herberta George Wellsa. W zbiorze *World brain* [Wells 1938] poświęconym rozważaniom na temat przyszłości organizacji wiedzy i edukacji zamieścił autor szkic *The Idea of a Permanent World Encyclopaedia*, w którym opisał ideę globalnej encyklopedii opartej na mikrofilmach. Miała ona wedle wizji autora obejmować uniwersum ludzkiej wiedzy i stać się motorem jednoczenia rasy ludzkiej. O krok dalej poszedł znany amerykański wynalazca Vannevar Bush, uważany za ojca idei sieci komputerowych i hipertekstu. Bush był zaniepokojony nadmiarem informacji, na jaki napotykają naukowcy i zaproponował kilka rozwiązań technologicznych, które – jago zdaniem – mogłyby w przyszłości rozwiązać ten problem. W sławnym artykule *As We May Think* opublikowanym na łamach *Atlantic Monthly* w 1945 roku przedstawił idee urządzenia Memex [Memory Extender], które składało się biurka (pulpitu) ze skośnymi ekranami, które umożliwiało wygodną lekturę. Mechanizm był swego

rodzaju czytnikiem mikrofilmów, z których obrazy były inteligentnie wyszukiwane, dekodowane i transportowane na ekran z dużą prędkością, nade wszystko jednak były zdolne do automatycznego przechodzenia od zawartych w nich odniesień do innych prac, czyli hipertekstowe [Tedd; Large 2005]. Kolejne ważne pomysły pojawiły się już w czasie, gdy istniały komputery; rozwijał je twórczo m.in. amerykański cybernetyk Joseph Licklider w systemem przetwarzania wiedzy *Procognitive* i uwiecznił m.in. w sławnej książce *Libraries of the future* (1965) [Hauben 2004]. Zapewne pod wpływem idei Busha był także Douglas Engelbart, twórca licznych wynalazków (np. myszy komputerowej), uważany za pioniera w dziedzinie interfejsów. Naukowiec ten prowadził intensywne prace w zakresie hipertekstu i szukał rozwiązań w zakresie komunikacji naukowej dla zespołów geograficznie rozproszonych [Engelbart 1963]; prace te uwieńczyło pierwsze wdrożenie pod nazwą oN-Line System (NLS). W grudniu 1968 r. Engelbart z grupą 17 naukowców pracujących w Augmentation Research Center (Stanford Research Institute), przedstawił udaną 90-minutową prezentację NLS na żywo [zapis filmowy eksperymentu – http://sloan.stanford.edu/mousesite/1968Demo.html; odczyt: 20.05.2011]; system był – warto dodać – wyposażony w interfejsy audio i wideo. Niebagatelny wpływ na rozwój technologii miały także prace Theodora Nelsona prowadzone w ramach projektu Xanadu. Późniejszy autor terminu *hipertext* rozwijał w ramach projektu ideę *idealnej przestrzeni pamięci literackiej*, która była w swej istocie zalążkiem obecnego systemu www. Koncepcje swe Nelson uwiecznił w dwu książkach: *Literary Machines* [Maszyny literackie] (1981) oraz *Computer Lib/Dream Machines* [Maszyny snów] (1974). Ukoronowaniem wskazanych idei były prace Tima Bernersa-Lee podjęte w latach 1989-1991, w których stworzył on podstawy teoretyczne i praktyczne współczesnych systemów hipertekstowych, czyli www. Idee wyżej wspomnianych uczonych stworzyły grunt dla pierwszych doświadczeń w zakresie organizacji i manipulacji dużych ilości danych, które z dzisiejszej perspektywy można uznać za prekursorów bibliotek cyfrowych [Tedd; Large 2005].

Odrębny tor rozważań obejmował rozwój technologii (głownie sprzętu) oraz realizowane w tym czasie pierwsze projekty biblioteczne. Warto zauważyć, że aż do końca lat 70. dostępna infrastruktura informatyczna była dopiero w fazie szkieletowej, co nie tworzyło sprzyjających warunków do jej wykorzystania. Przypomnijmy też, że aż do początku lat 80. dominowały komputery klasy mainframe, które były używane jedynie w dużych ośrodkach badawczych; niewielka była także liczba komputerów wpiętych do sieci (23 hosty w 15 węzłach w 1971 r., która wzrosła do 1000 dopiero w 1984 r.), zaś przepustowość sieci niska. Ogromnym problemem było też przechowywanie danych; pierwsze dyski produkcyjne twarde (Segate) pojawiły się w 1980 r., a płyty CD dopiero na przełomie 1982 i 1983 r. W efekcie potencjalna liczba użytkowników był niewielka. Mimo owej niesprzyjającej aury już w 1971 r. pojawił się pierwszy projekt biblioteki cyfrowej – student Uniwersytetu w Illinois Michael Hart rozpoczął tworzenie zbioru tekstów, które z czasem przerodziły się w *Gutenberg Project* [http://www.gutenberg.org/]. Z uwagi na brak uregulowań prawnych w tym zakresie Hart postanowił umieszczać jedynie utwory

z domeny publicznej, które aż do początku lat 90. XX w. rozpowszechniał pocztą elektroniczną w formacie tekstowym. Ręczne przepisywanie było pracochłonne, stąd mimo udziału woluntariuszy zbiór powiększał się wolno: w 1989 r. kolekcja liczyła 10 pozycji i wzrosła do 100 w 1994 r. Okres intensywnego rozwoju rozpoczął się dopiero po 1995 r., gdy na szerszą skalę wykorzystano OCR; w kolejnych latach kolekcja rosła dynamiczniej i już w 1997 r. przekroczono 1000 pozycji (1999 – 2000, 2003 – 10 000), zaś w 2008 r. – 25000 [Lebert 2008]. Pod koniec lat 90. XX w. zwiększono też liczbę dostępnych wersji (pojawiły się m.in. formaty multimedialne i mobilne), rozbudowano system metadanych oraz interfejs (wyszukiwarka od 1994). Już w latach 90. ubiegłego wieku projekt znalazł pierwszych naśladowców; bliźniacze inicjatywy powołano m.in. w Niemczech (1994) [http://projekt.gutenberg.de/], Australii (2001) [http://gutenberg.net.au/] i in. W 2004 pojawiła się również poszerzona wersja projektu *Project Gutenberg Consortia Center* [http://www.gutenberg.us/] (75 tys. poz.), która podjęła ścisłą współpracę z *WorldLibaryt.net*. Z dzisiejszej perspektywy *Gutenberg Project* reprezentuje nico przestarzałe podejście do zasobu i prymitywną technologię, niemniej sama idea, którą wdrożył Hart przez wiele lat inspirowała następców, zapisując się tym samym trwale w historii. Prócz opisywanego projektu do końca lat 80. pojawiło się jeszcze kilka innych inicjatyw, np. *Perseus Digital Library* [http://www.perseus.tufts.edu] założona w 1987 r. na Tufts University; zbiór ten początkowo deponowano na płytach CD, a od 1994 r. udostępniono w sieci. Z innych ważniejszych inicjatyw wczesnego etapu warto jeszcze wymienić: serwis skandynawski *Project Runeberg* [http://runeberg.org/] (od 1992) oraz *Bartleby* [http://www.bartleby.com/], *The On-Line Books Page* [http://digital.library.upenn.edu/books/], pierwsze pełne wydanie on-line dzieł Szekspira [http://shakespeare.mit.edu/], *MUSE* [http://muse.jhu.edu/] (od 1993), czy *The Humanities Text Initiative* [http://www.hti.umich.edu/] (1994) [Sauber 2009].

Z zupełnie z innych powodów ogromną rolę odegrał także serwis preprintowy *arXiv* [http://arxiv.org/] uruchomiony w 1991 r. System ten bowiem zapoczątkował pewien model naukowych zachowań publikacyjnych określanych potem jako *kultura preprintu*, które następnie znalazły szeroki odzew w świecie nauki (szczególnie w naukach ścisłych). System *arXiv* – należy podkreślić – otworzył także drogę do idei Open Access, co znalazło finał m.in. w późniejszej Deklaracji Berlińskiej z 2003 r. i szeregu innych uregulowaniach z tego zakresu, intensywnie potem promowanych przez wiele urzędów i agencji np. Radę Unii Europejskiej. *arXiv* stał się zatem prototypem instytucjonalnych repozytoriów, czyli systemów, które realizują funkcje *self-publishing* (składanie, przeglądanie, edytowanie, itd.) oraz rozpowszechniania dokumentów w formacie cyfrowym. Pierwsze tego typu systemy były nieskomplikowane i miały prostą architekturę; prócz *arXiv* były to m.in. baza dysertacji *Electronic Thesis and Dissertations Repositories* (od 1996) i ekonomiczny *RePEc* [http://repec.org/] (od 1997). Z kolei pierwszym systemem, który oferował usługi w zakresie *self-publishing* z wykorzystaniem PDF (od 1996) był istniejący od 1991 roku *Networked Digital Library of Theses and Dissertations* (NDLTD) [http://www.ndltd.org/]. Systemy te po

licznych ewolucjach przetrwały do współczesności i z obecnej perspektywy mogą być uważne za prototypy bibliotek cyfrowych [Candela; Castelli; Pagano 2011].

Obraz wczesnej fazy rozwoju bibliotek cyfrowych byłby niepełny, gdyby pominąć pierwsze serwisy czasopism elektronicznych i inicjatywy z tym związane. Istotną rolę odegrały tu w szczególności projekty *Mercury, CORE, TULIP i Red Sage*. Projekt *Mercury* (1989-1992) prowadzony w Mellon University w Pittsburgu był jednym z pierwszych prób udostępnienia na terenie kampusu artykułów z czasopism naukowych z zakresu informatyki [Mercury 1992]. Kolejnym dużym przedsięwzięciem stał się prowadzony w latach 1991-1995 projekt *Chemistry Online Retrieval Experiment* (CORE), w który zaangażowany był m.in. Cornell University, Chemical Abstracts Service i OCLC; owocem prac była digitalizacja i udostępnienie 20 czołowych czasopism chemicznych. Teksty te przygotowano w dwu wersjach: jako obrazy rastrowe oraz jako pliki tekstowe ze znacznikami SGML (razem 425 tys. stron tekstu). Wielkie znaczenie miał także koordynowany przez Elsevier Science projekt *TULIP* (1991-1995) [informacje o projekcie: http://www.elsevier.com/wps/find/librariansinfo.librarians/tulip; odczyt: 21.05.2011], w trakcie realizacji którego testowano rozmaite modele dystrybucji i rozwiązania techniczne. Na użytek testów i badań zeskanowano i poddano OCR ponad 500 tys. stron tekstu. Projekt ów prowadzony we współpracy w 15 czołowymi amerykańskimi uniwersytetami zakończył się spektakularnym sukcesem i z czasem stał się zalążkiem obecnego serwisu *ScieneceDirect* [Nahotko 2007]. Wymienione projekty, jakkolwiek nie były bibliotekami cyfrowymi *sensu stricto* pomogły wskazać korzyści, jakie można osiągnąć poprzez cyfrową dystrybucję wydawnictw oraz przyczyniły się do postępu technologicznego w zakresie techniki skanowania i OCR [Tedd; Large 2005].

W pierwszej połowie lat 90. podjęto także pierwsze próby refleksji teoretycznej nad bibliotekami cyfrowymi. Duże znaczenie w tym zakresie miały zorganizowane w 1991 r. warsztaty *Future Directions in Text Analysis, Retrieval and Understanding*. W latach 1992-1993 odbyło się kilka kolejnych konferencji i spotkań, na których kontynuowano i rozwijano zainicjowane w 1991 r. wątki. Pokłosiem wspomnianych imprez prac wydana sumptem National Science Foundation obszerna elektroniczna książka pt. *Source Book on Digital Libraries* [Fox 1993]. Pracę tę można postrzegać jako swoistą syntezę stanu wiedzy na temat bibliotek cyfrowych u schyłku 1993 r.; zawiera ona zarówno rozważania teoretyczne, jak i szczegółowe opisy projektów i kalendaria oraz bibliografię przedmiotu (28 pozycji). Najważniejszym jednak efektem wspomnianych wyżej spotkań było przygotowanie gruntu pod duże projekty; ukoronowaniem tych starań był rozpoczęty w 1994 r. *Digital Library Initiative*, który zapoczątkował nowy okres rozwoju bibliotek cyfrowych.

Prace badawczo-rozwojowe i ewolucja bibliotek cyfrowych w latach 1994-2001

Od początku lat 90. XX wieku, czyli wraz z narodzinami www rozpoczyna się okres dynamicznego rozwoju bibliotek cyfrowych i ich stopniowa ewolucja. Pierwsze

z nich w sensie technicznym były nieskomplikowane; zwykle tworzono je jako proste witryny HTML zawierające wykazy hiperłączy do publikacji (tzw. model *strona www*), które z czasem, w miarę powiększania się zbiorów, wyposażono w mechanizmy wyszukiwawcze i system metadanych (tzw. model *katalog obiektów cyfrowych*). Wadami tych rozwiązań był jednak brak mechanizmów konserwacyjnych oraz brak kontroli dostępu, zarządzania zasobem i metadanymi. Obydwa modele odzwierciedlały więc filozofię pierwszych bibliotek cyfrowych (np. *Gutenberg Project*); charakterystycznym ich rysem było programowe ukierunkowanie na udostępnienie kanonu kulturowego (klasyka literacka, dziedzictwo historyczne) oraz założenie, że dostęp do zbiorów winien być nielimitowany. W efekcie biblioteki te realizowały przede wszystkim cel archiwalny (nie biblioteczny); z drugiej zaś strony udostępniane zasoby nie wymagały stosowania żadnych mechanizmów zabezpieczeń i kontroli dostępu, gdyż były z natury swej wolne od ograniczeń wynikających z ochrony praw autorskich. Warto dodać, że nie interesowano się też szerzej problematyką użytkowania (interfejs) ani mechanizmami wyszukiwawczymi. Takie usytuowanie celów nie sprzyjało rozwojowi bibliotek cyfrowych. Było bowiem niewystarczające do realizacji funkcji bibliotecznych, gdyż stało w sprzeczności z jedną z zasadniczych cech biblioteki, zakładającej, że proces gromadzenia jest permanentny (stały, ciągły, powtarzalny), a także otwarty na nowe formy dokumentów. Z drugiej zaś strony istniała uzasadniona niechęć wydawców do nielimitowanego udostępniania, gdyż nie zabezpieczało to ich interesu ekonomicznego [Kolasa 2009].

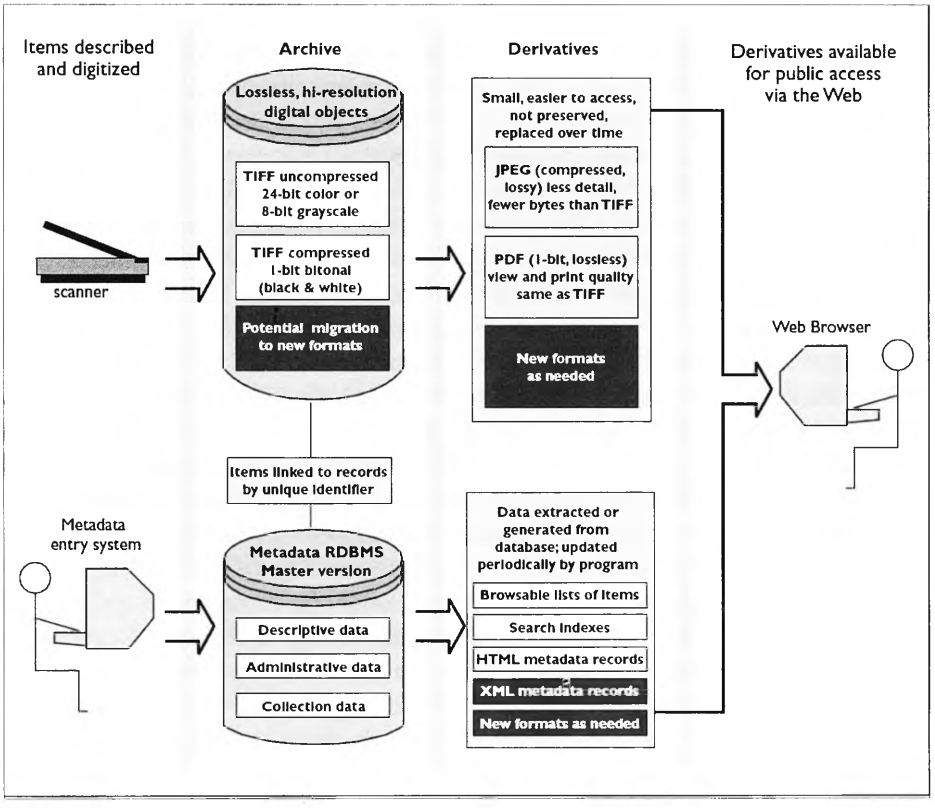
Czynniki te stały się motorem podjęcia licznych prac zmierzających do zmiany istniejącej sytuacji i rozpoczęła się stopniowa ewolucja systemów do form bardziej zaawansowanych. Prace te nabrały tempa w drugiej połowie ostatniej dekady XX wieku, gdy podjęto wiele ważnych inicjatyw mających charakter badań podstawowych. W ich wyniku powstało szereg standardów i rozwiązań, które – jak czas pokazał – okazały się trwałe. Rozmach w tym obszarze najlepiej ilustruje krótki wykaz wybranych dokonań tego okresu: uruchomiono duże projekty badawczo-rozwojowe ukierunkowane na biblioteki cyfrowe (*Digital Library Initiative* [1994-1998 i 1999-2004]; *eLib* [1998-2001]); narodowe i międzynarodowe strategie digitalizacji (np. amerykański *National Digital Library Program* [http://international.loc.gov/ammem/dli2/html/lcndlp.html; odczyt: 21.05.2011] [1995], czy *Memory of the World Programme* [1992]); powstały pierwsze stowarzyszenia bibliotek cyfrowych (*Digital Library Federation* [1995]) i organizacje koordynujące (np. *Joint Information Systems Committee* – JISC [1993], *Electronic Information for Libraries Direct* - eIFL [1999]); tematyczne konferencje (*Joint Conference on Digital Libraries* [1994]; *European Conference on Digital Libraries* (ECDL) [1997]; *International Conference on Asian Digital Libraries* [1998]); czasopisma fachowe (*D-Lib Magazine* [1995]; *Ariadne* [1996]) oraz standardy i specyfikacje (Dublin Core [1995]; DIENST [1995]; RDF [1999]; OAI-PMH [2001]). Atmosferę tego okresu wiernie obrazują materiały zbierane przez profesora Edwarda A. Foxa cyklicznie udostępniane na stronie Virginia Tech w latach 1998-2001 [http://ei.cs.vt.edu/~dlib/]. Dokumenty te (zebrane w formie *silva rerum*) zawierają wiele informacji historycznych, w tym zrzuty ekranowe nieistniejących już

serwisów. Interesujący jest wachlarz rozważanych wówczas problemów. Ich analiza przekonuje, że już u progu XXI wieku wiedza na temat bibliotek cyfrowych osiągnęła wysoki poziom dojrzałości. Udowadnia to krótki wykaz tematów z opracowań E. A. Foxa (2001):

> 1. Podstawy (definicje, publikacje, projekty); 2. Wyszukiwanie, przeglądanie (metody, indeksowanie, wyszukiwanie rozproszone, roboty/pająki, integracja linków); 3. Multimedia (tekst, audio, obraz, film, grafika, animacja, JPEG, MPEG, wyszukiwanie oparte na treści dokumentu); 4. Architektura (komponenty, protokoły, InfoBus, protokoły pośrednie, ramki, OAI); 5. Interfejs (psychologia, środowisko, taksonomia elementów, wizualizacja, design, użyteczność); 6. Metadane (ontologie, RDF, MARC, DublinCore, IMS, mapowanie, konwersja); 7. Publikacje elektroniczne (autorstwo, prezentacja, obiektowy model DOM, równoległe publikowanie, style XLS, e-książki, struktura, adiustacja, tagowanie); 8. Problemy baz danych (technologia, struktura informacji, dane multimedialne i hipermedialne, replikacja, magazynowanie); 9. Oprogramowanie agentowe do zarządzania informacją (problemy, protokoły, automatyczne indeksowanie); 10. Handel, ekonomia, wydawcy (zabezpieczenia, ochrona, problemy Open Access, modele płatności); 11. Prawa własności intelektualnej (ochrona praw autorskich, copyright, zarządzanie prawami); 12. Problemy społeczne (współpraca, ocena, użytek w edukacji, dziedzictwo kulturowe, organizacje).

Bliższa analiza wyliczonych tematów pozwala stwierdzić, że już w II połowie lat 90. XX w. podjęto niemal wszystkie wątki rozważane współcześnie.

Przez całe lata 90. równolegle ze wzrostem liczby repozytoriów (było ich wówczas ok. 300-500) następowała ewolucja systemów i doskonalenie standardów związanych z wspieraniem zapisu cyfrowego, następował rozwój formatów i nośników danych. Z czasem podjęto także próby udostępnienia modeli i narzędzi oferujących zakres usług zbliżonych do tradycyjnej biblioteki (gromadzenie, opracowanie, udostępnianie) oraz zainteresowano się interfejsem. Typowy model biblioteki cyfrowej tego okresu pokazano na rys. 2. Wyróżniano w nim najczęściej trzy poziomy: 1. poziom tworzenia (skanowanie i edycja metadanych), 2. poziom bazy danych: system archiwalny i prezencyjny oraz 3. poziom użytkownika (przeglądarka). Na styku poszczególnym poziomów wskazywano na rozmaite problemy, np.: jakość kopii archiwalnej i użytkowej, typy metadanych, mechanizmy wyszukiwawcze. Warto zwrócić uwagę, że w modelu tym zupełnie marginalne znaczenie miała współpraca z innymi systemami (XML, new formats), w efekcie system działał niejako w pojedynkę. Zmiana tej perspektywy dokonywała się jednak stopniowo od 1999 r., zaś po 2002 r. weszła na trwałe.



Rys. 2. Model biblioteki cyfrowej w 2001 roku [McCray; Gallagher 2001].

Jakkolwiek prace badawczo-wdrożeniowe nad bibliotekami cyfrowymi podjęto w wielu ośrodkach, jedynie trzy inicjatywy pozostawiły trwałe ślady w historii informatyki: *Digital Library Initiative* (DLI1 [1994-1998] i DLI2 [1999-2004]) oraz *eLib*. Bez wątpienia największy wpływ na liczbę wdrożeń i postęp w zakresie badań nad bibliotekami cyfrowymi miały projekty amerykańskie. Rozpoczął je w 1994 r. pierwszy etap *Digital Library Initiative* [DLI1], którego głównym inicjatorem była agencja rządowa National Science Foundation wspierana przez Department of Defense Advanced Research Projects Agency (DARPA) oraz National Aeronautics and Space Administration (NASA). Projekt trwał 4 lata i obejmował finansowanie prac realizowanych na sześciu uczelniach. Głównym celem DLI1 były i rozwój podstawowych technologii na rzecz bibliotek cyfrowych. Rozwijano w jej ramach trzech grup problemowych, które obejmowały: 1. utrwalanie, kategoryzację i porządkowanie informacji; 2. wyszukiwanie, przeglądanie, filtrowanie, prezentację i wizualizację; 3. protokoły i standardy sieciowe. Wykonawcami były czołowe amerykańskie uniwersytety: University of California w Berkeley, University of California w Santa Barbara, Carnegie Mellon University, University of Illinois w Urbana, University of Michigan oraz Stanford University. DLI1 okazała się sporym osiągnięciem, stąd w 1999 r. ogło-

szono jej drugi etap [DLI2], trwający kolejne sześć lat. W fazie drugiej skoncentrowano się na wdrożeniach wcześniej badanych technologii i testowaniem poszczególnych rozwiązań. Czas pokazał, że szczególne znaczenie miały prace z zakresu architektury i problemów użyteczności. Najistotniejsze zaś wydaje się przyjęcie koncepcji, w której architektura systemu jest zorientowana na użytkownika. W trakcie realizacji DLI2 do grona uczestników dołączyli nowi wykonawcy, m.in.: Library of Congress, National Library of Medicine, National Endowment for the Humanities i wielu innych [Fox 1999; Madalli 2003]. Po obu odsłonach projektu pozostały trwałe osiągnięcia. Sześć z nich – jak podkreślają znawcy technologii [Candela; Castelli; Pagano 2011] – było szczególnie istotnych: 1. *Projekt California Environmental Digital Library* [obecnie: California Environmental Digital Library Network (CalEDLN): http://caledln.projects.atlas.ca.gov/; odczyt: 21.05.2011], który wniósł wiele w zakresie technologii obsługi dużych, rozproszonych zbiorów ikonograficznych (fotografii, zdjęć satelitarnych, wideo, map); 2. *Alexandria Digital Library* [http://www.alexandria.ucsb.edu/], gdzie rozwijano problemy bibliotek rozproszonych oraz zagadnienia tworzenia kolekcji zawierających obiekty geograficzne (mapy, zdjęcia lotnicze, obrazy satelitarne) oraz badano zapytania zdefiniowane geograficznie; 3. *Projekt Digital Video Library* [http://www.informedia.cs.cmu.edu/], który zrealizował udaną koncepcję dużej cyfrowej kolekcji wideo obsługującej wyszukiwanie oparte na wiedzy; 4. Projekt *The Interspace,* koncentrujący się na budowie dużych kolekcji z zakresu techniki inżynieryjnej i fizyki, które były jednocześnie skutecznie przeszukiwane z poziomu jednolitego interfejsu; 5. *The University of Michigan Digital Library* [http://www.lib.umich.edu/digital-library-production-service-dlps], specjalizujący się w tworzeniu architektury systemów bibliotecznych, wykorzystujących oprogramowanie agentów; 6. *The Stanford Digital Library Project* [http://diglib.stanford.edu:8091/], badający zastosowanie zasad interoperacyjności w środowisku heterogenicznych usług i zbiorów, z użyciem specjalnego protokołu [InfoBus], który zapewniał dostęp do usług i informacji z różnych źródeł poprzez agentów. Jakkolwiek większość wspomnianych projektów nie przetrwała do współczesności w pierwotnej formie – należy podkreślić – że wniosły one podówczas ogromny wkład do wiedzy informatycznej.

Na marginesie warto podkreślić, że jednym z najważniejszych sukcesów wynikających z historii tych projektów jest Google. Późniejsi twórcy wyszukiwarki Lawrence Page i Sergey Brin rozpoczęli prace nad jej silnikiem w czasie, gdy byli doktorantami na Uniwersytecie Stanforda w trakcie realizacji Stanford Digital Library Project. DLI wykracza jednak daleko poza konkretne prace i wynalazki; ze współczesnej perspektywy można stwierdzić, że projekty te stworzyły podstawy do współczesnego oblicza bibliotekarstwa cyfrowego, jako nowej dyscypliny badawczej, która stała się sumą wysiłków teoretycznych kilku specjalności: informatyki, informacji naukowej, językoznawstwa, psychologii i in. [Candela; Castelli; Pagano 2011].

Równolegle z DLI podjęto w USA szereg innych inicjatyw i prac z interesującego nas zakresu, m.in. na University of Arizona badano wydajne systemy klasyfikacyjne – realizujące postulat *od wyszukiwania informacji do zarządzania wiedzą*; na Carnegie Mellon University rozpoczęto prace nad koncepcją późniejszego serwisu *Universal*

Library [http://www.ulib.org/]; na University of South Carolina testowano oprogramowanie do bibliotek; na Stanford University badano techologie na rzecz bibliotek cyfrowych zaś w Tufts University kontynuowano rozwój *The Perseus Digital Library Project*. Dużą rolę w konsolidacji wysiłków odegrało założone w 1995 r. konsorcjum *Digital Library Federation* (DLF) [http://old.diglib.org/]. Inicjatorami przedsięwzięcia były m.in. duże biblioteki realizujące własne projekty cyfrowe, m.in. New York Public Library [http://www.nypl.org/collections], Library of Congress (*American Memory* [http://memory.loc.gov/ammem/]) oraz instytucje nadzorujące dostęp do dokumentów, m.in. National Archives and Records Administration i Commission on Preservation and Access (CPA); ostatnia z nich zapewniała finansowanie pozyskane od IBM. W krótkim czasie liczba członków DLF urosła do 37, zrzeszając czołowe amerykańskie biblioteki oraz kilka książnic zagranicznych. Cele strategiczne DLF były jasno sformułowane: rozwój standardów, promocja najlepszych praktyk, stymulacja rozwoju poprzez wpływ na instytucje finansujące, ochrona własności intelektualnej, a nade wszystko budowa wspólnej infrastruktury [szczegółowy wykaz projektów: http://old.diglib.org/produce.htm; odczyt: 21.05.2011]. Ostatni z celów dalece wybiegał w przyszłość i zaowocował pracami ukierunkowanymi na rzecz tworzenia bibliotek rozproszonych (współpraca przy tworzeniu OAI-PMH oraz jego impementacji w serwisie *OAISTER* afiliowanym w University of Michigan od 2002 r.) [Greenstein; Thorin 2002].

Nie sposób przecenić także prac prowadzonych w ramach *Open Archive Initiative*, organizacji finansowanej przez Fundację Mellona, powołanej wspólnie przez DLF i partnerów społecznych. Owocem prowadzonych prac był wspomniany wyżej protokół OAI-PMH, którego pierwszą wersję opublikowano w 1999 r. Rozwiązanie to wychodziło naprzeciw idei interoperacyjności, czyli zdolności sieci do efektywnej współpracy w celu zapewnienia wzajemnego dostępu użytkowników do usług poprzez dedykowany model wymiany danych. Warto dodać, że podjęcie prac nad OAI-PMH wiązało się z kontrowersyjną decyzją zarzucenia eksperymentów z użyciem wcześniejszych standardów, m.in. z39.50, InfoBus [1995] oraz Dienst Protocol [1995][1]. Istotną rolę odegrały też prace grupy zaangażowanej w rozwój metadanych. Główną rolę odegrała w tym zakresie wspólna grupa OCLC i NCSA, która w 1995 roku zainicjowała powstanie Dublin Core Metadata Initiative (DCMI). Rozwijany w ramach DCMI standard Dublin Core, który prezentowano na cyklicznych warsztatach szybko wzbudził zainteresowanie społeczności międzynarodowej i zyskał dużą popularność wśród twórców bibliotek cyfrowych. Fakt ów sprawił, że Dublin Core zaczął w zastosowaniach cyfrowych stopniowo wypierać wcześniejsze specyfikacje (np. MARC21 i jego odmiany) jeszcze na długo przed jego oficjalną premierą (2001). Istotną rolę w rozwoju bibliotek cyfrowych spełniły projekty realizowane m.in. przez

[1] DIENST funkcjonuje w trzech znaczeniach: 1. koncepcja architektoniczna rozproszonych bibliotek cyfrowych; 2. protokół komunikacyjny w tej architekturze; 3. oprogramowanie, które realizuje funkcje tego protokołu. Zasadniczo Dienst to system konfiguracji szeregu indywidualnych usług na serwerach rozproszonych do współpracy między bibliotekami cyfrowymi – http://www.cs.cornell.edu/cdlrg/dienst/DienstOverview.htm [odczyt: 21.05.2011].

Virginia Tech i partnerów, szczególnie dwa repozytoria: baza preprintów z zakresu informatyki *Networked Computer Science Technical Reference Library* [http://www.ncstrl.org/] oraz międzynarodowa baza dysertacji *The Networked Digital Library of Theses and Dissertation* [http://www.ndltd.org/], działająca od 1996 r. Ogromny wpływ na kierunki rozwoju bibliotekarstwa cyfrowego wywarł też projekt edukacyjny *The National Science Digital Library* [http://nsdl.org/] realizowany przez kilka organizacji i udostępniony w 2000 r. [Madalli 2003].

Prace wdrożeniowe na wyspach brytyjskich koncentrowały się wokół projektów koordynowanych przez *Joint Information Systems Committee* (JISC) [http://www.jisc.ac.uk/] od 1993 r. Pierwsza faza prac (do 1998 roku) pod nazwą *eLib - The Electronic Libraries Programme* [http://www.ukoln.ac.uk/services/elib/], obejmowała kilkanaście projektów dotyczących szeroko pojętej digitalizacji, m.in. bibliotek hybrydowych (np. *Agora, Malibu*), systemów zabezpieczeń (CEDRAS) oraz szeroko pojętej technologii dokumentów cyfrowych (czasopisma elektroniczne, druk na żądanie, preprinty). W 1996 r. powołano też specjalny dwumiesięcznik *Ariadne* [http://www.ariadne.ac.uk/]. Inne prace ogniskowały się także w ramach powołanej przez *JISC Resource Discovery Network* [http://www.rdn.ac.uk/] [Hiom 2006]. Szerszą działalność JISC rozwinęła jednak dopiero po 2003 r. W Europie w latach 1994-1998 najważniejsze prace koncentrowały się w ramach II etapu projektu UE *Telematics for Libraries Program* [109 projektów] [Telematics for Libraries / Project: http://cordis.europa.eu/libraries/en/projects.html; odczyt: 19.06.201] oraz grupy roboczej DELOS [Raitt 2000]. Wysiłki tej grupy skupiły się początkowo nad pracami koncepcyjnymi, które twórczo rozwinięto w latach późniejszych, dopiero po 2001 r., czyli po przekształceniu grupy w sieć DELOS Thematic Network [Candela; Castelli; Pagano 2011].

W innych krajach wysiłki koordynacyjne podejmowały najczęściej biblioteki narodowe, np. w Kanadzie rolę tę pełnił afiliowany przy bibliotece narodowej spis - *Canadian Inventory of Digital Initiatives* (1999) [http://publications.gc.ca/]; w Australii - *Australian digitisation Project* [http://www.nla.gov.au/libraries/digitisation/] (od 1998); w Azji i Pacyfiku – doroczna konferencja ICADL [http://www.icadl.org/] (1998), zrzeszająca rozwinięte państwa regionu (Australia, Japonia, Chiny, Korea, Indie, Tajlandia, Wietnam, Indonezja). Osobno wysiłki Indii i Chin wspierało kilka projektów amerykańskich, szczególnie inicjatywa *Universal Digital Library* [http://www.ulib.org/] rozpoczęta przez Carnegie Mellon University oraz wspomniany wcześniej NTDL [http://www.ndltd.org/] wspierany przez Science and Technology Collaboration. Inne projekty azjatyckie to m.in. pilotażowy program malezyjski prowadzony przez Malaysian National Digital Library [http://www.mylib.com.my/) (od 2000) oraz koreański rozwijany od 1996 r. przez National Digital Library of Korea [http://www.dlibrary.go.kr/].

Duże znaczenie w rozwoju wiedzy o bibliotekach cyfrowych miały także międzynarodowe konferencje, które odbywały się cyklicznie już od połowy lat 90. Najaktywniejsze było środowisko badaczy amerykańskich, kanadyjskich i australijskich, które od 1994 r. cyklicznie urządzało branżowe imprezy. Najpierw organizowano je spontanicznie (1994-1995), potem osobno przez IEEE Computer Society (1995-

2000) oraz Association for Computing Machinery (1996-2000), a następnie wspólne pt. *Joint Conference on Digital Libraries* (JCDL) [http://www.jcdl.org/] (razem 24 konferencje). Równie bogate tradycje ma związana z DELOS impreza pt. *European Conference on Digital Libraries* (ECDL) [http://ecdlconference.isti.cnr.it/], która jest głównym europejskim forum z zakresie bibliotek cyfrowych; od 1997 r. odbyło się 14 edycji, których efektem były poważne prace teoretyczne drukowane w ramach prestiżowej serii Springera – *Lecture Notes in Computer Science*. Region Azji i Pacyfiku reprezentuje natomiast doroczna *International Conference on Asian Digital Libraries* [http://www.icadl.org/], która odbywa się od 1998 r. (razem 13 edycji).

Prócz wyliczonych uprzednio projektów (zwykle inspirowanych przez rządy lub organizacje naukowe) warto wspomnieć osobno o niektórych dużych realizacjach, które powstawały z inspiracji samych bibliotek. Wśród ok. 300-500 bibliotek cyfrowych działających w latach 1995-2001 najwyższy poziom reprezentowały amerykańskie projekty realizowane przez członków DLF[2]. Biblioteki te bowiem jako pierwsze testowały i wdrażały najnowsze rozwiązania oraz szybko powiększały zbiory. Spośród dużych przedsięwzięć, które z niewielkimi zmianami działają współcześnie na uwagę zasługują m.in. przedsięwzięcia Library of Congress (głownie *American Memory*), część serwisu Internet Archive oraz francuska *Galica*. Prace pilotażowe nad projektem *American Memory* (AM) [http://memory.loc.gov/ammem/] prowadzono w Library of Congress już w latach 1990-1994, zaś oficjalny debiut nastąpił w 1994 r. Portal gromadzi zasób, który odzwierciedla kanon kulturowy USA (głownie teksty i ikonografia) i składa się ze 140 części. Charakterystycznym rysem AM jest heterogeniczność kolekcji. W związku z tym, iż poszczególne części AM realizowały różne zespoły w czasach, gdy nie istniały jeszcze stabilne standardy, prawie każda kolekcja nosi na sobie indywidualne piętno twórców. Fakt ten utrudnia nawigację i jest uciążliwy dla czytelnika. Aby nieco złagodzić tę ułomność całość wyposażono z czasem w dedykowaną wyszukiwarkę. Drugą wadą AM jest archaiczne podejście do formatów dokumentów, gdyż nawet w zbiorach typowo tekstowych są one prezentowane albo jako grafika rastrowa, albo HTML. Mimo pewnych ułomności projekt zyskał wysoką renomę i szybko się rozwijał (ponad 9 milionów jednostek), dając przykład innym. Podobne podejście legło u podstaw narodowego projektu francuskiego *Gallica* [http://gallica.bnf.fr/] zainicjowanego w 1997 r. przez Bibliothèque nationale de France (BnF). Serwis prezentuje podejście typowe dla książnicy narodowej i gromadzi szeroko pojętą literaturę oraz multimedia francuskie i Francji dotyczące (razem ponad 1,4 mln jed-

[2] Wykaz członków: Bibliotheca Alexandrina; British Library; California Digital Library; Carnegie Mellon University; Columbia University; Cornell University; Council on Library and Information Resources; Dartmouth College; Emory University; Harvard University; Indiana University; Johns Hopkins University; Library of Congress; Massachusetts Institute of Technology; New York Public Library; New York University; North Carolina State University; Oxford University; Pennsylvania State University; Princeton University; Rice University; Stanford University; University of California, Berkeley; University of California, Los Angeles; University of Chicago; University of Illinois – Urbana-Champaign; University of Michigan; University of Minnesota; University of Pennsylvania; University of Southern California; University of Tennessee; University of Texas – Austin; University of Virginia; University of Washington; U.S. National Archives and Records Administration; U.S. National Library of Medicine; Yale University – źródło: http://old.diglib.org/about.htm [odczyt: 21.05.2011].

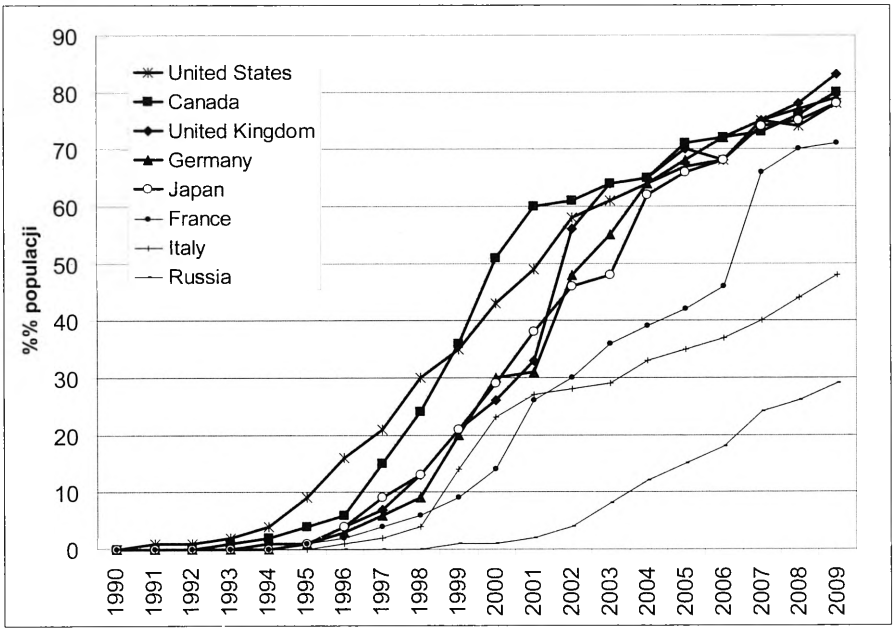
nostek). W odróżnieniu od AM *Gallica* jest jednak serwisem nowocześniejszym, który od niemal samego początku jest blisko związany z innymi serwisami BnF, oferuje doskonałe mechanizmy wyszukiwawcze oraz dostęp do zasobu w kilku wariantach i wspiera idee współpracy z użyciem protokołu OAI. Zupełnie inny model reprezentuje działające od 1996 r. serwis *Interent Archive* (IA) [http://www.archive.org/], którego znaczącą częścią jest kolekcja książek (IA/Texts). Serwis ten nie wytwarza własnych zasobów, lecz gromadzi kopie lustrzane z bibliotek współpracujących, które są replikowane w wielu formatach. Właściwy rozwój IA nastąpił dopiero od 2005 r., gdy założyciel przystąpił do *Open Content Alliance*, od tego czasu jego zasoby wzrosły z nieco ponad 150 tys. do ok. 2,8 mln. i nastąpiła integracja z *Open Library* [http://openlibrary.org/]. Warto też dodać projekty, które inspirowała inicjatywa UNESCO *Memory of the World*. Wśród nielicznych owoców tego przedsięwzięcia pozytywnym wyjątkiem jest m.in. czeski serwis udostępniający książki historyczne *Manuscriptorum* [http://manuscriptorium.com/] (od 1992) oraz kolekcja historyczna *Memory of the Netherlands* [http://www.geheugenvannederland.nl/], która może uchodzić za wzór narodowego repozytorium historycznego (od 1999). Spośród innych serwisów o zasięgu globalnym warto wymienić przedsięwzięcie komercyjne *World Public Library* [http://worldlibrary.net/], działające od 1996 r. (2 mln jednostek); serwis ten – co warto podkreślić – wyprzedził podobne przedsięwzięcia o niemal pełną dekadę.

Biblioteki cyfrowe po 2002 roku. Współczesność

U progu XXI wieku bibliotekarstwo cyfrowe wkroczyło w nowy etap rozwoju. Tendencja ta staje się łatwo uchwytna jeśli za miernik postępu przyjąć liczbę funkcjonujących serwisów i wzrost dostępnych obiektów (rys. 1). Źródeł tego zjawiska można dopatrywać się w kilku płaszczyznach. Z jednej strony zadziałał tu pozytywnie rozwój technologii i osiągnięcia organizacyjne w sferze bibliotek cyfrowych, czyli tendencje, które przedstawiono wcześniej. Z drugiej jednak – nałożył się na to ogólny trend rozwoju usług online, którego obiektywnym miernikiem wydaje się być wzrost liczby użytkowników. Na przestrzeni 2000-2003 nastąpił bowiem proces, który można nazwać przekroczeniem progu umasowienia, gdyż z usług sieciowych korzystało już ponad 50% populacji w najbogatszych krajach świata (rys. 3). Fakt ów był niewątpliwie motorem dalszych zmian, w szczególności przeobrażeń sposobu komunikacji w internecie i mechanizmów generowania treści. *Signum temporis* tego przełomu jest symboliczna data narodzin internetu określanego jako web 2.0 (2001).

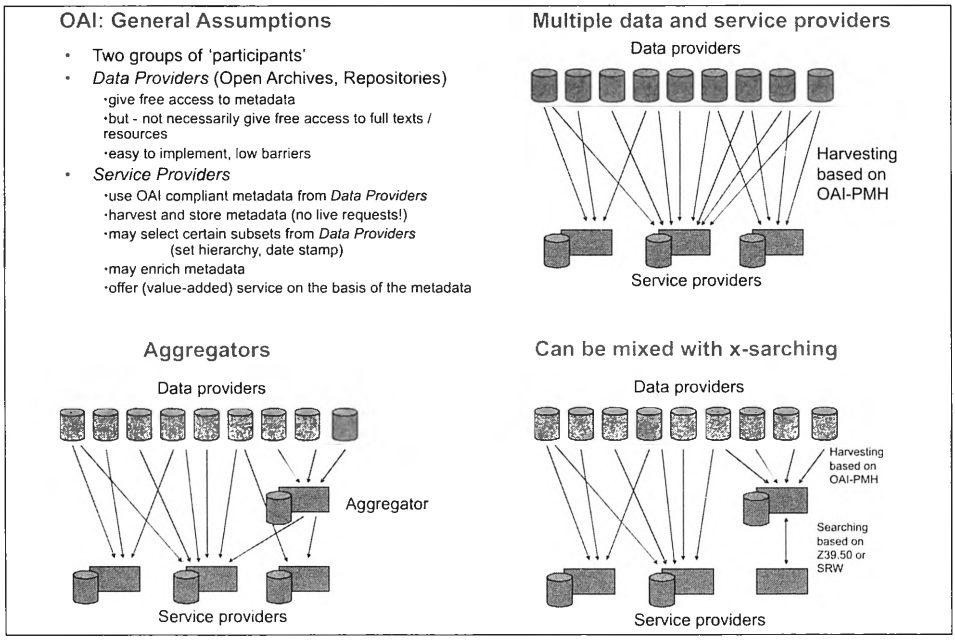
Z technicznego punktu widzenia idea web 2.0 oznaczała stosowanie nowych technologii zorientowanych na użytkownika, choć – jak czas pokazał – najistotniejszy okazał się aspekt społeczny, który oznaczał zrównanie użytkowników i twórców. W środowisku web 2.0 pojawiły się też inne zjawiska, istotne dla bibliotek cyfrowych, jak: tworzenie społeczności sieciowych, trend do używania folksonomii, nowy rodzaj komunikacji i zachowań; od strony technicznej zaś – powszechne dążenie do integracji wszelkich usług sieciowych oraz komunikacja maszynowa. Nie ulega wątpliwości, że idea web 2.0 postawiła bibliotekarzy w nowej sytuacji. Niekontrolowany zalew

łatwo dostępnych informacji generowanych przez użytkowników stał się realnym zagrożeniem dla tradycyjnego modelu biblioteki. Aby temu przeciwdziałać twórcy serwisów bibliotecznych zaczęli aktywnie uczestniczyć w środowisku web 2.0 oferując informację kwalifikowaną. Pierwszym i najłatwiej uchwytnym aspektem przemian w tym zakresie był powszechny trend do tworzenia serwisów rozproszonych.



Rys. 3. Użytkownicy internetu w krajach G8 (1990-2009). Przeliczenia na podstawie danych: The United Nations Statistics Division – http://data.un.org/ [odczyt: 28.05.2011].

Nad ideą komunikacji pomiędzy bibliotekami cyfrowymi podjęto prace już w połowie lat 90. ubiegłego wieku, zaś ich ukoronowaniem był opracowany w ramach Open Archive Initiative dedykowany protokół OAI-PMH (1999). Protokół ten [Open Archive Initiative Protocol for Metadata Harvesting] służy do selektywnego pobierania metadanych opisujących obiekty cyfrowe; umożliwia on automatyczną wymianę informacji pomiędzy systemami gromadzącymi takie metadane, a systemami chcącymi te dane uzyskać. Jest przy tym bardzo elastyczny, gdyż dane są zapisane w otwartym standardzie XML (zwykle RDF). Idea jego działania jest zatem zupełnie inna niż *z39.50 czy SRU*, gdyż nie funkcjonuje on w trybie rzeczywistym.



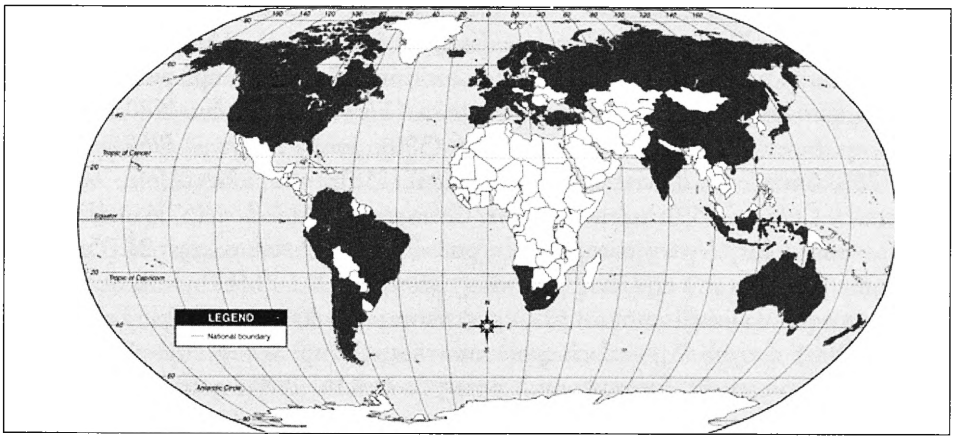
Rys. 4. Model biblioteki rozproszonej OAI-PMH [Hunter 2005].

W praktyce centrum serwisu rozproszonego [service provider] pełni w tym rozwiązaniu dedykowany serwer [harvester], który okresowo pobiera metadane z określonych bibliotek [data provider] (rys. 4). Serwis providera indeksuje własnymi mechanizmami metadane i pełni rolę bramki dostępowej dla użytkownika. Warto dodać, że mogą istnieć dwa rodzaje dostawców: pojedynczy serwis lub tzw. agregator – czyli lokalny pośrednik. Drugi typ dostawcy stwarza wielkie możliwości, gdyż biblioteki mogą się łączyć w dowolne grupy i elastycznie współpracować.

Największą udaną realizacją serwisu opartego na OAI-PMH był portal *OAISTER*, uruchomiony w lipcu 2002 r. na University of Michigan. Do października 2009 r. działał samodzielnie, następnie stał się częścią serwisu OCLC *Worldcat* [http://www.worldcat.org/], istnieje też odrębna bramka dostępowa [http://oaister.worldcat.org/]. Rozwój portalu następował szybko (rys. 1). W grudniu 2006 r. zrzeszał już 726 dostawców oferujących prawie 10 mln jednostek [Garrison 2010], najwięcej dostawców pochodziło z USA (272), Wielkiej Brytanii (78), Niemiec (72), Francji (32), Kanady (27), Brazylii (23), Włoch i Australii (po 21), Holandii (19), Hiszpanii (18), Japonii (16), Szwecji (14), Polski (12) oraz innych państw Europy (55)[3]. W maju 2011 r.

[3] OAIster Repositories: Argentina (3), Australia (21), Austria (6), Belgium (9), Brazil (23), Bulgaria (1), Canada (27), Chile (5), China (2), Columbia (2), Croatia (1), Denmark (3), England (63), Estonia (1), Finland (7), France (32), Germany (72), Greece (2), Hong Kong (3), Hungary (3), Iceland (1), India (11), Indonesia (1), Ireland (3), Italy (21), Japan (16), Lithuania (1), Namibia (1), Netherlands (19), New Zealand (3), Northern Ireland (1), Norway (1), Peru (3), Poland (12), Portugal (5), Russia (1), Scotland (13), Serbia (1), Slovenia (1), South Africa (4), Spain (18), Sweden (14), Switzerland (6),

zrzeszał ponad 1100 dostawców, oferując ponad 25 mln rekordów. Zaprezentowany rozkład, ze względu na reprezentatywność może stanowić dobrą podstawę do uogólnień geograficznych (rys. 5). Z przedstawionych danych wynika w szczególności lapidarna teza, że biblioteki cyfrowe funkcjonują jedynie w państwach rozwiniętych i rozwijających się oraz daleko od ognisk konfliktów zbrojnych.



Rys. 5. Biblioteki cyfrowe na świecie w rozkładzie terytorialnym (2006). Źródło: OAISTER – http://www.oaister.org/stats.html [odczyt: 13.04.2011].

Prócz bazy *OAISTER* powstało wiele innych serwisów rozproszonych o różnej wielkości. Oficjalny wykaz na stronach OAI [Registered Service Providers: http://www.openarchives.org/service/listproviders.html; odczyt: 21.05.2011] wylicza ich ponad 30, lecz z cała pewnością jest ich znacznie więcej, niektóre zaś pełną wyłącznie funkcję agregatów dla innych portali. Wśród samodzielnych serwisów na szczególną uwagę zasługują portale naukowe, m.in.: serwis *Scirius* [http://www.scirus.com] należący do Elseviera (410 mln jednostek, głownie artykułów); *ScientificCommons* [http://www.scientificcommons.org/], założony w 2007 r. na szwajcarskim Universität St. Gallen (38 mln jednostek, 1269 dostawców), repozytorium dysertacji *DART-Europe* [http://www.dart-europe.eu/] (338 dostawców) czy rozbudowana w 2006 r. wersja serwisu *Perseus* [http://www.perseus.tufts.edu/], która scala zasoby filologiczne. Nadto liczne serwisy narodowe: holenderski *DARE* (od 2004), norweska *NORA* [http://www.ub.uio.no/nora/] (2004); hiszpańska sieć muzealna *Hispana* [http://hispana.mcu.es/] (2006; 148 dostawców, 3 mln jednostek) czy polska sieć *Federacja Bibliotek Cyfrowych* [http://fbc.pionier.net.pl/] (2007; 69 dostawców, 648 tys. jednostek). Z kolei spisy dostawców, których aktualna liczba sięga zapewne 2-3 tysięcy gromadzi kilka dedykowanych serwisów, m.in.: *OpenDOAR* [http://www.opendoar.org/] (1966 dostawców) czy *Openarchives.eu*: [http://www.openarchives.eu/] (1200 dostawców).

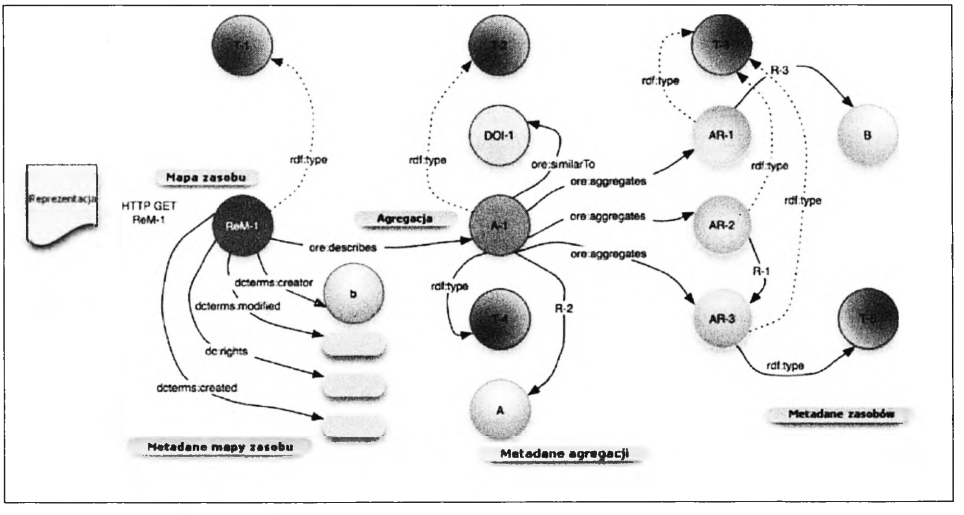
Na osobną uwagę zasługuje jedna z największych realizacji tego typu – serwis *Europeana* [http://www.europeana.eu/], który uruchomiony w listopadzie 2008 r. Projekt

Taiwan (2), Turkey (1), United States (272), Venezuela (2), Wales (2), West Indies (1) – http://www.lib.umich.edu/digital-library-production-service-dlps/oaister-statistics [odczyt: 21.05.2011].

stał się realizacją inicjatywy *i2010* finansowanej przez Komisję Europejską w ramach programu *eContentplus* (2005-2008) i ma na celu zgromadzenie całego dziedzictwa kulturowego Europy (zbiory biblioteczne, muzealne i archiwalne). Warto dodać, że projekt funkcjonował w postaci pilotażowej już od 2006 roku jako *EDLnet* [http://edlproject.eu/] (Ayris 2009). Podstawą *Europeany* były zasoby nieudanego przedsięwzięcia UE *The European Library*, które realizowano w latach 2001-2004 [http://www.theeuropeanlibrary.org]. Aktualnie *Europeana* zrzesza ok. 1500 instytucji (ok. 200 dostawców oferujących ponad 15 mln obiektów). Znaczną grupę partnerów (16) stanowią narodowi agregatorzy, m.in.: francuska *Culture.fr/Collections* (480 instytucji, w tym np. Luwr), brytyjska *Collections Scran* (380); polska *Federacja Bibliotek Cyfrowych* (250); niemiecka *Bayerische Staatsbibliothek* (158); rumuńska *cIMeC - Institutul de Memorie Culturala* (59); holenderska *Het Geheugen van Nederland* (57) i in. [Europeana / partnerzy: http://www.europeana.eu/portal/partners.html; odczyt: 21.05.2011].

Na marginesie warto dodać, że omawiany protokół OAI-PMH był szeroko wykorzystywany do lokalizacji zasobów przez popularne wyszukiwarki (*Google, Yahoo*) oraz do aktualizacji danych w portalach społecznościowych (np. *Wikipedii*).

Realizacje projektów rozproszonych opartych na OAI-PMH mają jednak pewną słabość. Protokół ten nie potrafi poprawnie reprezentować struktury obiektu wielopoziomowego i powiązań z innymi obiektami oraz wizualizować tych zależności. Wyjściem naprzeciw temu problemowi jest nowa specyfikacja OAI-ORE (ang. Open Archive Initiative – Object Reuse and Exchange) [OAI-ORE: http://www.openarchives.org/ore/; odczyt: 23.05.2011] ogłoszona w 2008 r. Specyfikacja ta (rys. 6) wprowadza nowy typ zasobu – agregację, która identyfikuje oraz reprezentuje zbiór powiązanych obiektów. Ma ona własny URI, lecz nie zgodnie z postulatami sieci semantycznej, nie posiada swojej reprezentacji. W tym zakresie agregację uzupełnia mapa zasobu utrzymana w postaci czytelnej do przetwarzania maszynowego czyli RDF.



Rys. 6. Schemat komunikacji przy użyciu OAI-ORE [Lewandowska; Werla 2009].

W praktyce więc mapa zasobu kieruje do agregacji, zaś ta na wskazuje na konkretny zasób. W mapie zasobu można zawrzeć: 1. opis relacji między obiektami; 2. informacje o własnościach konkretnych zasobów; 3. informacje o własnościach samej mapy zasobu [Lewandowska; Werla 2009]. Protokół nie doczekał się wprawdzie jeszcze większych wdrożeń, lecz jego możliwości są obiecujące i szeroko dyskutowane w kręgach specjalistów, np. *DELOS* i *Google*. Innym przykładem dużego systemu rozproszonego jest projekt *DRIVER* [http://www.driver-community.eu/] scalający zasoby naukowe. Jest to wynik prac dwóch projektów finansowanych przez Komisję Europejską w latach 2006-2009. Z technicznego punktu widzenia jest on najbardziej innowacyjny, gdyż opiera się na tzw. D-Net technology, czyli architekturze zorientowanej na usługi. Aktualnie serwis ten daje dostęp do ok. 3,5 mln z 38 krajów (295 repozytoriów).

Poza systemami rozproszonymi czołową rolę odgrywają współcześnie duże serwisy międzynarodowe, które można określić, jako megaprojekty. Do grupy tej należy zaliczyć: *Google Books* (2004), *Internet Archive/OCA* (2005), *Universal Digital Library – Million Book Collection* (2001) oraz omówioną wcześniej *Europeanę*. Istnieje nadto znaczna liczba projektów narodowych, regionalnych, specjalistycznych oraz trudna do oszacowania liczba projektów instytucjonalnych.

Jedną z największych współczesnych inicjatyw digitalizacyjnych był projekt *Google Print*, rozpoczęty formalnie jesienią 2004 r. (od 2005 r. pod nazwą *Google Books*). Na wstępie warto przypomnieć, że zainteresowanie bibliotekami twórców *Googla* Lawrence'a Page i Sergeya Brina nie było przypadkowe. Obaj rozpoczynali bowiem swą karierę naukową jako doktoranci w Stanford pracując przy projekcie *Stanford Digital Library Project* (część *Digital Library Initiative*). *Google* tuż po ogłoszeniu planu uruchomienia nowej usługi intensywnie zabiegał o współpracowników posiadających zasoby i w przeciągu trzech lat (2005-2007) pozyskał jako partnerów 19 dużych bibliotek amerykańskich i europejskich[4]. Szczegóły umów nie są znane, lecz wiadomo, że biblioteka partnerska otrzymywała cyfrową kopię każdej skanowanej przez *Google* książki z jej zbiorów. Prócz wyliczonych bibliotek do projektu w kolejnych latach przyłączyło się wiele innych podmiotów, w szczególności wydawców, którym Google oferował korzystne warunki współpracy [Toobin 2007]. Udostępniany zasób rósł szybko: w listopadzie 2008 r. osiągnięto 7 mln, zaś w czerwcu 2010 – 12 mln; wynik ten ośmielił twórców do karkołomnej deklaracji, wg której serwis zobowiązał się do zeskanowania w ciągu następnej dekady (2010-2020) wszystkich znanych książek na świecie – czyli ok. 130 mln. Serwis *Google Books* od początku swego działania wzbudzał liczne kontrowersje natury prawnej [Franke 2007]. Już rok po debiucie skierowano przeciw portalowi pierwsze pozwy zbiorowe w Stanach Zjednoczonych (w 2008 zakończyły się one ugodą); w następnych latach kolejne procesy (instytucje

[4] M.in.: Bavarian State Library, Princeton University, University of California, The National Library of Catalonia, University Complutense of Madrid, Harvard University, University of Michigan, The New York Public Library, Oxford University, Stanford University, University of Texas at Austin, University of Virginia, University of Wisconsin–Madison – http://books.google.pl/googlebooks/partners.html [odczyt: 21.05.2011].

w Niemczech, Francji i duże korporacje, np. Microsoft). Część z nich została zakończona ugodami, inne są w toku. System *Google Books* wydaje się być jednak skutecznie zabezpieczony przed łamaniem prawa, gdyż udostępnia książki w trzech trybach: pełnym – tylko pozycje z domeny publicznej; oraz ograniczonym – widoczne są fragmenty lub cytaty z zawartości. Abstrahując od innych cech serwisu warto podkreślić, że *Google Books* nie jest biblioteką cyfrową *sensu stricto*, lecz hybrydą (księgarnio-biblioteką), gdyż książki inne niż z domeny publicznej mogą być na wniosek wydawcy w każdej chwili usunięte z serwisu. Niemniej liczba pozycji z domeny publicznej jest znaczna (ok. 1 mln w 2008 r. i ok. 1,7 mln w 2010 r.), co sprawia że *Google Books* jest liczącym się graczem w omawianym obszarze. Niezaprzeczalnym walorem serwisu jest przytłaczająca popularność samego serwisu matki – wyszukiwarki *Googe* (ok. 90-93% rynku) [Search Engine Barometer: http://en.atinternet.com/Resources/Surveys/Search-engine-barometer/index-1-2-6-0.aspx; odczyt: 21.05.2011], jak i innych usług specjalistycznych, wśród których wschodzący *Google Scholar* zaczyna pełnić coraz donioślejszą rolę w świecie nauki.

W październiku 2005 r. w odpowiedzi na inicjatywę Google podjęto konkurencyjny projekt pod auspicjami konsorcjum *Open Content Alliance* (OCA), które zostało utworzone w wyniku porozumienia pomiędzy serwisem *Internet Archive*, *Yahoo*! oraz Adobe System, European Archive, University of Kalifornia, HP Labs, brytyjskim The National Archive) i wieloma innymi. Konsorcjum działa w oparciu o współpracę organizacji rządowych, non-profit, kulturalnych i technicznych z całego świata i ma zadanie trwałe zabezpieczenie szeroko pojętego dziedzictwa kulturowego [Urbaniec 2006]. OCA nie tworzył jednak odrębnego portalu, lecz rozbudował istniejący wcześniej serwis należący do *Internet Archive* (IA) [http://www.archive.org/]. Portal ten założył w 1996 r. Brewster Kahle dzięki dotacji Alexa Internet i miał początkowo na celu archiwizację szeroko pojętych zasobów internetowych (stron web, plików programów, wideo i audio oraz tekstów). Z punktu widzenia bibliotek interesujący jest jedynie serwis tekstowy (działający od 1999 r.), w którym udostępniane są zbiory książkowe. Serwis ten w sensie funkcjonalnym jest biblioteką cyfrową w ścisłym znaczeniu (jest nawet członkiem American Library Association). W 2005 r., po powstaniu OCA utworzono przy IA kilka grup roboczych (m.in. do straw metadanych, ochrony, formatu, skanowania i protokółów komunikacyjnych), nade wszystko jednak przystąpiono do rozbudowy zasobów. Zrąb główny portalu tworzą dzieła zdigitalizowane przez instytucje partnerskie (23 biblioteki i archiwa) [OCA Contributors: http://www.opencontentalliance.org/contributors/; odczyt: 21.05.2011] oraz zasoby z domeny publicznej pozyskane legalnie z innych portali (np. z *Google Books*). Aktualnie dominują zasoby z domeny, lecz udostępniana jest też pewna liczba zbiorów współczesnych objętych prawami, dla których autorzy udzielili wyraźną czytelną zgodę na publiczne wykorzystanie (dla niewielkiej części dostęp jest limitowany). Na początku 2005 r. zbiór książkowy IA był niewielki i liczył nieco ponad 25 tys., jednak po utworzeniu OCA szybko się powiększył (ok. 1 mln w 1998 r. oraz 2,8 mln w 2011 r.); warto przy tym dodać, że do zbioru IA od 2007 r. włączone się książki z domeny publicznej dostępne w *Google Books* (ok. 900 tys.). Warto też podkreślić, że IA

przyjęło specyficzne rozwiązanie techniczne, polegające na tworzeniu kopii zasobów pobranych od instytucji partnerskich i ich deponowania na własnej platformie w wielu formatach. Ostatnia cecha bardzo korzystnie wyróżnia IA na tle innych serwisów. Książki są tu bowiem dostępne w wielu równoległych wersjach: online, czyli renderowanie przez flasha, PDF, DJVu, czysty tekst, HTML, formaty mobilne (ePub, kindle, mobi) oraz przeznaczonym dla niewidomych i niedowidzących format DAISY (od 2009 r. dostępnych jest nieco ponad 1 mln pozycji w tym formacie). W miarę rozwoju projektu funkcjonalność kolekcji IA została poszerzona, gdyż zbiory zostały zintegrowane z *Open Library*, projektem rozwijanym przez OCA od 2006. *Open Library* jest uniwersalnym serwisem rozproszonym (funkcjonalnie zbliżonym do katalogu *WorldCat*), dla którego dostawcami danych stały się katalogi bibliotek amerykańskich i kanadyjskich (przede wszystkim Library of Congress) oraz dane Amazon [O'Leary 2009].

The Universal Digital Library – Million Book Collection (UDL) [http://www.ulib.org/] – to inicjatywa, nad którą prace rozpoczęły się w 2001 r., a oficjalny debiut nastąpił we wrześniu 2003 r. Projekt ów jest wspólnym przedsięwzięciem amerykańsko-chińsko-indyjskim. Ze strony amerykańskiej prace koordynuje Carnegie Mellon Uniwersity, zaś technicznie jest realizowany głównie w Zhejiang University (Chiny) i Indian Institute of Science (Indie) oraz siedmiu chińskich i ośmiu indyjskich uczelniach partnerskich; w projekcie uczestniczą też egipska Bibliotheca Alexandrina. UDL koncentruje się na książkach z domeny publicznej oraz dziełach, dla których pozyskano stosowne umowy od autorów i wydawców. Do grudnia 2007 r. udostępniono ponad 1,5 mln dzieł, wśród których dominują publikacje chińskie – 970 tys., nieco mniej udostępniono w języku angielskim – 360 tys., kilku językach indyjskich (telungu – 50 tys.; urdu – 17 tys.; sanskryt – 7 tys. i in.) oraz arabskim – 40 tys., perskim – 7 tys. i kilku innych. W UDL zastosowano kilka interesujących osobliwości technicznych: baza danych istnieje na trzech serwerach lustrzanych (USA, Chiny, Indie), serwis jest wyposażony w wielojęzyczny interfejs (w tym języki wschodnie), zaś publikacje są poddawane OCR – co dla języków azjatyckich jest dużym wyzwaniem, a zasób serwowany jest w kilku formatach, m.in. djvu i TIFF; do obsługi ostatniego typu służy dedykowana wtyczka AlternaTIFF [http://www.alternatiff.com/]. W listopadzie 2005 r. Indie, Chiny i USA uzgodniły, że UDL do przyłączy się do *Open Content Alliance*; w efekcie zbiory UDL stały równolegle dostępne w *Internet Archive* oraz katalogu *Open Library*.

Wśród licznych inicjatyw o charakterze międzynarodowym należy wspomnieć o kilku mniejszych projektach, m.in. inicjatywie Library of Congress i UNESCO – *World Digital Library* (WDL) [http://www.wdl.org/] (od 2009 r., ok. 1300 poz.); *International Children's Digital Library Foundation* (ICDL) [http://en.childrenslibrary.org/], tworzonej od 2002 r. na University of Maryland (ok. 2 tys. pozycji), która istotnie wzbogaciła zasób po 2006 r., gdy uzyskała finansowanie m.in. z Library of Congress i Microsoftu.

Liczba znaczących projektów o charakterze regionalnym i narodowym jest trudna do ustalenia. Punktem wyjściowym do oszacowania ich liczby może być baza

UNESCO i IFLA *Directory of Digitized Collections* [http://www.unesco.org/webworld/digicol/], związana z inicjatywą *Memory of the World*. Źródło to jest wprawdzie dalekie od kompletności (238 bibliotek), ale z uwagi na staranny dobór w dobrym przybliżeniu odzwierciedla proporcje regionalne. Według stanu z 2007 r. na Europę i Amerykę Północną przypadało 83% bibliotek cyfrowych, Azję i Pacyfik – 15%, Amerykę Południową – 1,6% oraz Afrykę – 0,4%. Dane te potwierdza omawiany wyżej *OAISTER* – gdzie na Europę i Amerykę Północną przypadało 88% bibliotek cyfrowych (w tym 47% – Europa 41% i Ameryka)[5].

W największym pod względem inicjatyw regionie – czyli Europie zdecydowana większość istniejących bibliotek uczestniczy w roli partnerów omówionej wcześniej *Europeany*. W przedsięwzięciu tym partycypuje ok. 200 bibliotek cyfrowych, które udostępniają zasoby samodzielnie oraz ponad tysiąc, które czynią to za pośrednictwem sieci agregujących (16 sieci) [Eurpeana Partners: http://www.europeana.eu/portal/partners.html; odczyt: 21.04.2011]. W grupie tej można wyróżnić kilka typów dostawców. Największą grupę tworzą repozytoria i agregatorzy narodowi, np. *Gallica* (Francja) [http://gallica.bnf.fr/]; *Digital Library of Slovenia* (Słowenia) [http://www.dlib.si/v2/]; *The British Library* (Wielka Brytania) [http://www.bl.uk/eresources/main.shtml]; *Federacja Bibliotek Cyfrowych* (Polska) [http://fbc.pionier.net.pl/]; *Biblioteca Nacional Digital* (Portugalia) [http://bnd.bn.pt]; *NLL Digital Library* (Łotwa) [http://www.lnb.lv/digitala_biblioteka/index_anglu.htm]; *Zentrales Verzeichnis Digitalisierter Drucke* (Niemcy) [http://www.zvdd.de/startseite/]; *Электронная Ленинка РГБ* (Moskwa) [http://www.rsl.ru/] oraz *Электронная библиотека РНБ* (St. Petersburg) [http://www.nlr.ru/eng/] (obie w Rosji); *Magyar Elektronikus Könyvtár* (Węgry) [http://mek.oszk.hu/indexeng.phtml]; *Helmi Digitised* (Finlandia) [http://www-db.helsinki.fi/helmi/en/]. Inny, dość liczny typ biblioteki to narodowe repozytoria czasopism, np.: *ANNO AustriaN Newspapers Online* (Austria) [http://anno.onb.ac.at/]; *Deutsche Exilpresse 1933-1945 Digital* (Niemcy) [http://deposit.ddb.de/online/exil/exil.htm]; *EPA Elektronikus Periodika Adatbázis* (Węgry) [http://epa.oszk.hu/]; *Internet Library of Early Journals* (Wielka Brytania) [http://www.bodley.ox.ac.uk/ilej/]. Wreszcie typy specjalne; najczęściej występują narodowe repozytoria historyczne – np. holenderski *Memory of the Netherlands* [http://www.geheugenvannederland.nl/gvnnl/] [Bakker 1999] lub narodowe repozytoria literackie – np. *Poésie Française* [http://poesie.webnet.fr/home/] [Pełniejszy wykaz inicjatyw europejskich i amerykańskich zob. Kolasa 2010]. Mozaikę tę uzupełnia duża liczba mniejszych portali narodowych, specjalnych, regionalnych i instytucjonalnych.

Obraz bibliotek cyfrowych w Stanach Zjednoczonych nie jest tak skonsolidowany, jak w Europie, gdyż zabrakło narodowego reprezentanta. Niemniej większość ważnych podmiotów o charakterze narodowym skupiło się wokół Library of Congress lub *Interent Archive*. Największy zasób skupia opisany szczegółowo wcześniej projekt *American Memory* (AM) [http://memory.loc.gov/ammem/], którego uzupełnieniem jest kilka innych serwisów powstałych w czasie realizacji *National Digital*

[5] Por. przyp. 3: OAIster Repositories.

Library Program (NDLP)[6]; warto dodać że poszczególne kolekcje AM [wykaz kolekcji AM: http://international.loc.gov/ammem/dli2/html/list.html; odczyt: 21.05.2011] są tworami na tyle samodzielnymi, że są rozróżniane w literaturze, np. *Votes for Women 1848-1921*. Bardziej klarowną konstrukcję ma literatura amerykańska w serwisie *Internet Archive*; tworzy ją subkolekcja *American Libraries* [http://www.archive.org/details/americana] – która scala w jednolitym interfejsie ogromny zasób *americanów* (ok. 1,5 mln) udostępnionych przez najważniejsze książnice i muzea amerykańskie. Nie jest to bynajmniej komplet amerykańskich inicjatyw. Wiele starszych projektów funkcjonuje niezależnie, np. *The Berkeley Digital Library SunSITE* [http://sunsite.berkeley.edu/]; *The Humanities Text Initiative* [http://www.hti.umich.edu/]; *Historic American Sheet Music* [http://scriptorium.lib.duke.edu/sheetmusic/]. Istnieją też odrębne projekty tematyczne, np. kolekcja gazet *Historic American Newspapers* [http://chroniclingamerica.loc.gov/] czy serwis historyczny *Making of America* [http://quod.lib.umich.edu/m/moagrp/]. Obraz ten dopełnia ogromna liczba dużych bibliotek stanowych lub uniwersyteckich, których zbiory są niemal w komplecie dostępne w IA, wyróżniają się na tym tle: *California Digital Library* [http://www.cdlib.org/]; *The University of Pittsburgh Digital Research Library Collections* [http://digital.library.pitt.edu/]; *The Indiana University Digital Library Program* [http://www.dlib.indiana.edu/]; *Kentuckiana Digital Library* [http://kdl.kyvl.org/]; *LOUISiana Digital Library* [http://louisdl.louislibraries.org/]; *Michigan State University Libraries Digital and Multimedia Center* [http://digital.lib.msu.edu/]; *NYPL Digital Library Collections* [http://www.nypl.org/digital/]; *Oregon State University Library Digital Collections* [http://digitalcollections.library.oregonstate.edu/]; *The University of Michigan Digital Library Production Service Collection* [http://www.umdl.umich.edu/] czy *The University of Chicago Library Digital Activities and Collections* [http://www1.lib.uchicago.edu/cgi-bin/nand/search/diglist] [pełniejszy wykaz dostawców IA: http://www.archive.org/details/americana; odczyt: 21.05.2011]. Podobną strukturę mają zasoby Kanady, które skupione są w dwu serwisach: IA oraz przy książnicy narodowej. Warto podkreślić, że ogromne znaczenie w tym obszarze miał narodowy programu *The Canadian Initiative on Digital Libraries* [CIDL: http://epe.lac-bac.gc.ca/100/206/301/lac-bac/cidl-ef/2007-09-28/cidl/040021-300-e.html; odczyt: 1.06.2011], który realizowano w latach 2000-2006. Aktualnie 38 największych kanadyjskich repozytoriów (łącznie 338 tys. jednostek) tworzy zwartą kolekcję wydzieloną w IA [http://www.archive.org/details/toronto]; pozostałe inicjatywy (w tym 40 dużych kolekcji) [Canada's Digital Collections: Sharing the Canadian Identity on the Internet: http://www.collectionscanada.gc.ca/publications/002/015002-2170-e.html; odczyt: 1.06.2011] wyczerpująco prezentuje serwis [http://www.collectionscanada.gc.ca/a-z-index/index-e.html] prowadzony przez Library and Archives Canada.

Liczba bibliotek cyfrowych na kontynencie afrykańskim jest symboliczna. W zasadzie jedynym dużym przedsięwzięciem jest udział egipskiej Bibliotheca Alexandrina w opisanym uprzednio projekcie UDL [http://www.ulib.org/] oraz kilka egipskich kolekcji zlokalizowanych na jej serwerach [http://www.bibalex.org], m.in. *Modern*

[6] Por. przyp. 3: OAIster Repositories

Egipt [http://modernegypt.bibalex.org]. Inne inicjatywy afrykańskie pojawiają się incydentalnie, np. powstały na University of Natal w RPA serwis *DISA* [http://disa.nu.ac.za/]. Stan ten nieco łagodzą portale zagraniczne interesujące się Afryką. Najważniejszym projektem tego typu jest powstały w 2008 r. w USA dzięki inicjatywie serwisu JSTOR serwis *ALUKa* [http://www.aluka.org/] – portal w całości poświęcony Afryce i gromadzący *africana*. Do końca 2007 r. serwis zgromadził pokaźny zasób: 76 kolekcji, 323 tys. jednostek pochodzących z 31 krajów (w tym z 14 państw afrykańskich).

Proces tworzenia bibliotek cyfrowych w Ameryce Łacińskiej znajduje się dopiero na etapie wstępnym. W większości państw pojawiły się już kolekcje eksperymentalne, lecz nie istnieje ani jedno duże repozytorium. Według bazy *OAISER* najwięcej serwisów istnieje Brazylii (23) – są to jednak gównie repozytoria naukowe, zaś uniwersalnych kolekcji działa niewiele, np. *Biblioteca Digital de Obras Raras e Especiais* [http://www.obrasraras.usp.br/]. Stan ten jednak stopniowo ewoluuje i od niedawna pojawiają się inicjatywy skoordynowane, np. *Brazilian Government Document Digitization Project* [http://www.crl.edu/brazil]. W pozostałych państwach regionu (Chile, Boliwia, Kuba, Nikaragua, Peru, Urugwaj, Argentyna, Kolumbia, Wenezuela) istnieje po 1-5 bibliotek cyfrowych, najczęściej są to małe kolekcje eksperymentalne afiliowane przy bibliotekach narodowych lub ministerstwach, np. czylijska *Memoria Chliena* [http://www.memoriachilena.cl/elibros/e-libros.asp] lub argentyńska *Biblioteca Digital de Tratados* [http://tratados.cancilleria.gob.ar/index.php]. Obraz dopełniają inicjatywy zagraniczne, np. działający przy University of Miami portal poświęcony Kubie – *Cuban Heritage Collection* [http://www.library.miami.edu/chc/].

Główną inicjatywą azjatycką, która od 2003 roku zogniskowała wysiłki chińskie i indyjskie jest omówiona wcześniej *Universal Digital Library* (UDL) [http://www.ulib.org/]. Podkreślić jednak należy, że główną siłą koordynująca prace w tym zakresie były cykliczne konferencje ICADL, które odbywały się w regionie od 1998 r. Dynamikę przemian odzwierciedla liczba członków ICADL na przestrzeni 1998-2004. W okresie tym wzrosła ich liczba z 10 do ok. 40. Projekty działające w poszczególnych krajach są bardzo zróżnicowane, lecz przeważają projekty scentralizowane. Model taki reprezentują Chiny, gdzie prace nad głównymi inicjatywami narodowymi koordynuje rząd, wynikiem tych działań jest narodowy projekt *Digital Library of China* [http://www.d-library.com.cn/] oraz jej dopełnienie *China Digital Library for Visual Impairment* [http://www.cdlvi.cn/]. Prace przygotowawcze nad ich realizacją trwały od 2000 roku, zaś oficjalne otwarcie w 2005 r. W okresie tym na potrzeby pierwszego projektu przygotowano imponujący zasób: 1,1 mln książek, 450 tys. obrazów, 12 tys. tyt. czasopism i 600 tyt. gazet [Leye; Ping 2009]; zasób ten rośnie w imponującym tempie i do końca 2009 r. urósł do ponad 27 mln [Zhen 2010]. Główna inicjatywa w Indiach to *The Digital Library of India* [http://www.dli.cdacnoida.in/], wspólny projekt 21 narodowych i kilku zagranicznych instytucji, nad którymi prace koordynują Indian Institute of Science oraz International Institute of Information Technology (Hyderabad). Istnieją jednak przesłanki, by sądzić, że zawartość DLI pokrywa się z kolekcjami UDL [Swanepoel 2008]. Znaczne osiągnięcia w zakresie budowy infrastruktury

cyfrowej ma także Korea Południowa, gdzie istnieje 6 dużych projektów kierowanych przez National Digital Library of Korea [http://www.dlibrary.go.kr/] (od 2006). W pozostałych krajach regionu istnieją jedynie repozytoria w formie zalążkowej, np. *Filipiniana* [http://www.filipiniana.net/] czy malezyjska *myLibrary* [http://www.mylib.com.my/] (od 2000). Istotnym dopełnieniem tego obrazu są tematyczne serwisy zagraniczne; wyróżnia się w tym zakresie *Portal to Asian Internet Resources* [http://digicoll.library.wisc.edu/PAIR/] – amerykańska inicjatywa prowadzona od 2005 r. pod kierunkiem University of Wisconsin-Madison.

Odrębnie należy rozpatrywać projekty japońskie, gdyż mają one własną specyfikę. W Japonii dominują bowiem rozwiązania zbliżone do Europy lub USA, tj. działają tu w szczególności sieci informacji specjalistycznej i istnieje ogromna liczba repozytoriów instytucjonalnych. Serwisy tego typu działają przy większości ośrodków badawczych i uniwersytetów, np. *Kyoto University RIR* [http://repository.kulib.kyoto-u.ac.jp/dspace/]; *Nagoya University Academic Knowledge Factory* [http://akf.nul.nagoya-u.ac.jp/]; *Hiroshima University IR* [http://ir.lib.hiroshima-u.ac.jp/portal/index_e.html]. Istnieje nadto kilka serwisów uniwersalnych, które można uznać za narodowe; taki charakter na sześć kolekcji prowadzony przez National Diet Library [http://www.ndl.go.jp/en/data/endl.html], wśród których wyróżnia się serwis dawnych książek *Digital Library from the Meiji Era* [http://kindai.ndl.go.jp/] (150 tys. jednostek) oraz centralny portal cyfrowy archiwów japońskich: *PORTA* [http://porta.ndl.go.jp/portal/dt].

Zainteresowanie bibliotekami cyfrowymi w regionie Australii i Oceanii ma długą tradycję. Już od połowy lat 90. XX w. z inicjatywy bibliotek narodowych Australii i Nowej Zelandii uruchomiono wiele lokalnych projektów realizowanych przede wszystkim na uniwersytetach [Ianella 1996]. Inicjatywy te zorientowane były zarówno na ochronę dziedzictwa kulturowego, np. *The Australian Museums On Line* czy *Preserving Oral History Recordings Project*, jak i dotyczyły dystrybucji wydawnictw współczesnych np. *Electronic Journal Project*. W efekcie infrastruktura cyfrowa regionu prezentuje poziom zbliżony do europejskiego [Swanepoel 2008]. Wśród znaczących inicjatyw australijskich należy wyliczyć m.in.: *Project Gutenberg of Australia* [http://gutenberg.net.au/] realizowany od 2001 r.; *Picture Australia* uruchomiony w 1998 r. [http://www.pictureaustralia.org/] (1,8 mln obrazów), *Australian Periodical Publications 1840-1845* [http://www.nla.gov.au/ferg/] oraz ponad 90 innych zestawionych w specjalnym przewodniku *Australian Digitisation Project* [http://www.nla.gov.au/libraries/digitisation/projects.html]. Znaczące osiągnięcia ma także Nowa Zelandia, która w ramach narodowego projektu, który rozpoczęto w 1994 r. zbudowała wiele kolekcji, w szczególności portal narodowy *New Zealand Digital Library* [http://www.nzdl.org/], realizowany w University of Waikato. Warto dodać, że na potrzeby projektu zaprojektowano oprogramowanie Greenstone [http://www.greenstone.org/], rozwijane później we współpracy z UNESCO, które jest dystrybuowane bezpłatnie za licencji GNU/GPL. W oparciu o system Greenstone powstało kilkaset bibliotek cyfrowych na całym świecie, np. *Biblioteca Nacional de Uruguay* [http://coleccionesdigitales.bibna.gub.uy], *Armenian Rare Books* [http://greenstone.flib.sci.am/gsdl/cgi-

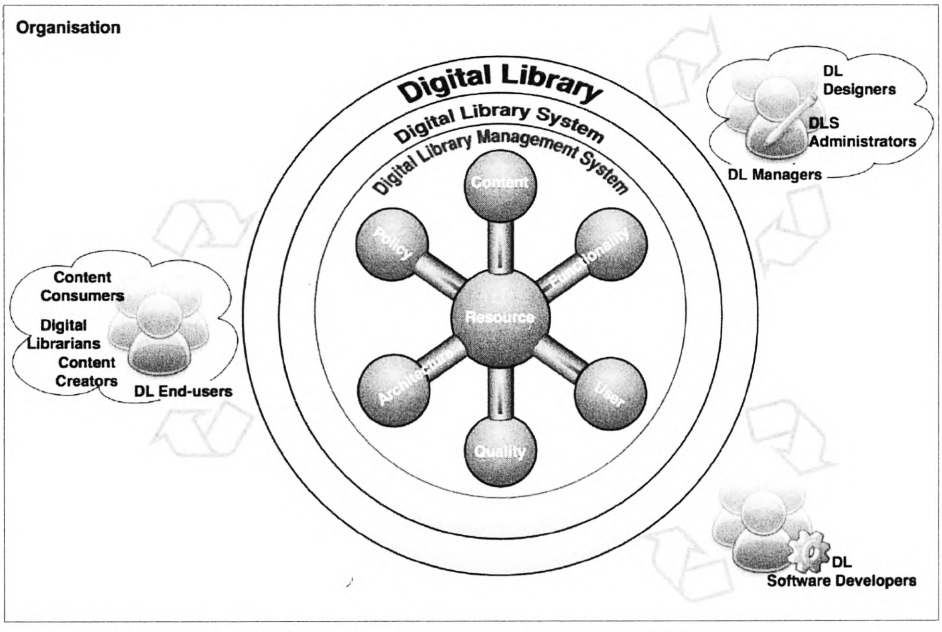
bin/library.cgi], czy *Oxford Digital Library* [http://www2.odl.ox.ac.uk/]. Wśród wielu innych przedsięwzięć prowadzonych w Nowej Zelandii warto też wspomnieć o kilku kolekcjach utrzymywany przez bibliotekę narodową, m.in. serwisach dawnych czasopism: *AtoJs* [http://atojs.natlib.govt.nz] i *Papers Past* [http://paperspast.natlib.govt.nz/].

Perspektywy

Rozwój prac w zakresie digitalizacji i budowy bibliotek cyfrowych, który rozpoczął się w latach 90. ubiegłego wieku i przyspieszył w pierwszej dekadzie bieżącego stulecia zaowocował znaczącym przyrostem projektów (por. rys. 1). Warto jednak zauważyć, że są one mocno zróżnicowane. Zdecydowana większość z nich to serwisy ukierunkowane na treści udostępniania (a nie funkcjonalność), zaś ich model organizacyjny opiera się na współpracy dostawców treści. Takie podejście było przez wiele lat dominujące i najczęściej uzasadniane jako strategia w celu zmniejszenia kosztów utrzymania biblioteki cyfrowej. Pierwszym krokiem ku integracji serwisów (spójność na poziomie danych) było upowszechnienie się protokółów OAI-PMH i budowa sieci rozproszonych. Wnikliwa analiza przekonuje jednak, że znacznie trudniej osiągnąć spójność na poziomie funkcjonalności ze względu na wielką różnorodność modeli i ontologii przyjętych przez istniejące systemy. Innymi słowy jest to negatywny skutek braku systematycznego podejścia do interoperacyjności przez projektantów. Barierę tę od niedawna próbują przełamać niektóre projekty najnowszej generacji, np. *DL.org* [http://www.dlorg.eu/] (społeczność w ramach DELOS), której prace torują drogę dla pełnej realizacji interoperacyjności w systemach, co jest warunkiem koniecznym dla realizacji spójnej infrastruktury globalnej [Candela; Castelli; Pagano 2011].

Dyskusja nad budową takiej infrastruktury, nazwanej często globalną biblioteką cyfrową [Nahotko 2010] toczy się w dyskursie naukowym już od kilku lat. Warto jednak zaznaczyć, że w odróżnieniu od poprzedniej dekady, gdy prym wiodły ośrodki amerykańskie, w XXI wieku najbardziej wartościowe prace w tym zakresie powstały w Europie (głównie w ramach DELOS). Szczególnie interesująca jest dyskusja dotycząca modelowania bibliotek cyfrowych w zmieniających się warunkach. Jeszcze na początku XXI wieku najpopularniejszy był model OASIS (2002), który dzielił aktorów występujących w otoczeniu biblioteki na producentów, konsumentów i administratorów [Dudczak 2007]; konkurowały z nim inne rozwiązania, m.in. model CKESS dedykowany do współpracy społecznościowej (2002), model Bessera dotyczący współpracy rozproszonej z wykorzystaniem agentów (2002) i model 5S (2005), nazwany od pierwszych liter elementów modelu – streams (strumienie), structures (struktury), spaces (przestrzenie), scenarios (scenariusze) i societies (społeczności) oraz inne [Nahotko 2010, s. 22-23]. Warto podkreślić, że żaden z nich nie operował na odpowiednio wysokim poziomie abstrakcji, co sprawiało, że w rozmaitych okolicznościach modele okazywały się ułomne. Zmianę przyniósł dopiero rok 2006, wraz z pierwszą wersją modelu referencyjnego opracowanego w DELOS. Początkowo głównym celem DELOS (jako grupy ekspertów) było podnoszenie poziomu wiedzy

w dziedzinie bibliotek cyfrowych poprzez koordynację prac europejskich zespołów badawczych. Jednym z pierwszych osiągnięć było nawiązanie współpracy z amerykańską National Science Foundation i powołanie wspólnych grup roboczych. Zespoły interesowały się m.in. techniką oraz kwestiami społecznymi i ekonomicznymi oraz – co najważniejsze – opublikowano zestaw zaleceń w odniesieniu do interoperacyjności, m.in. metadanych, własności intelektualnej i ekonomii, zasobów oraz informacji wielojęzycznych; zalecenia te opublikowano w specjalnym czasopiśmie *International Journal of Digital Libraries* [http://www.dljournal.org/] [Griffin; Peters; Thanos 2005]. Po przekształceniu DELOS z grupy w sieć (DELOS Netwok), jej główną misją stałą się integracja i koordynacja działań badawczych europejskich zespołów badawczych w zakresie bibliotek cyfrowych. Głównym zaś jej osiągnięciem stał się opracowany w 2006 r. model referencyjny biblioteki cyfrowej *Digital Library Reference Model* [DLRM], opisujący jej ramy koncepcyjne na bardzo wysokim poziomie abstrakcji [Candela; Castelli; Pagano 2007; Candela; Nardi 2011].



Rys. 7. Model referencyjny biblioteki cyfrowej [Candela; Nardi 2011].

Według DLRM biblioteka cyfrowa to – wirtualna organizacja, która kompleksowo gromadzi, zarządza i długotrwale zachowuje treści cyfrowe oraz oferuje docelowym społecznościom wyspecjalizowane funkcje, dbając o ich jakość i zgodność z przyjętymi kierunkami polityki [Candela; Nardi 2011]. W dokumencie przedstawiono specyfikację trzech warstw modelu biblioteki opisanego wcześniej w *Digital Libraries Manifesto* [Candela; Castelli 2006] oraz podjęto próbę zdefiniowania funkcjonalności na poziomie każdej z tych warstw. W DLRM bibliotekę cyfrową tworzą trzy odrębne warstwy (rys. 7): 1. Biblioteka cyfrowa (Digital Library); System biblioteki cyfrowej

(Digital Library System) i System zarządzania biblioteką cyfrową (Digital Library Management System). Rozróżnienie to jest istotne z poziomu administratora, zaś z punktu widzenia użytkownika warstwy te są tożsame. Należy zwrócić uwagę, że komponenty: content, policy, architecture, quality, user, functionality [odpowiednio: zawartość, polityka, architektura, jakość, funkcjonalność] znajdują się wewnątrz systemu i podlegają zarządzaniu. Zakres funkcjonalny został z kolei podzielony na pięć podstawowych grup, w ramach których wyróżniono konkretne funkcje: 1. Uzyskaj dostęp do informacji - przeglądaj, wizualizuj, zachowaj, przetłumacz, przeszukuj, dostarcz; 2. Zarządzaj informacją – stwórz, uaktualnij, przekształć, nanieś adnotacje, złóż, prześlij, waliduj; 3. Zarządzaj biblioteką cyfrową – rozpowszechnij, publikuj, wycofaj, przechowaj, zarządzaj zasadami, zarządzaj kolekcjami, zarządzaj użytkownikami (rejestruj użytkowników, zarządzaj rolami); 4. Personalizuj – wyraź zainteresowanie, zaloguj się, zapisz się, zastosuj profil, dostarcz informację zwrotną; 5. Współpracuj [Dudczak 2007]. Dalej idące wnioski można wyścignąć analizując zdefiniowane klasy aktorów (content consumers, digital librarians, content creator, end user, designers, administrators, managers, software developers [odpowiednio: konsumentów treści cyfrowych bibliotekarzy, twórcy treści, użytkowników końcowych, projektantów, administratorów, menedżerów, twórców oprogramowania]. Każda z tych grup definiuje wobec biblioteki cyfrowej własne oczekiwania i interakcje, np. użytkownik końcowy – język zapytań. Zaprezentowany model jest od strony technicznej zdecydowanie bardziej rozbudowany i pozwala na precyzyjne projektowanie biblioteki cyfrowej. DLRM jest także wyjściem naprzeciw heterogeniczności współczesnych repozytoriów, które z definicji są tak projektowane, aby móc długotrwale przechować dane cyfrowe w dowolnym formacie, opisane przy pomocy różnych schematów metadanych, a ich architektura coraz częściej zorientowana jest na usługi (SOA)[7]. Takie podejście do projektowania bibliotek umożliwia ich elastyczną integrację z innymi systemami [Dudczak 2007], czyli sprawia, że są zdolne do współpracy z dowolnymi systemami, które powstaną w przyszłości.

[7] SOA (ang. Service-Oriented Architecture) – architektura systemów informatycznych, w której podstawową zasadą jest wykorzystanie luźno powiązanych usług do realizacji logiki systemu. Poszczególne usługi udostępniane są jako niezależne od siebie, a ich użycie nie wymaga wiedzy o szczegółach ich realizacji [Dudczak 2007].

Bibliografia

1. Athanasopoulos, George; Ioannidis, Yannis; Fox, Edward (2010). *A functionality perspective on digital library interoperability.* „Lecture Notes in Computer Science" vol. 6273, s. 405-408 [dok. elektr.] Dostęp: doi: 10.1007/978-3-642-15464-5_43 [odczyt: 1.06.2011].
2. Ayris, Paul (2009). *The EU-funded Europeana Travel project.* „Liber Quarterly: The Journal of European Research Libraries" 19 (2), s. 63-102 [dok. elektr.] http://www.europeanatravel.eu/downloads/articleeTravel.pdf [odczyt: 2.06.2011].
3. Bakker, Trix (1999). *Ongoing digital library projects in the Netherlands (1995-2001).* „Bibliothek Forschung und Praxis" vol. 23 nr 3, s. 375-379 [dok. elektr.] 10.1515/bfup.1999.23.3.375 [odczyt: 1.06.2011].
4. Bush, Vannevar (1945). *As we may think.* „The Atlantic Monthly" nr 176, s. 101-108 [dok. elektr.] http://www.ps.uni-saarland.de/~duchier/pub/vbush/vbush-all.shtml [odczyt: 19.05.2011].
5. Candela, Leonardo; Castelli, Donatella; Ioannidis, Yannis; Koutrika, Georgia; Pagano, Pasquale; Ross, Seamus; Schek, Hans-Jörg; Schuldt, Heiko (2006). *Digital Libraries Manifesto.* [S.l.]: DELOS. Project no. 507618 [dok. elektr.] http://146.48.87.21/OLP/UI/1.0/Disseminate/1322686809I1OHFq6Tol/a221322686809g5YYBzA8 [odczyt: 25.05.2011].
6. Candela, Leonardo; Castelli, Donatella; Pagano, Pasquale (2007). *A reference architecture for digital library systems: principles and applications.* „Lecture Notes in Computer Science" vol. 4877, s. 22-35 [dok. elektr.] doi: 10.1007/978-3-540-77088-6_3 [odczyt: 21.05.2011].
7. Candela, Leonardo; Castelli, Donatella; Pagano Pasquale (2011). *History, evolution and impact of digital libraries.* [W:] Iglezakis, Ioannis; Synodinou, Tatiana-Eleni; Kapidakis, Sarantos ed. (2011). *E-Publishing and digital libraries: legal and organizational issues.* IGI Global, Hershey, PA: Information Science Reference, s. 1-30. [dok. elektr.] doi: 10.4018/978-1-60960-031-0.ch001 [odczyt: 30.11.2011].
8. Candela, Leonardo; Nardi, Alessandro red. (2011). *The Digital Library Reference Model* [ver. D 3.2b]. [dok. elektr.] http://www.dlorg.eu/index.php/outcomes/reference-model [odczyt: 25.05.2011].
9. Dudczak, Adam; Heliński, Marcin; Mazurek, Cezary; Parkoła, Tomasz; Werla, Marcin (2007). *Analiza funkcjonalności wybranych modeli i systemów zarządzania bibliotekami cyfrowymi.* „Zeszyty Naukowe Wydziału ETI Politechniki Gdańskiej" nr 5 Seria: Technologie Informacyjne, s. 1-10.
10. Englebart, Dauglas C. (1963). *Conceptual framework for the augmentation of man's intellect.* [W:] Howerton, Paul W.; Weeks, David C. red. (1963). *Vistas in information handling.* Washington, D.C: Spartan Books, s. 1-20.
11. Fox, Edward A. red. (1993). *Source book on digital libraries.* Blacksburg: Virginia Tech. [dok. elektr.] http://fox.cs.vt.edu/DLSB.html [odczyt: 20.05.2011].
12. Fox, Edward A. (1999). *The digital libraries initiative: update and discussion.* „Bulletin of the American Society of Information Science" vol. 26 nr 1, s. 7-11.
13. Fox, Edward A. (2001). *Digital libraries: topical outline.* Blacksburg: Virginia Tech. Department of Computer Science. [dok. elektr.] http://ei.cs.vt.edu/~dlib/tut/FoxAll.pdf [odczyt: 21.05.2011].

14. Franke, Jerzy (2005). *Warianty rozwoju cyfrowych kolekcji.* [W:] *Bibliologia, biblioteki, bibliotekarze* (2005). Warszawa: Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich, s. 73-89.
15. Franke, Jerzy (2007). *Googletheca Universalis?* [W:] *Biblioteki cyfrowe: projekty, realizacje, technologie* (2007). Warszawa: Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich, s. 121-166. Także: http://bbc.uw.edu.pl/publication/4 [odczyt: 21.05.2011].
16. Garrison, Adriene (2010). *OAISTER: the roots and the resource.* „PNLA Quarterly" 74 (2), s. 46-51 [dok. elektr.] http://www.pnla.org/quarterly/Winter2010/PNLAQ74-2winter2010.pdf [odczyt: 1.06.2011].
17. Greenstein, Daniel; Thorin, Suzanne E. (2002). *The digital library: a biography.* Washington: Digital Library Federation. [dok. elektr.] www.clir.org/pubs/reports/pub109/pub109.pdf [odczyt: 21.05.2011].
18. Griffin, Stephen; Peters, Carol; Thanos, Costantino (2005). *Toward the new-generation digital libraries: recommendations of the NSF/EU-DELOS working groups.* „International Journal on Digital Libraries" 5 (4), s. 253-254. [dok. elektr.] doi: 10.1007/s00799-004-0093-9 [odczyt: 2.06.2011].
19. Hauben, Jay (2004). *Vannevar Bush and JCR Licklider: Libraries of the Future 1945-1965,.* Conference Paper – Conference in the Institute for Library Science of Humboldt University in Berlin [dok. elektr.] http://*www.ais.org/~jrh/acn/acn15-2.articles/jhauben.pdf* [odczyt: 21.05.2011].
20. Hiom, Debra (2006). *Retrospective on the RDN.* „Ariadne" nr 47 (April) [dok. elektr.] http://www.ariadne.ac.uk/issue47/hiom/ [odczyt: 21.05.2011].
21. Hunter, Philip (2005). *OAI and OAI–PMH for absolute beginners: a non-technical introduction.* Geneva: CERN. [dok. elektr.] http://hdl.handle.net/10760/7143 [2011.05.21].
22. Ianella, Renato (1996). *Australian Digital Library Initiatives.* „D-Lib Magazine" vol. 2 nr 12. [dok. elektr.] http://www.dlib.org/dlib/december96/12iannella.html [odczyt: 1.06.2011].
23. Kolasa, Władysław M. (2007). *dLibra Digital Library Framework – platforma do budowy bibliotek cyfrowych.* [W:] *Biblioteki cyfrowe: projekty, realizacje, technologie.* Warszawa: Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich, s. 67-88. http://bbc.uw.edu.pl/publication/4 [odczyt: 21.05.2011].
24. Kolasa, Władysław M. (2009). *Biblioteki cyfrowe: wstęp do problematyki.* [W:] *Nowe technologie w bibliotekach publicznych.* Warszawa: Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich, s. 183-196.
25. Kolasa, Władysław M. (2010). *Największe światowe biblioteki cyfrowe i repozytoria OA* [dok. elektr.] http://mbc.malopolska.pl/dlibra/text?id=dlib_world [odczyt: 21.05.2011].
26. Lebert, Marie (2008). *Project Gutenberg (1971-2008)* [dok. elektr.] http://www.gutenberg.org/cache/epub/27045/pg27045.html [odczyt: 20.05.2011].
27. Lewandowska, Agnieszka; Werla, Marcin (2009). *OAI-ORE: Interoperacyjność bibliotek cyfrowych.* „PIONIER Magazine", nr 3 (04), s. 22-23 [dok. elektr.] http://www.pionier.net.pl/magazine/pl/artykuly/1042/OAI-ORE.html [odczyt: 23.05.2011].
28. Leye, Yao; Ping, Zhao (2009). *Digital libraries in China: progress and prospects.* „Electronic Library" 27(2), s. 308-318 [dok. elektr.] doi: 10.1108/02640470910947656 [odczyt: 1.06.2011].
29. Madalli, Devika P. (2003). *Digital Libraries and Digital Library Initiatives.* [W:] *Digital Libraries: Theory and Practice.* Bangalore: Documentation Research & Training Centre, s. 1-8.

30. The Mercury Electronic Library and Library Information System II (1992). *Mercury Technical Reports Series* nr 6 [dok. elektr.] http://www.cs.cornell.edu/wya/papers/Mercury6.doc [odczyt: 21.05.2011].
31. McCray, Alexa T.; Gallagher, Marie E. (2001). *Principles for digital library development*. „Communications of the ACM" vol. 44 nr 5, s. 49-53 [dok. elektr.] doi: 10.1145/374308.374339 [odczyt: 1.06.2011].
32. O'Leary, Mick (2009). *Open Content Alliance Embodies Open Source Movement*. „Information Today" 26 (1), s. 37-43.
33. Nahotko, Marek (2007). *Naukowe czasopisma elektroniczne*. Warszawa: Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich, 180 s.
34. Nahotko, Marek (2010). *Komunikacja naukowa w środowisku cyfrowym : globalna biblioteka cyfrowa w informatycznej infrastrukturze nauki*. Warszawa: Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich, 283 s.
35. Raitt, David (2000). *Digital initiatives across Europe*. „Computers Libraries" vol. 20 nr 11/12, s. 26-34 [dok. elektr.] http://www.infotoday.com/cilmag/nov00/raitt.htm [odczyt: 25.05.2011].
36. Sauber, Peter (2009). *Timeline of the Open Access Movement* [dok. elektr.] http://www.earlham.edu/~peters/fos/timeline.htm [odczyt: 13.03.2011].
37. Swanepoel, Marinus (2008). *Digitzation initiatives: A reconnaisance of the global landscape*. [Lethbridge]: University of Lethbridge. [dok. elektr.] http://uleth.academia.edu/MarinusSwanepoel/Papers [odczyt: 20.05.2011].
38. Tedd, Lucy A.; Large, Andrew (2005). *Digital libraries: principles and practice in a global environment*. München: K.G. Saur, XIX+280 s.
39. Toobin, Jeffrey (2007). *Google's moon shot: the quest for the universal library*. "The New Yorker" (5 February).
40. Urbaniec, Marta (2006). *Open Content Alliance – cyfrowa biblioteka dla wszystkich*. „EBIB" nr 3 (73) [dok. elektr.] http://www.ebib.info/2006/73/urbaniec.php [odczyt: 21.05.2011].
41. Wells, Herbert George (1938). *World Brain*. Garden City, New York: Doubleday, Doran & Co. Także: https://sherlock.ischool.berkeley.edu/wells/world_brain.html [odczyt: 19.05.2011].
42. Zhen, Xihui (2010). *Overview of digital library development in China*. „D-Lib Magazine" vol. 16 nr 5/6 [dok. elektr.] doi:10.1045/may2010-zhen [odczyt: 23.05.2011].

Biblioteki cyfrowe w świetle międzynarodowych konferencji 1993-2011

Monika Krakowska

Wstęp

Wymiarem praktyczno-teoretycznym rozwoju badań oraz rozważań teoretycznych w obszarze dynamicznie rozwijających się bibliotek cyfrowych stały się różnorodne i liczne, regularne i czasowe konferencje, warsztaty i spotkania naukowe, odbywające się na całym świecie. Spotkania te, organizowane wśród m.in. decydentów, praktyków, badaczy, bibliotekarzy, a także twórców, stanowią konieczny stymulant ich rozumienia, implementacji, dobrych praktyk, a także analizy wszystkich elementów tworzących ten specyficzny system informacyjno-wyszukiwawczy. Wielokierunkowy rozwój bibliotek cyfrowych, zainicjowanie rezultatów rudymentarnych analiz oraz doświadczeń w procesie indywidualnego i instytucjonalnego projektowania bibliotek cyfrowych, wykorzystywanych w wirtualnym procesie edukacyjnym lub organizacjach cyfrowych, stanowi obszar badań prowadzonych globalnie. Wzorcowe modele bibliotek cyfrowych, wymiana praktycznych inklinacji w zakresie formowania tych systemów informacyjnych pozwolą na zharmonizowane, jednolite i systematyczne udoskonalanie obiektów, a także podejmowanie nowych wyzwań w procesie digitalizacji zasobów oraz realizowania polityki cyfryzacji dziedzictwa kulturowego. W kontekście wdrażania repozytoriów cyfrowych oraz kolejnych etapów w rozwoju bibliotek cyfrowych, w artykule zostaną przedstawione najważniejsze konferencje związane z problematyką zasobów cyfrowych.

Pierwsze konferencje – lata 90.

Pierwsze konferencje związane z wdrażaniem bibliotek cyfrowych związane są ściśle z inicjatywami dotyczącymi rozwoju technologii i nowych narzędzi informacyjno-komunikacyjnych. Rozwój i eksperymenty dotyczące implementacji i wzrostu repozytoriów cyfrowych wiąże się przede wszystkim z ekspansją technologiczną, nierównomierną, na całym świecie. Lata 90. XX w. stanowią okres realizacji pierwszych projektów organizujących biblioteki cyfrowe, co zostało scharakteryzowane w artykule o historii bibliotek cyfrowych. W czasie ekspansji pierwszego etapu rozwoju repozytoriów cyfrowych konferencje dedykowane bibliotekom cyfrowym, zawierają zagadnienia odwołujące się do tego obszaru, pojawiają się także incydentalnie podczas licznych konferencji poświęconych zarządzaniu wiedzą, rozwoju nowych form i narzędzi komunikacji i informacji, ekspansji Interentu, hipertekstów i hipermediów, problematyki procesu wyszukiwania informacji, katalogów OPAC, interakcji człowieka z komputerem [Fox, dok.elektr.].

Jedną z pierwszych konferencji, na których podejmowano zagadnienia związane z rozwojem oraz implementacją bibliotek cyfrowych, stanowiących systemy informacyjno-wyszukiwawcze w perspektywie zastosowanych technologii była Międzynarodowa Konferencja Zarządzania Informacją i Wiedzą – *Conference on Information and Knowledge Management* (CIKM). Po raz pierwszy została zorganizowana w 1993 r. i odbyła się w Baltimore, w Stanach Zjednoczonych[1]. Wśród pierwszych warsztatów znalazł się *National Science Foundation Workshop on Visual Information Management Systems* (Warsztaty Narodowej Fundacji Nauki Systemów Zarządzania Informacją Wizualną) z 1992 r., które zorganizowano w Redwood, w Kalifornii. Dotyczyły one systemów organizacji zasobów wizualnych oraz wdrażania projektów akademickich związanych z zasobami sieciowymi, opartych na tych systemach. Kolejne warsztaty promujące biblioteki cyfrowe odbyły się w 1993 r. w Pittsburghu, w amerykańskim stanie Pensylwania. Information *Access and the Networks – A Research Workshop* (Warsztaty naukowe Dostępu do Informacji i Sieci), związane były z wykorzystywaniem nowych technologii do tworzenia cyfrowych repozytoriów wiedzy i archiwów sieciowych. Wydarzenia związane z rozwojem bibliotek cyfrowych związane były z realizowanymi projektami, w tym finansowanymi przez Unię Europejską, pierwszymi inicjatywami mającymi na celu organizację systemów informacyjno-wyszukiwawczych w cyberprzestrzeni. Jednym z takich wydarzeń były warsztaty organizowane w Amsterdamie (Holandia), w ramach programu IMPACT zatytułowane *Workshop on Hypermedia and Hypertext Standards* (Warsztaty Standardów Hipermediów i Hipertekstów). W 1993 r., ponownie w Pittsburgu, podczas *16th International Conference on Research and Development in Information Retrieval* (16. Międzynarodowej Konferencji Badań i Ewolucji Wyszukiwania Informacji) pojawiły się zagadnienia związane z procesami wyszukiwania informacji, multimediami i nowymi systemami informacyjno-wyszukiwawczymi.

[1] Konferencja odbywa się zawsze na terenie Stanów Zjednoczonych, w różnych miastach. W 2012 r. odbędzie się w Maui na Hawajach. http://fox.cs.vt.edu/DLSB.html

W Wielkiej Brytanii organizowano od 1990 r. serię konferencji ELVIRA: *International Conference on Electronic Library & Visual Information Research* (Międzynarodowa Konferencja Biblioteki Elektroniczne i Wyszukiwanie Informacji Wizualnej), które poruszały istotne kwestie związane z bibliotekami cyfrowymi, strukturą systemów informacyjnych, procesów wyszukiwania i przechowywania informacji wizualnej [ARIADNE Links Index, dok. elektr.]. W 1995 r., w japońskiej miejscowości Ibaraki, odbyło się międzynarodowe sympozjum ISDL'95: *International Symposium on Digital Libraries* (Międzynarodowe Sympozjum Bibliotek Cyfrowych), gdzie prezentowano kwestie poświęcone procesom digitalizacji, tworzenia, przechowywania, wyszukiwania zasobów multimedialnych, definiowania i rozumienia bibliotek cyfrowych, wizualizacji informacji, interfejsów i użytkowników [http://www.dl.slis.tsukuba.ac.jp/ISDL95/CFPE.html].

Przykładem różnorodnych, cyklicznych seminariów organizowanych przez instytucje digitalizujące zasoby, jest *Seminar on Cataloging Digital Documents* (Seminarium Katalogowania Dokumentów Cyfrowych) z 1994 r. przygotowaną przez Bibliotekę Uniwersytetu Wirginii oraz Bibliotekę Kongresu (*University of Virginia Library, Charlottesville, Library of Congress*).

W 1994 r. w Genewie, w Szwajcarii, po raz pierwszy zorganizowano międzynarodową konferencję *International Conference on the World-Wide Web, IWWW (Międzynarodowa Konferencja na temat Światowej Sieci Komputerowej). Przedsięwzięcie kontynuowane jest do dzisiaj i odbywa się w wybranych miastach na wszystkich kontynentach. 20. edycja konferencji odbyła się w 2011 r., w Indiach, w miejscowości Hyderabad, pod auspicjami International Institute of Information Technology Hyderabad (Międzynarodowego Instytutu Technologii Informacyjnych w Hyderabad), International Institute of Information Technology Bangalore (Międzynarodowego Instytutu Technologii Informacyjnych w Bangalore) oraz Indian Institute of Technology Bombay (Indyjskiego Instytutu Technologicznego w Bombaju). Zagadnienia prezentowane w licznych wystąpieniach odnoszą się także do tematyki zastosowania nowych technologii informacyjnych w opracowywaniu repozytoriów cyfrowych, a także wymiaru społeczno-kulturowego przez tworzenie i rozwijanie bibliotek cyfrowych* [Gross, dok. elektr.; IWWW Conferences Committee, dok. elektr.]. W 1996 r. Szwajcaria była także organizatorem warsztatów *Special Interest Group on Information Retrieval* (Specjalnej Grupy Zainteresowań Wyszukiwania Informacji): *SIGIR-96 Workshop on Networked Information Retrieval* podczas *19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (19. Corocznej Międzynarodowej Konferencji ACM SIGIR na rzecz Badań i Rozwoju Wyszukiwania Informacji). Problematyka warsztatów dotyczyła przede wszystkim bibliotek cyfrowych, interfejsów tego typu systemów informacyjnych, relewancji zasobów sieciowych i tworzenia repozytoriów cyfrowych [CIIR, dok. elektr.]

Istotnym z punktu widzenia charakterystyki bibliotek cyfrowych w paradygmacie kognitywnym i zachowań użytkowników, procesów informacyjnych oraz wyszukiwania informacji w repozytoriach cyfrowych było forum *Artificial Intelligence in Digital Libraries (Sztuczna Inteligencja w Bibliotekach Cyfrowych) z 1997 r., które stanowiło*

część International Joint Conference on Artificial Intelligence Workshop Series (Serii Warsztatów Międzynarodowej Wspólnej Konferencji Sztucznej Inteligencji), którą zorganizowano w Nagoyi, w Japonii. Dwa lata wcześniej w 1995 r., również w Japonii, odbyło się międzynarodowe sympozjum ISDL'95: International Symposium on Digital Libraries (Międzynarodowe Sympozjum Bibliotek Cyfrowych), którego tematem były kluczowe zagadnienia w zakresie tworzenia repozytoriów cyfrowych, gromadzenia zasobów cyfrowych, ich opracowywania, przechowywania i wykorzystywania, jak również wizualizacji informacji i interfejsów użytkowników [Takehisa, University of Tsukuba, 1995, dok.elektr.].

W 1997 r. na uwagę zasługuje także europejska konferencja o charakterze międzynarodowym, która została zorganizowana przez UKOLN, *The United Kingdom Office for Library and Information Networking* (Brytyjski Urząd Sieci Bibliotecznych i Informacyjnych). Do zagadnień poruszanych podczas tego forum znalazły się kwestie cyfrowych środowisk informacyjnych, cyfrowych zasobów, przechowywania repozytoriów cyfrowych [UKOLN, dok. elektr.].

Lista różnorodnych wydarzeń z lat 90. ubiegłego wieku, znajduje się na stronie europejskiego projektu Unii Europejskiej Ariadne [Ariadne, dok. elektr.] oraz na blogu Bena Grossa, inżyniera systemowego zajmującego się m.in. badaniami nad systemami informacyjnymi i bibliotekami cyfrowymi [Ben Gross, dok. elektr.].

W międzyczasie rozpoczęto organizację pierwszych cyklicznych konferencji europejskich, amerykańskich, kanadyjskich, azjatyckich, które zostaną scharakteryzowane w dalszej części tekstu.

Najważniejsze konferencje europejskie

Zagadnienia bibliotek cyfrowych pojawiają się obecnie w obszarze tematów, które podejmowane są podczas rozlicznych konferencji odbywających się w Europie. Na uwagę zasługuje m.in. chorwacka konferencja *Libraries in the Digital Age, LIDA (Biblioteki w erze cyfrowej), organizowana cyklicznie od 2000 r. Od jedenastu lat poruszane są kwestie praktyczne i teoretyczne, związane z technologicznym, społecznym i kultowym wymiarem zastosowania nowych technologii informacyjno-komunikacyjnych do tworzenia m.in. repozytoriów cyfrowych, automatyzacji bibliotek, wyszukiwaniem informacji elektronicznej, edukacji i szkoleń z zastosowaniem bibliotek cyfrowych. Konferencja ma charakter pan-europejski [LIDA, dok. elektr.]. Podobnym sympozjum jest International Conference on Conceptions of Library and Information Science*, CoLIS (Międzynarodowa Konferencja Koncepcje w Bibliotekoznawstwie i Informacji Naukowej). Konferencja ta jest organizowana od 1991 r. (pierwsza edycja miała miejsce w Tampere, w Finlandii) i do tej pory odbyło się już siedem edycji, organizowanych głównie co trzy lata – w Kopenhadze, Danii (1996 r.), Dubrowniku, w Chorwacji (1999 r.)[2], Seattle, w Stanach Zjednoczonych (2002 r.), Glasgow, Wielkiej Brytanii (2005 r.),

[2] Podczas konferencji CoLIS w 1999 r. artykuł dotyczący aspektu bibliotek cyfrowych był prezentowany przez prof. UJ, dr hab. Wande Pindlową oraz prof. UJ, dr hab. Marię Kocójową, co zostało już wymienione w niniejszej publikacji w artykule dr Małgorzaty Janiak.

Boras, w Szwecji (2007 r.), Londynie, W Wielkiej Brytanii (2010 r.). W 2013 r. CoLIS będzie odbywał się ponownie w Kopenhadze, w Danii [Royal School of LIS, CoLIS8, dok. elektr.]. Zagadnienia dotyczącej konsytuacyjności w informacji naukowej i bibliotekoznawstwie, prezentowane podczas kolejnych edycji tego sympozjum odnoszą się także do problematyki funkcjonowania bibliotek cyfrowych, w wielu odmiennych wymiarach, perspektywach. Ogólnym celem jest zapewnienie szerokiemu gremium praktyków i teoretyków, prezentacji wyników badań historycznych, teoretycznych, empirycznych i technicznych informacji naukowej i bibliotekoznawstwa, odnoszących się do zrozumienia i wykorzystywania informacji, jak również promowanie interdyscyplinarnego podejścia do analizowania różnorodnych zagadnień. Podczas konferencji w Dubrowniku tematem przewodnim były sprawy związane właśnie z bibliotekami cyfrowymi [Vrana, University of Zagreb, 2002, dok. elektr.].

Wśród spotkań organizowanych w Europie, mających także charakter międzynarodowych wydarzeń skupiających wszystkich użytkowników – administratorów, bibliotekarzy, naukowców, decydentów, jak również odbiorców końcowych na uwagę zasługują, oprócz Europejskiej Konferencji Bibliotek Cyfrowych, liczne warsztaty, sympozja i spotkania *Network of Excellance for Digital Libraries DELOS (Sieci Doskonałości Bibliotek Cyfrowych DELOS). Konsorcjum to, częściowo finansowane przez Komisję Europejską, realizowało komplementarne prace w latach 2004-2007, zaś wyniki i rezultaty działań Sieci stosowane są do dnia dzisiejszego przez różnorodne ośrodki akademickie, instytucje kultury i inne jednostki implementujące biblioteki cyfrowe. Organizowane przez DELOS warsztaty i szkolenia realizowane były między innymi podczas wielu europejskich konferencji, w tym ECDL, Europejskiej Konferencja na rzecz Badań i Rozwiniętej Technologii Bibliotek Cyfrowych. Organizacja jest również odpowiedzialna za wdrożenie Modelu Referencyjnego Bibliotek Cyfrowych (Reference Model for Digital Libraries) [Candela i in. 2007, dok. elektr.] charakteryzowanego już w niniejszej publikacji. Zorganizowane zostały warsztaty, dotyczące zagadnień modelu referencyjnego tych specyficznych systemów informacyjno-wyszukiwawczych, które odbywały się między innymi:*

- 1-2 czerwca 2006, *Reference Model Workshop* (Warsztaty Modelu Referencyjnego), Villa Grazioli, Frascati – Rzym, Włochy;
- 23 czerwca 2007, *First International Workshop on „Digital Libraries Foundations"* (Pierwsze Międzynarodowe Warsztaty Podstaw Bibliotek Cyfrowych), Vancouver, British Columbia, w Kanadzie;
- 20 września 2007, *Second Workshop on Foundations of Digital Libraries* (Drugie Warsztaty Podstaw Bibliotek Cyfrowych), Budapeszt, Węgry;
- 18 września 2008, *Third Workshop on Foundations of Digital Libraries (Trzecie* Warsztaty Podstaw Bibliotek Cyfrowych), Aarhus, Dania [DELOS a, dok. elektr.].

Sieć DELOS organizowała także dwie konferencje w 2007 r., które odbyły się dwukrotnie: w lutym, w Grecji oraz w grudniu, we Włoszech. Podczas obu forów poruszone zostały kwestie związane z perspektywą badań Sieci DELOS, którą zaprezentowano w poniższej tabeli. W szczególności zajmowano się kwestiami 1) modelu

referencyjnego biblioteki cyfrowej, 2) wyszukiwaniu zasobów w bibliotekach cyfrowych, wyszukiwaniu zasobów wizualnych, 3) ochronie danych, 4) personalizacji, 5) ewaluacji bibliotek cyfrowych, 6) zarządzania zasobami cyfrowymi, obiektów 3D, 7) ontologiami, wielojęzyczności 8) architekturą systemu [DELOS b, dok. elektr.]. W Tabeli 1 znajdują się obszary badawcze analizowane w Sieci DELOS, których omówienie i prezentacja najważniejszych rezultatów i przykładów podejmowano na konferencjach i warsztatach reprezentatywnego projektu.

Tabela 1. Obszary badań Sieci DELOS prezentowane podczas dwóch edycji Konferencji DELOS Biblioteki Cyfrowe w 2007 r. [opracowanie za: DELOS b, dok. elektr.].

DELOS Obszary badawcze	
Badania dotyczące podstaw tworzenia bibliotek cyfrowych	***Badania w perspektywie użytkowników***
• Model referencyjny dla systemów bibliotek cyfrowych	• Interfejsy użytkowników bibliotek cyfrowych
Badania w perspektywie technologicznej	• Wizualizacja bibliotek cyfrowych
• Architektura systemów bibliotek cyfrowych	• Personalizacja w biblistach cyfrowych
• Dostęp do informacji w bibliotekach cyfrowych	***Problematyka horyzontalna***
• Audio/Wizualne biblioteki cyfrowe	• Przechowywanie i odzyskiwanie zasobów cyfrowych
• Semantyczna interoperacyjność w bibliotekach cyfrowych	• Repozytoria cyfrowe
	• Metody ewaluacji bibliotek cyfrowych
	• Infrastruktura ewaluacji bibliotek cyfrowych

Wśród najważniejszych wydarzeń – forów, sympozjów i konferencji na szczeblu krajowym i międzynarodowym, znajdują się organizowane w latach 2004-2007, m.in.: 2004 r. – Krajowe forum DELOS: *Verso la Biblioteca Digitale Italiana della Scienza e della Tecnica* (W kierunku Włoskiej Cyfrowej Biblioteki Naukowej), Rzym, Włochy; warsztaty DELOS: Semantic Interoperability in Digital Libraries (Semantyczna interoperacyjność w bibliotekach cyfrowych), Bath, Wielka Brytania; warsztaty DELOS: The Evaluation of Digital Libraries (Ewaluacja bibliotek cyfrowych), Piza, Włochy; krajowe przedsięwzięcie DELOS: *1st Italian Research Conference on Digital Library Management Systems* (Pierwsza Włoska Konferencja Naukowa Systemów Zarządzania Biblioteką Cyfrową), Padwa, Włochy; 2006 r. - *Reference Model Workshop*, warsztaty modelu referencyjnego BC, Rzym, Włochy; *The Open DELOS Seminar on Digital Libraries and Digital Preservation* (Otwarte Seminarium DELOS Bibliotek Cyfrowych i Przechowywania Zasobów Cyfrowych), Tallin, Estonia; 2007 r. - *DELOS International Summer School on Digital Preservation in Digital Libraries* (Międzynarodowa Letnia Szkoła DELOS na temat Przechowywania Zasobów Cyfrowych w BC), Pisa, Włochy; *1st Workshop on Digital Libraries Foundations* (Pierwszy Warsztat Bibliotek Cyfrowych), Vancouver, Kanada. Partnerzy Sieci DELOS współorganizowali również warsztaty, sesje i panele podczas międzynarodowych konferencji poświęconych holistycznie pojmowanej bibliotece cyfrowej, forach dedykowanych paradygmatom technologicznym, kognitywnym, systemowym, społecznym [DELOS a, dok. elektr.].

W trakcie realizacji projektu DELOS wśród organizowanych licznych warsztatów, wymienić należy *International Workshop of the EU Network of Excellence DELOS on Audio-Visual Content And Information Visualization In Digital Libraries* (Międzynarodowe Warsztaty Sieci Doskonałości Unii Europejskiej DELOS Zasoby Audio-wizualne w Bibliotekach Cyfrowych), podczas których prezentowano przede wszystkim zagadnienia dotyczące tworzenia, przechowywania, zarządzania kolekcjami multimedialnymi, wyszukiwania informacji wizualnej, tworzenia interfejsów w bibliotekach cyfrowych [DELOS AVIVDiLib, dok. elektr.].

Istotnym wydarzeniem realizowanym także w ramach działalności Sieci DELOS były spotkania *Workshop on Very Large Digital Libraries* (Warsztaty Bardzo Duzych Bibliotek Cyfrowych). Zostały one włączone w zakres 12. Europejskiej Konferencji Bibliotek Cyfrowych (ECDL) w 2008 r. Dotychczas odbyły się trzy spotkania Warsztatów – w Aarhaus, w Danii (2008 r.), w Corfu, w Grecji (2009 r.) oraz w Glasgow w 2010, wspólnie z 14. edycją ECDL [VLDL 2010, dok. elektr.]. Tematyka dotyczyła m.in.: architektury usług / infrastruktury dla dużych systemów bibliotek cyfrowych (BC), doświadczenia w budowie i eksploatacji dużych systemów bibliotek cyfrowych, integracji i federacji istniejących systemów BC, budowy i zarządzania rozproszonymi środowiskami informacyjnymi, metodologii zrównoważonego rozwoju BC na dużą skalę, zarządzania skalowalnością w BC na dużą skalę, interoperacyjności w BC na dużą skalę, jakości usług w BC na dużą skalę, zarządzania danymi w DL na dużą skalę, modeli organizacyjnych dla dużych bibliotek cyfrowych [VLDL 2010, dok. elektr.].

Najważniejszą i jedną z najstarszych konferencji europejskich jest *The European Conference on Research and Advanced Technology on Digital Libraries ECDL. Europejska Konferencja na rzecz Badań i Rozwiniętej Technologii Bibliotek Cyfrowych organizowana jest od czternastu lat. Pierwsza edycja odbyła się w 1997 r., we włoskiej Pizie, częściowo sfinansowana przez Unię Europejską i została przygotowana w związku z działaniami podejmowanymi przez Grupę Roboczą Sieci Doskonałości DELOS (Network of Excellence on Digital Libraries)*.Stanowi najważniejszą obecnie konferencję pan-europejską, dotyczącą technologicznego, społecznego oraz praktycznego wymiaru bibliotek cyfrowych. Kolejne cykliczne spotkania odbyły się w Heraklionie na Krecie (1998 r, Grecja.), Paryżu (1999 r., Francja), Lizbonie (2000 r., Portugalia), Darmstadcie (2001 r., Niemcy), Rzymie (2002 r., Włochy), Trondheim (2003 r., Norwegii), Bath (2004 r., Wielka Brytania), Wiedniu (2005 r., Austria), Alicante (2006 r., Hiszpania), Budapeszcie (2007 r., Węgry), Aarhus (2008 r., Danii), Korfu (2009 r., Grecja), Glasgow (2010 r., Szkocja, Wielka Brytania).

W 2011 r. nazwa konferencji została zmieniona, przez co piętnastą edycję zrealizowaną w Berlinie (Niemcy) przekształcono w *International Conference on Theory and Practice of Digital Libraries* (TPDL), Międzynarodową Konferencję Teoria i Praktyka Bibliotek Cyfrowych [TPDL2011, dok. elektr.].

W trakcie pierwszej konferencji ECDL podjęto się charakterystyki zagadnień związanych z: 1) wspierania interakcji z użytkownikiem, 2) metadanych, 3) wyszukiwania informacji, 4) architektury systemów informacyjno-wyszukiwawczych, 5) wielojęzycznego wyszukiwania informacji, 6) repozytoriów strukturalnych, 7) mo-

delowania danych strukturalnych i przeszukiwania repozytoriów (uwzględniając OCR – *optical character recognition*) oraz 8) prezentacji przykładów implementacji bibliotek cyfrowych (*case study*) [Peters, Thanos 2007, dok.elektr.]. Kolejna edycja, zorganizowana w Heraklionie, w 1998 r. została poświęcona m.in.: 1) architekturom i strukturom bibliotek cyfrowych, 2) wizualnym bibliotekom cyfrowym (szczególnie repozytoriom zasobów wideo), 3) wielojęzyczności, 4) technologii specjalistycznej dla bibliotek cyfrowych, 5) nawigacji, 6) interakcjom człowiek-komputer, 7) semantycznemu wyszukiwaniu informacji, 8) wyszukiwaniu w języku naturalnym, 9) włączeniu bibliotek elektronicznych do programów edukacyjnych, a także przedstawieniu dobrych praktyk, organizacji paneli oraz warsztatów projektu DELOS [Nicolaou, Stephanidis 1998, dok. elektr.]. Konferencja w 1999 r. została poszerzona, oprócz zagadnień wyszukiwania informacji, prezentacji wdrożeń, o problematykę: 1) kategoryzacji i dostępowi do obiektów wizualnych, 2) obiektów audio i wideo w repozytoriach cyfrowych, 3) wymianie wiedzy, 4) perspektywie użytkowników bibliotek cyfrowych, 5) wyszukiwaniu w wielu językach (*cross-language information retrieval*), 6) modelowania, dostępności i powiązań między obiektami [Abiteboul, Vercoustre 1999, dok. elektr.].

W 2000 r. podczas konferencji pojawiły sie nowe zagadnienia dotyczące: 1) rozpoznawania optycznego, 2) struktur (*frameworks*) bibliotek cyfrowych, 3) multimediów. Ponadto ponownie zajmowano się obszarem wyszukiwania informacji, użytkowników i meta danych [Borbinha, Baker 2000, dok. elektr.]. W latach 2001-2011 na konferencji ECDL poruszano wielorakie zagadnienia dotyczące holistycznie pojmowanej biblioteki cyfrowej, a zwłaszcza związanej z technologią wykorzystywana podczas tworzenia repozytoriów sieciowych, zasobów cyfrowych, typologii bibliotek cyfrowych, ich rodzajów, ontologii stosowanych podczas indeksowania i klasyfikowania zasobów, a także kwestii dotyczących użytkowników i interakcji człowieka z komputerem. Szczegółowy wykaz najistotniejszych tematów podejmowanych podczas konferencji w latach 2001-2011 znajduje się w tabeli 2.

Tabela 2. Problematyka bibliotek cyfrowych poruszana podczas konferencji ECDL za lata 2001-2011 [oprac. autorki].

OBSZAR	NAJWAŻNIEJSZE ZAGADNIENIA poruszane podczas konferencji ECDL w latach 2001-2011
TECHNOLOGIA	aplikacje dla BC
	architektura systemu, infrastruktura
	modele gridowe, peer-to-peer
	interfejsy
	interoperacyjność
	modelowanie BC
	XML
ZASOBY, TYPY BC	archiwa i repozytoria
	dziedzinowe, specjalne biblioteki cyfrowe

	Europeana
	multimedia
	tekstowe biblioteki cyfrowe
	zasoby audiowizualne
ONTOLOGIE	ontologie
	semantyka
	społeczne tagowanie
	technologie języków informacyjno-wyszukiwawczych
	wielojęzyczność
PROCESY, FUNKCJONOWANIE BC	analiza cytowań
	aspekt prawny
	cytowanie i adnotacje
	digitalizacja, interpretacja
	ewaluacja systemu
	klasyfikowanie obiektów cyfrowych
	linkowanie obiektów cyfrowych
	metadane, aplikacje metadanych
	przechowywanie zasobów cyfrowych
	rekomendowanie zasobów cyfrowych
	tworzenie bibliotek cyfrowych
	tworzenie kolekcji cyfrowych
	usługi BC
	wyszukiwanie w BC
	zarządzanie informacją personalną
	zarządzanie kolekcją cyfrową
UŻYTKOWNICY	analiza logowań
	badania użytkowników
	e-learning
	integrowanie BC ze środowiskiem społecznościowym użytkowników
	interfejs użytkownika i prezentacja danych
	modelowanie użytkowników
	personalizacja BC
	systemy skoncentrowane na użytkownika
	tworzenie skoncentrowane na użytkowniku (user centered)
	użytkownicy i interakcje użytkowników, wizualizacja informacji przez użytkowników
	Web 2.0, sieci społecznościowe

Wśród inicjatyw mających na celu edukowanie w zakresie bibliotek cyfrowych, wymianę doświadczeń pomiędzy teoretykami i praktykami, prezentowanie wyników badań repozytoriów cyfrowych w wymiarze społeczno-kulturowym, wyszczególnić należy także konferencję *International Conference on Qualitative and Quantitative Methods in Libraries, QQML* (Międzynarodową Konferencję Metod Jakościowych i Ilościowych w Bibliotece). Konferencja jest organizowana od 2009 r. (dwie edycje odbyły się w Chani, na Krecie w Grecji, edycja z 2011 r. miała miejsce w Atenach, także w Grecji, zaś od 2012 r. konferencja będzie realizowana w róznych miastach europejskich rozpoczynając od Limerick w Irlandii). W 2010 r. odbyła się w ramach QQML równorzędne forum Digital Library Multi-Conference (MultiKonferencja Biblioteki Cyfrowe), która w całości poświęcona była tematyce repozytoriów cyfrowych, następnie obszar tematyczny włączono do nadrzędnej konferencji [ISAST 2011, dok. elektr.]. Wśród zagadnień poruszanych podczas konferencji QQML oraz DCML znalazły się kwestie repozytoriów instytucjonalnych, archiwów cyfrowych, eksploracji danych, organizacji i zarządzania bibliotekami cyfrowymi, digitalizacji i przechowywania zasobów cyfrowych, oprogramowaniem dla BC, aspektów legislacyjnych i prawa autorskiego, ontologii, semantyki i sieci semantycznych [DLMC 2011, dok. elektr.]. W 2012 r. nie zachowano konwencji osobnego sympozjum współorganizowanego wraz z QQML i ponownie, jak w latach 2009 i 2010 włączono ten zakres tematyczny do ogólnych treści konferencji.

Konferencje amerykańskie i kanadyjskie

Rozwój nowych form komunikacji i informacji na skalę światową oraz tworzenie bibliotek cyfrowych przez liczne biblioteki, archiwa i muzea, zapoczątkowała serię wydarzeń, które były organizowane w Stanach Zjednoczonych oraz Kanadzie. Najpopularniejszą i jedną z największych konwencji jest *Joint Conference on Digital Libraries* (JCDL) – Wspólna Konferencja Bibliotek Cyfrowych. Kongres ten jest jednym z najistotniejszych forów o charakterze międzynarodowym, koncentrującym się na problematyce bibliotek cyfrowych oraz związanych z nimi kwestii technicznych, praktycznych oraz społecznych. Wydarzenie powstało jako połączenie trzech najważniejszych, cyklicznych konferencji amerykańskich, które były organizowane od 1994 r. przez *Institute of Electrical and Electronics Engineers Computer Society, IEEE CS (Stowarzyszenie Komputerowe Instytutu Inżynierów Elektryki i Elektroniki), Association for Computer Machinery, ACM (Stowarzyszenie dla Maszyn Liczących), największego towarzystwa naukowego informatycznego oraz dwóch edycji konferencji Digital Libraries. The First Annual Conference on the Theory and Practice of Digital Libraries* (Biblioteki Cyfrowe. Pierwsza Coroczna Konferencja na temat Teorii i Praktyki Bibliotek Cyfrowych) z lat 1994-1995 [JCDL, George Washington University 2011, dok. elektr.][3]. Sympozja poświęcone były zagad-

[3] Obie konferencje Digital Libraries zostały zorganizowane w Teksasie, w Stanach Zjednoczonych (College Station – w 1994 r. oraz Austin – w 1995 r.). Sympozja były współfinansowane przez liczne stowarzyszenia amerykańskie, w tym Amerykańskie Stowarzyszenie Informacji Naukowej, jak: ACM SIGLINK, SIGBIT, SIGCAPH, SIGCUE, SIGIR, SIGOIS, American Society for Information Science

nieniom różnorodnie pojmowanej biblioteki cyfrowej, rozumianej jako a) system informacyjny i wyszukiwawczy, organizujący zasoby cyfrowe, b) serwis o nowym znaczeniu, posadowiony w środowisku cyfrowym, obejmujący wielorakie funkcje – gromadzenia, organizowania i dystrybucji repozytoriów cyfrowych, c) obiekt, model teoretyczny i narzędzie transferu informacji, d) rozszerzenie bibliotek tradycyjnych i usług oferowanych przez te instytucje [JCDL, dok. elektr.]. Wśród zagadnień, które omawiano podczas konferencji *Digital Libraries w latach 1994-1995, oprócz prezentacji konkretnych przykładów implementacji i organizacji bibliotek cyfrowych, znalazły się kwestie związane m.in. z: 1) modelami operacyjnymi, modelami prototypowymi, strukturą i architekturą, 2) zarządzania systemem, 3) własności intelektualnej, 4) ontologii, 5) użyteczności, 6) polityki tworzenia i udostępniania zasobów, 7) ewaluacji użytkowników, 8) organizacji i zarządzania wiedzą, 9) struktury zasobów cyfrowych, wizualizacji i nawigacji katalogów, heterogeniczności danych, 10) interfejsu, 11) struktury danych zasobów cyfrowych [Schnase i in. 1994, dok. elektr.; Shipman 1995, dok. elektr.].*

W 1995 r. organizacja dorocznej konferencji poświęconej bibliotekom cyfrowym została przejęta przez Instytut Inżynierów Elektryki i Elektroniki. Forum o nazwie *IEEE Advances in Digital Libraries Conference* (ADL, Konferencja IEEE Procesy Bibliotek Cyfrowych) przygotowane w McLean, w stanie Wirginia poświęcone zostało: 1) wizualizacji, 2) wyszukiwaniu informacji, 3) odnajdywaniu zasobów oraz informacji sieciowej, 4) tworzeniu bibliotek cyfrowych i ich prototypom [Adam i in. 1995, dok. elektr.]. Druga edycja (Waszyngton, 1996 r.) dotyczyła przede wszystkim procesów wyszukiwania informacji, dostępności serwisów oraz aplikacjom medycznym [ADL 1996, dok. elektr.]. Z kolei trzecia konferencja (Waszyngton, 1997 r.) odnosiła się do prezentacji cyfrowych bibliotek europejskich (skany Niemieckiej Bibliografii Narodowa 1601-1700, Biblioteka Wirtualna), zasobom wizualnym w bibliotekach, projektom tych systemów informacyjno-wyszukiwawczych oraz dostępnej technologii wspomagającej tworzenie i implementację bibliotek [ADL 1997, dok. elektr.]. W 1998 r. podczas konferencji ADL, która odbyła się w Santa Barbara, podjęto się zaprezentowania przede wszystkim problemów dotyczących 1) automatyzacji procesów ekstrakcji, 2) użyteczności bibliotek cyfrowych, indeksowania i dostępności informacji, 3) elektronicznego publikowania w przestrzeni biblioteki cyfrowej, 4) technikom zastosowania sztucznej inteligencji, 5) procesów wyszukiwania, 6) meta informacji oraz 7) wizualizacji informacji [ADL 1998, dok. elektr.]. Rok później, w 1999 r., w Baltimore podjęto się charakterystyki 1) ontologii, 2) przeszukiwania baz w bibliotekach cyfrowych, 3) metadanych, 4) technologii bibliotek cyfrowych, 5) wyszukiwania informacji, 6) aplikacji muzycznych, matematycznych, wizualnych (wideo) [ADL 1999, dok. elektr.]. W ostatnim roku organizowania konferencji IEEE (Waszyngton, 2000 r.) zagadnienia dotyczyły m.in. 1) systemów bibliotek cyfrowych, 2) interakcjom człowiek-komputer, 3) reprezentacji wiedzy i bezpieczeństwa, 4) interoperacyjności i ontologii, 5) modelom dokumentów, 6) filtrowaniu statystyki [ADL 2000, dok. elektr.].

(ASIS), Association of Research Libraries (ARL), Knowledge Systems, Inc., School of Information Studies at Syracuse University, Xerox PARC.

Równolegle, od 1996 r. konferencje dotyczące bibliotek cyfrowych były organizowane przez ACM. Pierwsze forum odbywające się w Bethesda, w stanie Maryland dedykowane zostało podczas sesji roboczych, posterowych oraz warsztatów, problematyce: 1) multimedialnych bibliotek cyfrowych, 2) polityce tworzenia repozytoriów cyfrowych i perspektywach ich rozwoju, 3) interakcjom człowiek-komputer, w szczególności przeglądaniu i wizualizacji informacji, a także dokumentom wizualnym i organizacji przestrzennej, 4) obiektom cyfrowym, 5) wyszukiwaniu informacji oraz 6) indeksowaniu i analizowaniu [Digital Libraries 1996, dok. elektr.]. Filadelfijska konferencja, zorganizowana rok później, w 1997 r., została zadedykowana kwestiom: 1) reprezentacji wiedzy, 2) systemowym agentom, 3) multimediom i dokumentom wizualnym, 4) środowisku biblioteki cyfrowej, 5) problematyce systemów informacyjno-wyszukiwawczych, 6) dostępności cyfrowych zasobów, 7) kolekcjom cyfrowym, 8) bazom danych, 9) społeczności użytkowników, 10) nawigowaniu i przeszukiwaniu bibliotek cyfrowych [ACM Digital Libraries 1997, dok. elektr.] W latach 1998-2000, podczas ostatnich trzech edycji konferencji ACM tematyka dotyczyła przede wszystkim:

1) technologii, w tym a) różnorodnych aspektów technologicznych (serwery, nowe technologie), interfejsów, b) modeli prototypowych bibliotek cyfrowych, c) oprogramowania dla bibliotek cyfrowych, d) wizualizacji zasobów multimedialnych, różnorodności cyfrowych dokumentów, ich przechowywania, e) użyteczności i dostępności
2) użytkowników i społeczności odbiorców, a zwłaszcza f) aspektów psychologii i zachowań informacyjnych (w procesie wyszukiwania informacji w bibliotekach cyfrowych, g) modeli użytkowników, h) perspektyw społecznych w tworzeniu i implementacji bibliotek cyfrowych, i) zagadnień zachowań informacyjnych w środowisku bibliotek cyfrowych; analizy adnotacji i cytowań
3) procesów, czyli j) wyszukiwania informacji, narzędzi wyszukiwawczym, usług w bibliotekach cyfrowych k) indeksowania, klasyfikowania i linkowania zasobów; metadanych l) interoperacyjności i globalnej współpracy bibliotek cyfrowych,
4) organizacji informacji, a przede wszystkim m) ontologii, n) reprezentacji wiedzy, jej zarządzania oraz prezentacji informacji.

Wśród zagadnień odnoszących się do bibliotek cyfrowych, po raz pierwszy podczas konferencji ACM i IEEE, uwagę poświęcono sprawom legislacji i praw autorskich oraz jakości tych systemów informacyjno-wyszukiwawczych [ACM Digital Libraries 1998,1999, 2000, dok. elektr.].

W 2001 r. postanowiono połączyć oba, coroczne wydarzenia organizowane przez ACM i IEEE. W tym też roku, w Roanoke (stan Wirginia, Stany Zjednoczone) odbyła się po raz pierwszy Wspólna Konferencja Bibliotek Cyfrowych (*Joint Conference on Digital Library*). Bazując na doświadczeniu poprzednich sympozjów, które stały się podstawą integracji w jedno wydarzenie, JCDL jest kluczowym przedsięwzięciem w Stanach Zjednoczonych i Kanadzie, mającym międzynarodową specyfikę, rozszerzając swoje oddziaływanie na międzynarodowe środowisko teoretyków i praktyków.

Od początku działalności Wspólna Konferencja Bibliotek Cyfrowych gromadzi specjalistów z różnych dziedzin, w tym informatyki, informacji naukowej i bibliotekoznawstwa, nauk medycznych, społecznych i humanistycznych, pracowników archiwów i muzeów. Organizatorzy zachęcają także do czynnego uczestnictwa zarówno przedstawicieli sektora rządowego, środowiska akademickiego, branże nowych technologii informacyjno-komunikacyjnych. Zakres tematyczny obejmuje zagadnienia związane z bibliotekami cyfrowymi, obejmujące perspektywę infrastrukturalną, instytucjonalną, społeczną, polityki rozwoju kolekcji cyfrowych, usług oferowanych przez tego typu system informacyjno-wyszukiwawczy, a także wszelkich interakcji pomiędzy komputerem i człowiekiem. Do obszarów prezentowanych podczas sympozjów zaliczono również ewaluację bibliotek cyfrowych, prawa własności intelektualnej oraz komplementarne analizy wszystkich użytkowników zainteresowanych w tworzenie i wdrażanie bibliotek cyfrowych. Po zorganizowaniu pierwszego sympozjum w Roanoke, kolejno JCDL odbywały się w 2002 r. – Portland, stan Oregon, 2003 r. – Houston, stan Teksas, 2004 r. – Tucson, stan Arizona, 2005 r. – Denver, stan Kolorado, 2006 r. – Chappel Hill, stan Północna Karolina, 2007 r. – Vancouver w Kanadzie, 2008 r. – Pittsburgh, stan Pensylwania, 2009 r. – Austin, stan Teksas, 2010 r. – Go Coast, Queensland, w Australii 2011 r. - Ottawa, Kanada [ACM/IEEE JCDL, dok. elektr.]. W tabeli poniżej przedstawiono pogrupowane według obszarów zagadnieniowych, które były podejmowane przez uczestników Wspólnej Konferencji Bibliotek Cyfrowych za lata 2001-2011 [JCDL, dok. elektr.].

Tabela 3. Problematyka bibliotek cyfrowych poruszana podczas konferencji JCDL za lata 2001-2011 [oprac. autorki].

OBSZAR	NAJWAŻNIEJSZE ZAGADNIENIA poruszane podczas konferencji JCDL w latach 2001-2011
UŻYTKOWNICY	aspekt społeczny
	badania użytkowników i interfejsy odbiorców
	badanie czytelnictwa w bibliotekach cyfrowych
	ewaluacja użytkowników bibliotek cyfrowych - ocena formatywna i sumatywna
	interakcje użytkownika z systemem
	interakcje użytkowników
	jakościowe badania użytkowników
	personalizacja
	społeczności bibliotek cyfrowych
	udoskonalanie oddziaływania poprzez badania użytkowników, potrzeby informacyjne, zachowania informacyjne
	użytkownicy - potrzeby informacyjne i percepcja
	użytkownicy i wymiar społecznościowy
	zachowania informacyjne, wyszukiwanie informacji i personalizacja

TECHNOLOGIA	algorytmy dla muzycznych bibliotek cyfrowych
	analiza dokumentów
	aplikacje dla bibliotek cyfrowych w naukach społecznych i humanistycznych
	aplikacje Open Initiative Archives (OAI)
	archiwizacja i narzędzia sieciowe dla bibliotek cyfrowych
	aspekty technologiczne
	automatyczna klasyfikacja
	automatyczne mechanizmy rozumienia tekstu
	automatyczne narzędzia dla bibliotek cyfrowych
	automatyczne tworzenie metadanych
	automatyczne zarządzanie zasobami w bibliotece cyfrowej
	cyfrowe odzyskiwanie i ochrona zasobów
	dane sieciowe
	geograficzne systemy informacyjne w bibliotekach cyfrowych
	hipertekst
	identyfikacja haseł (nazw osób i miejsc)
	implementacja systemów
	infrastruktura bibliotek cyfrowych i systemy informacyjno-wyszukiwawcze
	integrowanie danych fizycznych i cyfrowych
	interfejsy i nawigacja
	mechaniczne zarządzanie kolekcją cyfrową
	metadane
	metody klasyfikowania i organizowania zasobów w bibliotekach cyfrowych
	narzędzia do tworzenia i użytkowania bibliotek cyfrowych
	nowoczesne interfejsy użytkowników
	ochrona i archiwizacja danych
	ochrona zasobów cyfrowych
	pozyskiwanie informacji (ekstrakcja)
	standardy, metadane
	strukturalizowanie i translacja zapytań wyszukiwawczych
	systemy rekomendowania zasobów cyfrowych
	udoskonalanie zarządzania kolekcjami
	wspomaganie techniczne klasyfikowania
	zarządzanie zasobami i usługami
	zastosowanie maszynowego uczenia do rozwoju kolekcji

KOLEKCJE CYFROWE, ZASOBY	artefakty
	kolekcje na wielką skalę (large scale)
	kolekcje osobiste
	modele zawartości kolekcji cyfrowych
	nazewnictwo obiektów cyfrowych
	oddziaływanie zasobów (interakcje z użytkownikami)
	ontologie
	polityka ochrony zasobów cyfrowych
	polityka tworzenia kolekcji cyfrowych
	publikacje naukowe w bibliotekach cyfrowych
	reprezentacja wiedzy
	ujednolicanie zasobów
	wizualizacja informacji w bibliotekach cyfrowych
	zasoby tekstowe w bibliotece cyfrowej
	zawartość kolekcji cyfrowych
FUNKCJE i PROCESY	eksploracja danych
	ewaluacja zasobów cyfrowych
	interoperacyjność
	modele i narzędzia do generowania bibliotek cyfrowych
	nowe środowiska wyszukiwawcze
	rekomendacje i rozpowszechnianie
	schemat tworzenia bibliotek cyfrowych
	strategie wyszukiwawcze
	tworzenie systemów i ewaluacja dla środowisk akademickich, naukowe biblioteki cyfrowe
	współpraca bibliotek cyfrowych
	wyszukiwanie i przeszukiwanie bibliotek cyfrowych, przeszukiwanie bibliotek cyfrowych
	wyszukiwanie informacji i generowanie danych
	zastosowanie i relacyjność
EDUKACJA	biblioteki cyfrowe dla edukacji
	biblioteki cyfrowe w cyberprzestrzeni i aspekt edukacyjny
	biblioteki cyfrowe w programach nauczania
	biblioteki cyfrowe w systemie komunikacji naukowej
	edukacyjne biblioteki cyfrowe
	perspektywa bibliotek cyfrowych w edukacji - badania technologii, serwisów i użytkowników
	wspieranie edukacji przez biblioteki cyfrowe

TEORIA, MODELE BC, LEGISLACJA	heurystyka bibliotek cyfrowych
	prawo autorskie, prawa udostępniania i wykorzystywania
	teoria bibliotek cyfrowych
	aspekt geograficzny bibliotek cyfrowych
BIBLIOTEKI CYFROWE (TYPY), PARADYGMAT	aspekt multimedialny
	audiowizualne biblioteki cyfrowe
	historyczne biblioteki cyfrowe
	muzyczne biblioteki cyfrowe
	przestrzenne biblioteki cyfrowe
	tworzenie reprezentacji informacji dla nauk humanistycznych
	wizualne i kulturowe biblioteki cyfrowe
	wykorzystanie bibliotek cyfrowych w naukach humanistycznych

Wśród innych wydarzeń o charakterze międzynarodowym, podczas którego prezentowane są zagadnienia związane z problematyka bibliotek cyfrowych jest *International Conference on Open Repositories* (Międzynarodowa Konferencja Otwarte Repozytoria). *K*onferencja ta, organizowana każdego roku w Stanach Zjednoczonych, jest wydarzeniem skupiającym międzynarodową społeczność, osób zainteresowanych oraz zaangażowanych w rozwój, implementację, zarządzanie i zastosowanie cyfrowych repozytoriów. Dotychczas miało miejsce 6 edycji konferencji, od 2006 r. odbyły się one w Sydney (2006 r., Australia), teksaskim San Antonio (2007 r., Stany Zjednoczone), Southampton (2008 r., Wielka Brytania), Atlancie, w stanie Georgia (2009 r., Stany Zjednoczone), Madrycie (2010, Hiszpania), Austin w stanie Teksas (2011 r., Stany Zjednoczone). W roku 2012 konferencję zaplanowano w Edynburgu, w Szkocji [Open Repositories, dok. elektr.]. Podczas cyklicznych spotkań, przedstawiane są zagadnienia dotyczące strategicznych, technicznych, teoretycznych i praktycznych rozwiązań dotyczących bibliotek cyfrowych. Prowadzone są warsztaty dla użytkowników otwartych repozytoriów (oprogramowania DSpace, EPrints, Fedora). Podczas konferencji są eksplorowane i prezentowane zagadnienia dotyczące m.in.: charakterystyki i oceny całego cyklu informacji, tworzenia i zarządzania treściami cyfrowymi, umożliwiającego stosowanie, ponowne wykorzystanie i wzajemne połączenie zasobów oraz zapewnienie długoterminowej konserwacji i archiwizacji repozytoriów; tworzenia, skalowania i utrzymywania między instytucjonalnych działań w przestrzeni cyfrowej biblioteki; tworzenia serwisów skoncentrowanych na użytkowniku, jak również modeli komunikacji dla odbiorców szkolnych i akademickich, naukowców; tworzenia modeli długoterminowego utrzymywania systemów informacyjno-wyszukiwawczych oraz biznesowych rozwiązań wspomagających rozwój cyfrowych repozytoriów; kluczowych kryteriów, probierzy współpracy wszystkich użytkowników systemu – administratorów, informatyków, bibliotekarzy i odbiorców końcowych; narzędzi, technologii i dobrych praktyk wspierający tworzenie zespołów o różnorodnych kompetencjach; ewoluowania i wzbogacania umiejętności twórców, kadry za-

rządzającej, deweloperów repozytoriów cyfrowych; obszarów efektywnej współpracy w przestrzeni biblioteki cyfrowej [Open Repositories, dok. elektr.].

Konferencje azjatyckie

Ekspansywny rozwój bibliotek cyfrowych spowodował wzrost zainteresowania wdrożeniami azjatyckie środowiska uniwersyteckie i bibliotekarskie, gdzie odbyło się do tej pory trzynaście edycji międzynarodowej konferencji dedykowanej repozytoriom cyfrowym w perspektywie krajów azjatyckich (uwzględniając doświadczenia państw azjatyckich tworzących zasoby cyfrowe). Pierwsza Międzynarodowa Konferencja Azjatyckich Bibliotek Cyfrowych (*International Conferences of Asian Digital Libraries*, ICADL) została przygotowana w formie warsztatów. Odbyły się one w sierpniu 1998 r., w Hong Kongu i nosiły nazwę *First Asia Digital Library Workshop* (Pierwsze Azjatyckie Warsztaty Bibliotek Cyfrowych). Tematem tego spotkania, zatytułowanego *East Meets West* (Wschód spotyka Zachód) była problematyka tworzenia, wdrażania i doskonalenia repozytoriów cyfrowych. Obrady i wymiana wiedzy toczyła się w środowisku naukowców z krajów Azji i Pacyfiku, jak w Chiny, Hong Kong, Tajwan, Singapur, Korea i Nowa Zelandia. W 1999 r. sympozjum ICADL odbyło się w Taipei, na Tajwanie i stanowił także formę warsztatów. W 2000 r. konferencja została zorganizowana w Seulu, w Korei, gdzie nastąpił wzrost znaczenia oraz zainteresowania krajów, które nie były pierwotnie związane z sympozjum. Czwarta edycja ICADL z 2001 r., odbyła się w Bangalore, w Indiach. Kolejna konferencja odbyła się w Singapurze, w 2002 r., zaś uczestniczyli w niej reprezentanci Azji i Pacyfiku, Europy, Stanów Zjednoczonych Ameryki Północnej, Malezji i Nepalu. W 2003 r., do szóstej edycji ICADL, która została zrealizowana w Kuala Lumpur, dołączyli uczestnicy z Iranu. Kolejne sympozjum miało miejsce w Szanghaju, w Chinach (2004 r.), następnie w Bangkoku, w Tajlandii (2005 r.), Kioto, w Japonii (2006 r.), Hanoi, w Wietnamie (2007 r.), Bali, w Indonezji (2008 r.). Konferencja nie odbyła się w 2009 r. Ostatnie przedsięwzięcie zostało zorganizowane w ścisłej współpracy zarówno ICADL, jak i amerykańskiego sympozjum JCDL [Artificial Intelligence Lab ICADL, dok. elektr.]. Zorganizowane w australijskim Brisbane, w 2010 r., międzynarodowe forum Wspólna Joint JCDL/ICADL *International Digital Libraries Conference* (Międzynarodowa Konferencja Bibliotek Cyfrowych), zatytułowane zostało *Digital Libraries in a Time of Global Change* (Biblioteki Cyfrowe w Okresie Globalnych Zmian). Wśród zagadnień poruszanych na konferencji znalazły się przede wszystkim problemy: 1) tworzenia bibliotek cyfrowych, 2) wyszukiwania zasobów cyfrowych, 3) użytkowników, 4) zarządzania biblioteką cyfrową, 5) funkcji, procesów, jak adnotowania i markowania zasobów, 6) zasobów cyfrowych dziedzictwa kulturowego, 7) publikacji naukowych w bibliotekach cyfrowych, 8) perspektywie międzynarodowej bibliotek cyfrowych, 9) migracji zasobów i ich mobilności, 10) języków naturalnych w systemach informacyjno-wyszukiwawczych, jakimi są biblioteki cyfrowe, 11) cyfrowych zasobów historycznych, 12) metadanych, 13) użyteczności i nawigowaniu, 14) współpracy śro-

dowisk informacyjnych, 15) kolekcji personalizowanych, 16) teoretycznych aspektów bibliotek cyfrowych, 17) wizualizacji, eksploracji danych i obiektom wizualnym, 18) architektury i struktury bibliotek cyfrowych, 19) integracji zasobów fizycznych i cyfrowych, 20) aspektów społecznych [University of Queensland 2010, dok. elektr.]. W trakcie konferencji zorganizowano także trzy warsztaty poświęcone muzycznym repozytoriom cyfrowym (*Music for the Masses*), międzynarodowego aspektu rozwoju bibliotek cyfrowych (*Digital Libraries for International Development*)oraz perspektywie wykorzystania bibliotek cyfrowych w edukacji (*Digital Libraries and Education*) [University of Queensland 2010, dok. elektr.]. Tematyka konferencji, panele i sesje oraz warsztaty nie odbiegały rzeczowo od zagadnień podejmowanych podczas amerykańskiej edycji JCDL z 2010 r.

Wśród konferencji azjatyckich, na uwagę zasługuje sympozjum zorganizowane w Indiach, w Bangalore, które przeniesiono w 2009 r. do Trento, we Włoszech. Było to forum *International Conferences on Digital Libraries and the Semantic Web, ICSD* (Pierwsza Międzynarodowa Konferencja Bibliotek Cyfrowych i Sieci Semantycznych), które zostało zorganizowane w 2007 r. przez *Indian Statistical Institute in Bangalore* – Indyjski Instytut Statystyczny w Bangalore, w Indiach. Duże zainteresowanie konferencją, zagadnieniami skupiającymi się wokół problematyki sieci semantycznych, zastosowaniu ontologii w bibliotekach cyfrowych, wykazało potrzebę stworzenia obszaru komunikacji między społecznościami praktyków i teoretyków. Celem jest ułatwienie tworzenia różnorodnych aplikacji bibliotek cyfrowych, które mogłyby wspomagać funkcje semantyczne. W 2009 r., konferencja ICSD została przygotowana we włoskim Trento [ICSD, dok. elektr.].

Konferencje rosyjskie

Na terenie Rosji także podjęto się różnorodnych działań mających na celu wyłonienie kluczowej konferencji, nie odbiegającej tematycznie od światowych trendów w dziedzinie bibliotek cyfrowych. Należy jednak zaznaczyć, że rosyjski odpowiednik JCDL oraz ECDL, RCDL, intensywnie przez te dwie organizacje wspierany, przede wszystkim jest zainteresowany wymiarem technologicznym i strukturalnym repozytoriów cyfrowych, zdecydowanie mniej miejsca poświęcając samym użytkownikom.

Celem serii *All-Russian Research Conference Digital Libraries: Advanced Methods and Technologies, Digital Collections*, RCDL (Ogólnorosyjska Konferencja Naukowa Bibliotek Cyfrowych: Zaawansowane Metody i Technologie, Kolekcje Cyfrowe), organizowanej od 1999 r., jest stymulowanie rozwoju bibliotek cyfrowych krajów Wspólnoty Rosyjskiej, a także wspierania badań naukowych w dziedzinie formowania, wdrażania i doskonalenia tych systemów informacyjno-wyszukiwawczych. RCDL stanowi osobliwe forum do wymiany doświadczeń, wyników badań, teoretycznych i praktycznych rozwiązań, jak również komunikacji i współpracy pomiędzy ekspertami w danej dziedzinie[4]. Konferencja wspierana jest przez *Russian Foundation*

[4] Podczas trzynastu edycji konferencji RCDL prezentowano doświadczenia specjalistycznych, dzie-

for Basic Research (Rosyjską Fundację Badań Podstawowych) oraz moskiewski oddział *Association for Computing Machinery Special Interest Group on Management Of Data* ACM SIGMOD (Stowarzyszenie dla Maszyn Liczących. Specjalna Grupa Zainteresowanych Zarządzaniem Danymi). Organizatorzy konferencji RCDL pozyskują również uczestników zagranicznych, a także poprzez rozpowszechnianie wiedzy o repozytoriach cyfrowych, ułatwia współpracę międzynarodową. Język angielski stanowi drugi język urzędowy konferencji, która jest także przedsięwzięciem interdyscyplinarnym. RCDL ściśle współpracuje z europejskim i amerykańskimi oraz kanadyjskimi konferencjami ECDL i JCDL, stymulując tym samym globalną współpracę w zakresie bibliotek cyfrowych jako systemów informacyjno-wyszukiwawczych. Dotychczas coroczne sympozja RCDL odbywały się w: 1999 r. – Sankt Petersburg, 2000 r. – Protwino, 2001 r. – Pietrozawodsk, 2002 r. – Dubna, 2003 r. – Sankt Petersburg, 2004 r. – Puszkino, 2005 r. – Jarosław, 2006 r. – Suzdal, 2007 r. – Peresław, 2008 r. – Dubna, 2009 r. – Pietrozawodsk, 2010 r. – Kazań, 2011 r. – Woroneż [RCDL, dok. elektr.].

Od 2003 r. do konferencji dołączono warsztaty organizowane każdego roku przez *Russian Information Retrieval Evaluation Seminar* ROIMP (Rosyjskie Seminarium Ewaluacyjne Wyszukiwania Informacji) [RCDL 2003, dok. elektr.]. W 2005 r. podczas sympozjum w Jaroslawiu odbyły się dodatkowe warsztaty opracowane i przeprowadzone przez Yandex, czołową firmę internetową w Rosji [RCDL 2005, dok. elektr.]. Od 2009 w ramach konferencji RCDL prowadzone są warsztaty *Doctoral Research on Information Technologies Related with Digital Libraries* (Badania Doktoranckie na rzecz Technologii Informacyjnych Związanych z Bibliotekami Cyfrowymi) podczas których istnieje możliwość przedstawienia bieżących wyników badań, a także uzyskania konstruktywnej krytyki ze strony teoretyków i praktyków, uczestników RCDL. Specyfika zagadnień konferencyjnych dotyczy przede wszystkim, rozwiązań technologicznych w zakresie tworzenia i implementowania bibliotek cyfrowych, interoperacyjności i syntezy informacji, tworzenia bibliotek cyfrowych oraz opracowywania kolekcji cyfrowych, automatyzacji wszystkich procesów, metadanych, ontologii, i narzędzi prezentacji informacji semantyki zasobów cyfrowych, ontologii i wyszukiwania informacji. Systemowe podejście uwzględnia większość zagadnień związanych ze strukturą danych, integralnością zasobów cyfrowych, inteligentnych systemów informacyjno-wyszukiwawczych. Prezentowane problemy dotyczą obszaru architektury systemu, funkcjonalności i użyteczności, mniej uwagi poświęcając zagadnieniom użytkowników i paradygmatu społeczno-kulturowego, a także kognitywnego. Szczegółowe przedstawienie najważniejszych kwestii eksponowanych podczas wszystkich edycji konferencji RCDL, od 1999 r. do ostatniego sympozjum zorganizowanego w 2011 r. ukazano w poniższej tabeli.

dzinowych repozytoriów cyfrowych z zakresu nauk medycznych, rolniczych, biologicznych, biofizyki, geologii, geofizyki, nauk humanistycznych, w tym filologicznych, a także ścisłych.

Tabela 4. Problematyka bibliotek cyfrowych poruszana podczas konferencji RCDL za lata 1999-2011 [oprac. autorki].

OBSZAR TEMATYCZNY	NAJWAŻNIEJSZE ZAGADNIENIA poruszane podczas konferencji RCDL w latach 1999-2011
TECHNOLOGIA	agenci i zasoby sieciowe
	algorytmy rangowania
	analiza systemów heterogenicznych
	architektura bibliotek cyfrowych
	automatyczna kategoryzacja zasobów cyfrowych
	automatyczna klasyfikacja informacji i treści zasobów cyfrowych, klasyfikacja pełnotekstowych obiektów cyfrowych
	automatyczne tworzenie i wspieranie XML
	bezpieczeństwo informacji i systemów informacyjno-wyszukiwawczych
	biblioteka cyfrowa jako informacyjny system korporacyjny
	generowanie danych strukturalnych
	integracja narzędzi eksploracji danych
	integracja heterogenicznych modeli strukturalnych danych, zasobów cyfrowych, eksploracja danych
	integracja cyfrowych archiwów
	integracja danych personalnych
	integracja systemów informacyjnych i bibliotek cyfrowych
	inteligentny agent i tworzenie kolekcji cyfrowych
	lingwistyka korporatywna w oparciu o specjalne biblioteki cyfrowe
	metadane i standardy międzynarodowe
	metody automatycznego gromadzenia informacji w bibliotece cyfrowej
	model produkcji i rozpowszechniania danych cyfrowych
	model statystyczny dokumentów XML
	modele obiektów w języku XML
	modelowanie baz danych
	narzędzia publikacji metadanych w zintegrowanych systemach informacyjnych
	nowe technologie, narzędzia w bibliotekach cyfrowych
	ontologie, automatyczne tworzenie terminologicznych baz danych, specyfikacja relewantności zasobów cyfrowych
	procesy digitalizacji zasobów do bibliotek cyfrowych
	przetwarzanie danych strukturalnych w bibliotece cyfrowej
	relacyjne systemy zarządzania wiedzą, relacyjne bazy danych i ontologie
	różnice semantyczne, ontologie
	sieć semantyczna
	spójność danych strukturalnych
	standardy i technologie dla usług i integrowania danych

TECHNOLOGIA	struktura danych w systemach informacyjno-wyszukiwawczych
	struktura i tworzenie bibliotek cyfrowych
	struktura inteligentnych systemów informacyjnych
	systemy generowania dynamicznych zasobów internetowych
	systemy ochrony danych
	systemy wyszukiwania informacji
	systemy zarządzania biblioteką cyfrową
	tworzenie rozproszonych bibliotek cyfrowych, środowiska rozproszonej informacji
	złożoność danych systemowych i danych użytkownika
	lokalizacja zasobów sieciowych
BIBLIOTEKA CYFROWA, MODELE, KONCEPCJE, RODZAJE	archiwa i muzea cyfrowe
	biblioteki cyfrowe specjalistyczne, dziedzinowe (nauki humanistyczne, nauki medyczne, nauki biologiczne, astrofizyka, geofizyka, geografia)
	elektroniczna biblioteka
	implementacja modeli bibliotek cyfrowych
	modele bibliotek cyfrowych, modele koncepcyjne i prototypowe
	modele działalności informacyjnej
	modelowanie bibliotek i zasobów cyfrowych
	monitorowanie i rekomendacje dla tworzenia akademickich bibliotek cyfrowych, polityka tworzenia bibliotek cyfrowych
	Open Archives Initiative (OAI)
	naukowe i edukacyjne biblioteki i kolekcje cyfrowe
	polityka tworzenia bibliotek cyfrowych
	rola bibliotek wirtualnych
	zintegrowane modele bibliotek cyfrowych
ZASOBY I KOLEKCJE CYFROWE	identyfikacja kolekcji cyfrowych
	jakość metadanych
	mechanika transkrypcji, ontologie
	różnorodność zasobów cyfrowych, charakterystyka danych
	semantyczna integracja zasobów cyfrowych
	specjalistyczne rodzaje danych dla bibliotek cyfrowych
	środowisko rozproszonej informacji
	ujednolicanie zasobów cyfrowych, struktury danych, środowiska gridowe
	wizualizacja informacji w bibliotekach cyfrowych
	zasoby naukowe w bibliotekach cyfrowych
	zasoby wizualne w bibliotekach cyfrowych
	heterogeniczne zasoby cyfrowe
	heterogeniczność zasobów cyfrowych

PROCESY, USŁUGI BIBLIOTEK CYFROWYCH, PROCESY INFORMACYJNE	adnotacje
	analizy statystyczne, badania bibliotek cyfrowych
	dystrybucja informacji w bibliotekach cyfrowych
	ekstrakcja informacji
	gromadzenie i przechowywanie zasobów cyfrowych
	indeksowanie danych strukturalnych, indeksowanie i przechowywanie zasobów cyfrowych
	interoperacyjność
	kategoryzacja i klasyfikowanie zasobów cyfrowych
	meta-wyszukiwanie informacji
	naturalne języki wyszukiwawcze
	nawigacja i przeglądanie zasobów cyfowych
	organizacja informacji
	organizacja wiedzy w bibliotekach cyfrowych
	przechowywanie zasobów cyfrowych
	rangowanie zasobów cyfrowych
	relacyjność, klasyfikowanie i indeksowanie zasobów cyfrowych
	reprezentacja wiedzy w bibliotekach cyfrowych
	sztuczna inteligencja i wyszukiwanie informacji
	usługi i serwisy bibliotek cyfrowych
	wyszukiwanie informacji, systemy wyszukiwania informacji
	zarządzanie biblioteką cyfrową
	zarządzanie kolekcjami cyfrowymi
	zarządzanie wiedzą
UŻYTKOWNICY BIBLIOTEK CYFOWYCH	aspekt użytkownika w architekturze bibliotek cyfrowych
	badania użytkowników, zachowania informacyjne
	interakcje użytkowników z biblioteką cyfrową
	interfejsy użytkownika
	personalizacja bibliotek cyfrowych
	użytkownicy, percepcja dokumentów elektronicznych
WYMIAR EDUKACYJNY BIBLIOTEK CYFROWYCH	aspekt edukacyjny i zastosowanie bibliotek cyfrowych w kształceniu
	ewaluacja zasobów cyfrowych
	naukometria w oparciu o zasoby bibliotek cyfrowych
	wymiar edukacyjny bibliotek cyfrowych
	wymiar edukacyjny bibliotek cyfrowych (e-learning)

Konferencje polskie

Dynamiczny rozwój polskiej infrastruktury bibliotek cyfrowych i aktywny udział polskich środowisk w projektach europejskich, spowodował rosnące zainteresowanie wszystkich użytkowników repozytoriów cyfrowych, zarówno twórców, decydentów, jak i bibliotekarzy i naukowców problematyką tych systemów informacyjno-wyszukiwawczych. Stał się także impulsem do podjęcia się czynnego udziału w szkoleniach, warsztatach i konferencjach poświęconych tematyce tworzenia, doskonalenia i implementowania bibliotek cyfrowych w Polsce. Wzrastająca liczba bibliotek i instytucji kultury, które dołączyły do społeczności formującej biblioteki cyfrowe, w tym europejskiej grupy tworzącej *Europeanę*[5], utworzenie Federacji Bibliotek Cyfrowych oraz powiększająca się liczba wdrożeń nowych obiektów, powoduje także chęć wymiany doświadczeń nie tylko na forum międzynarodowym, ale także krajowym. Pierwsze działania digitalizacji dóbr kultury rozpoczęto w 1995 r., gdy dopiero w 2002 r. zainaugurowano pierwszą bibliotekę cyfrową (Polska Biblioteka Cyfrowa) [Werla, 2010, dok. elektr.]. Ogólnopolska konferencja, której zadaniem byłoby połączenie wszystkich grup zainteresowanych praktycznym i teoretycznym wymiarem bibliotek cyfrowych zorganizowano dopiero w 2008 r. Forum to zostanie przedstawione w dalszej części artykułu. Do tego czasu liczne instytucje, jak biblioteki, archiwa i muzea przygotowywały lokalne warsztaty i szkolenia.

Jednym z kluczowych przedsięwzięć były organizowane regularnie przez Poznańskie Centrum Superkomputerowo-Sieciowe i Poznańską Fundację Bibliotek Naukowych od 2004 r., *Warsztaty Biblioteki cyfrowe*. Najważniejszym celem warsztatów była popularyzacja bibliotek cyfrowych oraz aktywizacja dynamicznego rozwoju tych systemów informacyjno-wyszukiwawczych w Polsce. Stało się to możliwe poprzez aktywne uczestnictwo wszystkich osób zainteresowanych tworzeniem, wdrażaniem i udoskonalaniem bibliotek cyfrowych, a także przez stworzenie możliwości wymiany doświadczeń oraz integracji środowiska użytkowników repozytoriów cyfrowych (twórców, administratorów, bibliotekarzy, naukowców). Coroczne edycje warsztatów mają specyficzny charakter poprzez powiązanie tematyki prezentowanej podczas sesji wraz z aktualnym stanem rozwoju bibliotek cyfrowych. Do tej pory odbyło się osiem edycji, podczas których charakteryzowano zagadnienia, problemy i praktyki dotyczące: 1. w 2004 r. – możliwości rozwoju bibliotek cyfrowych w Polsce, realizowane projekty (Wielkopolska Biblioteka Cyfrowa), proces digitalizacji (skanowanie, weryfikowanie, opracowywanie, publikowanie), zarządzanie zawartością biblioteki cyfrowej (struktura, klasyfikowanie, indeksowanie, uprawnienia dostępu, czytelnicy); 2. w 2005 r. - Biblioteki cyfrowe w infrastrukturze europejskiej, przykłady implementacji projektów (Dolnośląska Biblioteka Cyfrowa, Zielonogórska Biblioteka Cyfrowa, Kujawsko-Pomorska Biblioteka Cyfrowa), proces digitalizacji, funkcjonalność platform dla bibliotek cyfrowych – aplikacje administratora, czytelnika, oprogramowa-

[5] Europeana stanowi jedną z inicjatyw Komisji Europejskiej. Nadrzędnym celem projektu jest stworzenie portalu internetowego umożliwiającego dostęp do zdigitalizowanego zasobu kulturowego i naukowego dziedzictwa Europy.

nie dla budowy bibliotek cyfrowych, rozproszone biblioteki cyfrowe, zaawansowane przetwarzanie tekstu dla potrzeb bibliotek cyfrowych; 3. w 2006 r. – a) wprowadzenie do tematyki bibliotek cyfrowych – systemy bibliotek cyfrowych, organizacja biblioteki cyfrowej opartej, przegląd cech funkcjonalnych oprogramowania dla bibliotek cyfrowych (dLibra), b) czytelnik w bibliotece cyfrowej, c) tworzenie zasobu biblioteki cyfrowej, d) zarządzanie biblioteką cyfrową, e) praktyczne aspekty tworzenia bibliotek cyfrowych (m.in. interfejs użytkownika, implementacja technologii Web 2.0, badania statystyczne, społeczność użytkowników); 4. w 2007 r. – a) budowa i wykorzystanie bibliotek cyfrowych – wprowadzenie do problematyki, tworzenie zasobu biblioteki cyfrowej, biblioteki cyfrowe w środowisku sieciowym, czytelnicy bibliotek cyfrowych, b) mechanizmy automatycznego grupowania wyników wyszukiwania w bibliotekach cyfrowych, c) systemy identyfikacji obiektów cyfrowych w środowisku sieciowym, d) rozproszone systemy uwierzytelniania użytkowników, e) zaawansowane usługi dla rozproszonych bibliotek cyfrowych; 5. w 2008 r. a) wprowadzenie do tematyki bibliotek cyfrowych, b) tworzenie i zarządzanie biblioteką cyfrową, c) czytelnik w bibliotece cyfrowej, d) biblioteki cyfrowe w środowisku sieciowym, 6. w 2009 r. – a) zachowania informacyjne użytkowników bibliotek cyfrowych, b) pozycjonowanie zasobów cyfrowych, c) struktura i funkcjonowanie systemów długoterminowej archiwizacji dziedzictwa kulturowego, d) przechowywanie danych, e) edukacja w zakresie implementacji bibliotek cyfrowych, f) społeczności Web 2.0 w bibliotece cyfrowej (moduł użytkownika i bibliotekarza), g) metadane, Federacja Bibliotek Cyfrowych; 7. w 2010 r. – a) aspekt technologiczny bibliotek cyfrowych – skalowanie i monitorowanie działania serwera systemu, archiwizacja danych, masowe przesyłanie danych, b) automatyzacja procesu publikowania w bibliotece cyfrowej, c) przechowywanie danych – nowoczesne rozwiązania, d) organizacja repozytorium instytucjonalnego, e) interfejs użytkownika, f) systemy dla bibliotek cyfrowych; 8. w 2011 r. a) nowe usługi dla twórców i użytkowników bibliotek cyfrowych, najnowszym trendom i rozwiązaniom wirtualizacji, b) nowe technologie – zdalne zarządzanie, energooszczędność, wirtualizacja, rodziny serwerów, przeznaczenie i pozycjonowanie, c) system obsługi biblioteki cyfrowej, d) system do zarządzania procesem digitalizacji.

Od 2008 r. ze względu na aktywny rozwój bibliotek cyfrowych w Polsce oraz zainteresowanie tematyką tych obiektów Poznańskie Centrum Superkomputerowo-Sieciowe postanowiło rozszerzyć organizowane we współpracy z Fundacją Bibliotek Naukowych z Poznania warsztaty, których formułę rozszerzono w konferencję *Polskie biblioteki cyfrowe*. Organizatorami konferencji są: Biblioteka Kórnicka PAN, Poznańska Fundacja Bibliotek Naukowych oraz Poznańskie Centrum Superkomputerowo-Sieciowe. Konferencję powiązano także ze szkoleniami wprowadzającymi w problematykę bibliotek cyfrowych. Podczas czterech dotychczasowych edycji konferencji, odbywającej się w latach 2008-2011, połączono doświadczenia zarówno teoretyków, jak i praktyków, dając możliwość przedstawienia referatów oraz podjęcia się dyskusji na temat komplementarnie pojmowanej biblioteki cyfrowej, jej tworzenia i rozwoju, udoskonalania oraz implementowania do różnych środowisk naukowych i dla społeczeństwa informacyjnego. Forum połączone jest z prezentacjami rozwiązań

technologicznych, nowych narzędzi informacyjno-komunikacyjnych, stosowanych do organizowania repozytoriów cyfrowych, jak również spotkaniami przedstawicieli Konsorcjum *Polskie Biblioteki Cyfrowe* oraz działających w jego ramach zespołach koordynujących prace instytucji digitalizujących i formujących krajowe zasoby cyfrowe. Wśród zagadnień poruszanych podczas konferencji Polskie Biblioteki Cyfrowe do najważniejszych zaliczono problemy dotyczące technologii, modeli bibliotek cyfrowych, polityki rozwoju, typów repozytoriów i kolekcji cyfrowych, najistotniejszych procesów związanych z implementowaniem zasobów cyfrowych, a także zagadnień związanych z użytkownikami bibliotek cyfrowych oraz zastosowaniu repozytoriów w procesie kształcenia i przygotowywania naukowych zasobów elektronicznych. Najważniejsze kwestie, prezentowane podczas konferencji z lat 2008-2011 przedstawiono w tabeli 5.

Tabela 5. Problematyka bibliotek cyfrowych poruszana podczas konferencji Polskie Biblioteki Cyfrowe za lata 2008-2011 [oprac. autorki].

OBSZAR TEMATYCZNY	**NAJWAŻNIEJSZE ZAGADNIENIA poruszane podczas konferencji *Polskie Biblioteki Cyfrowe* w latach 2008-2011**
TECHNOLOGIA	metadane w bibliotekach cyfrowych
	pozyskiwanie treści dla bibliotek cyfrowych
	systemy OCR
	techniczne aspekty bibliotek cyfrowych
	technologie bibliotek cyfrowych - automatyzacja procesów, systemy organizacji wiedzy, integracja metadanych
	zintegrowana infrastruktura sieciowa wiedzy w Polsce
	zintegrowany system informacji archiwalnej
BIBLIOTEKI CYFROWE - modele, polityka, rodzaje, zasoby cyfrowe	biblioteka cyfrowa dla niewidomych
	biblioteka cyfrowa jako narzędzie animacji współpracy środowiskowej
	Europeana i otwarte dane
	infrastruktura bibliotek cyfrowych w Polsce
	kierunki rozwoju polskich bibliotek cyfrowych
	repozytorium cyfrowe instytutów naukowych
	rozwój bibliotek cyfrowych
	współpraca europejska
PROCESY	digitalizacja
	organizacja bibliotek cyfrowych
UŻYTKOWNICY WYMIAR EDUKACYJNY	aspekt edukacyjny bibliotek cyfrowych, naukowe biblioteki cyfrowe
	biblioteki cyfrowe w pracy naukowej i w dydaktyce
	użytkownicy
	użytkownicy oraz ocena jakości bibliotek cyfrowych
	wizualizacja bibliotek cyfrowych

Inne wydarzenia i konferencje

Na osobną uwagę zasługują spisy wydarzeń, uwzględniające rozliczne konferencje, warsztaty i sympozja organizowane na całym świecie.

Konferencje dotyczące bibliotek cyfrowych i aspektów odnoszących się do tych obiektów, systemów, przez konsorcja projektów finansowanych przez Unię Europejską, lub poszczególne kraje biorące udział w digitalizacji dziedzictwa kulturowego, tworzeniu rozproszonych systemów informacyjno-wyszukiwawczych, platform repozytoryjnych w środowiskach naukowych.

Warto też podkreślić, że polskie i zagraniczne środowiska bibliotekarskie poprzez fachowe stowarzyszenia, federacje, instytucje (np. biblioteki narodowe) organizują zarówno sekcje i odziały w krajowych i międzynarodowych organizacjach zajmujących się problematyką bibliotek cyfrowych, a także realizujących politykę tworzenia standardów, formowania modelowych rozwiązań, określających tzw. dobre praktyki i opracowujące publikacje na temat podstaw wdrażania i tworzenia bibliotek cyfrowych, indywidualnie lub podczas fachowych konferencji naukowych przygotowują sesje w zakresie repozytoriów cyfrowych i digitalizacji. Dotyczy to m.in. *International Federation of Library Associations* IFLA (Międzynarodowej Federacji Stowarzyszeń Bibliotekarskich), opracowującej linki referencyjne, odsyłające do zasobów dotyczących bibliotek cyfrowych[6] (http://archive.ifla.org/II/diglib.htm), a także DL.Org *Digital Library Interoperability, Best Practices and Modelling Foundations* (Fundacja na rzecz Interoperacyjności Bibliotek Cyfrowych, Najlepszych Praktyk i Modelowania, http://www.dlorg.eu/), UNESCO *United Nations Educational, Scientific and Cultural Organization* (Organizacja Narodów Zjednoczonych do Spraw Oświaty, Nauki i Kultury) opracowującego m.in. *World Digital Library* (Światową Bibliotekę Cyfrową, http://www.wdl.org/en/), *Digital Library Federation* wspierającej tworzenie bibliotek cyfrowych (http://www.diglib.org/about/).

Informacje dotyczące wszelkich inicjatyw w dziedzinie bibliotek cyfrowych oraz organizowanych konferencji, forów, warsztatów i szkoleń o zasięgu globalnym można także odnaleźć w czasopismach online zajmujących się w sposób holistyczny, problematyką repozytoriów cyfrowych, takich jak *Ariadne Magazine* (http://www.ariadne.ac.uk/), D-Lib Magazine – Digital Library Magazine (www.dlib.org), International Journal on Digital Libraries. *Quarterly* (link.springer.de/link/service/journals/00799/), *Journal of Electronic Publishing* (www.press.umich.edu/jep/), The National Digital Library Program. *Library of Congress* (lcweb.loc.gov/ndl/per.html), RLG DigiNews (www.rlg.org/toc.html).

[6] Uruchomiono również listę dyskusyjną dla wszystkich bibliotekarzy tworzących i korzystających z bibliotek cyfrowych DIGLIB (diglib@infoserv.inist.fr; http://www.ifla.org/en/mailing-lists).

Podsumowanie

Całkowita liczba kluczowych, znaczących konferencji o charakterze narodowym, jak i ogólnoświatowym jest dość trudna do ustalenia. Przedstawione w artykule międzynarodowe sympozja JCDL, ECDL, RCDL i ICADL są wymiernym wskaźnikiem stanu rozwoju bibliotek cyfrowych i stanowią pierwszoplanowe, doniosłe precedensy w obszarze wymiany wiedzy i doświadczeń w zakresie tworzenia repozytoriów cyfrowych. Ciągła ekspansja prac w zakresie digitalizacji i formowania bibliotek cyfrowych, zainicjowana w latach 90. ubiegłego wieku, rozwój nowych technologii informacyjno-komunikacyjnych, pojawienie się kluczowych międzynarodowych i krajowych dokumentów, wytycznych, określających politykę tworzenia społeczeństwa informacyjnego, cyfryzacji, digitalizacji dziedzictwa kulturowego spowodowała pojawienie się nie tylko nowych rozwiązań w dziedzinie implementacji repozytoriów cyfrowych, projektów, ale także wzrastającą ilość konferencji, forów, warsztatów i szkoleń poświęconych tematyce bibliotek cyfrowych.

Bibliografia

1. Abiteboul, Serge, Vercoustre Anne-Marie red. (1999): *Research and Advanced Technology for Digital Libraries, Third European Conference, ECDL'99, Paris, France, September 22-24, 1999, Proceedings.* „Lecture Notes in Computer Science" 1696 Springer 1999 [dok. elektr.] http://www.informatik.uni-trier.de/~ley/db/conf/ercimdl/ercimdl99.html [odczyt: 20.01.2012].
2. Adam, Nabil R., Bhargava, Bharat K., Halem, Milton Yesha, Yelena red. (1995): *Digital Libraries, Research and Technology Advances, ADL ,95 Forum, McLean, Virginia, USA, May 15-17, 1995, Selected Papers. Lecture Notes in Computer Science 1082 Springer 1996* [dok. elektr.] http://www.informatik.uni-trier.de/~ley/db/conf/adl/adl95.html [odczyt: 20.01.2012].
3. *ACM/IEEE JCDL. ACM/IEEE Joint Conference on Digital Libraries (JCDL)* [dok. elektr.] http://www.informatik.uni-trier.de/~ley/db/conf/jcdl/index.html [odczyt: 15.01.2012].
4. *ADL 1996, DBLP Computer Science Bibliography* (1996). Proceedings of the Third Forum on Research and Technology Advances in Digital Library, ADL ,96, Washington, DC, USA, May 13-15, 1996. IEEE Computer Society 1996 [dok. elektr.] http://www.informatik.uni-trier.de/~ley/db/conf/adl/adl96.html [odczyt: 20.01.2012].
5. *ADL 1997, DBLP Computer Science Bibliography* (1997). 4th International Forum on Research and Technology Advances in Digital Libraries (ADL ,97), Washington, DC, USA, May 7-9, 1997. IEEE Computer Society 1997 http://www.informatik.uni-trier.de/~ley/db/conf/adl/adl97.html [odczyt: 20.01.2012].
6. *ADL 1998, DBLP Computer Science Bibliography* (1998). Proceedings of the IEEE Forum on Reasearch and Technology Advances in Digital Libraries, IEEE ADL ,98, Santa Barbara, California, USA, April 22-24, 1998. IEEE Computer Society 1998 [dok. elektr.] http://www.informatik.uni-trier.de/~ley/db/conf/adl/adl98.html [odczyt: 20.01.2012].

7. *ADL 1999, DBLP Computer Science Bibliography* (1999). Proceedings of the IEEE Forum on Research and Technology Advances in Digital Libraries, Baltimore, Maryland, USA, May 19-21, 1999. IEEE Computer Society 1999 [dok. elektr.] http://www.informatik.uni-trier.de/~ley/db/conf/adl/adl99.html [odczyt: 20.01.2012].
8. *ADL 2000, DBLP Computer Science Bibliography* (2000). Proceedings of IEEE Advances in Digital Libraries 2000 (ADL 2000), Washington, DC, USA, May 22-24, 2000. IEEE Computer Society 2000 [dok. elektr.] http://www.informatik.uni-trier.de/~ley/db/conf/adl/adl2000.html [odczyt: 20.01.2012].
9. *ACM Digital Libraries 1997, DBLP Computer Science Bibliography* (1997). Proceedings of the 2nd ACM International Conference on Digital Libraries, July 25-28, 1997, Philadelphia, PA, USA. ACM 1997 [dok. elektr.] http://www.informatik.uni-trier.de/~ley/db/conf/dl/dl97.html [odczyt: 10.01.2012].
10. *ACM Digital Libraries, DBLP Computer Science Bibliography* (1998) Proceedings of the 3rd ACM International Conference on Digital Libraries, June 23-26, 1998, Pittsburgh, PA, USA. ACM 1998 [dok. elektr.] http://www.informatik.uni-trier.de/~ley/db/conf/dl/dl98.html [odczyt: 15.01.2012].
11. *ACM Digital Libraries, DBLP Computer Science Bibliography* (1999) Proceedings of the Fourth ACM conference on Digital Libraries, August 11-14, 1999, Berkeley, CA, USA. ACM 1999 [dok. elektr.] http://www.informatik.uni-trier.de/~ley/db/conf/dl/dl99.html [odczyt: 15.01.2012].
12. *ACM Digital Libraries, DBLP Computer Science Bibliography* (2000). Proceedings of the Fifth ACM Conference on Digital Libraries, June 2-7, 2000, San Antonio, TX, USA. ACM 2000 [dok. elektr.] http://www.informatik.uni-trier.de/~ley/db/conf/dl/dl2000.html [odczyt: 15.01.2012].
13. *Ariadne. ARIADNE Links Index. Workshops & Conferences* [dok. elektr.] http://www.comp.lancs.ac.uk/computing/research/cseg/projects/ariadne/links.html [odczyt: 28.12.2011].
14. *Artificial Intelligence Lab ICADL* (2010). ICADL International Conference on Asian Digital Libraries [dok. elektr.] http://www.icadl.org/ [odczyt: 17.01.2012].
15. Borbinha, José Luis, Baker, Thomas red. (2000): *Research and Advanced Technology for Digital Libraries, 4th European Conference, ECDL 2000, Lisbon, Portugal, September 18-20, 2000, Proceedings. Lecture Notes in Computer Science* 1923 Springer 2000 [dok. elektr.] http://www.informatik.uni-trier.de/~ley/db/conf/ercimdl/ecdl2000.html [odczyt: 20.01.2012].
16. *CIIR, Center for Intelligent Information Retrieval* (1996) [dok. elektr.] http://ciir.cs.umass.edu/nir96/ [odczyt: 20.12.2011]Candela, Leonardo i in. (2007). Setting the Foundations of Digital Libraries. The DELOS Manifesto. D-Lib Magazine Volume 13 Number 3/4 [dok. elektr.] http://www.dlib.org/dlib/march07/castelli/03castelli.html [odczyt: 09.01.2012].
17. *DELOS a. DELOS events* [dok.elektr.] http://www.delos.info/index.php?option=com_content&task=blogcategory&id=19&Itemid=74 [odczyt: 20.12.2011].
18. *DELOS b. DELOS Conference I. DELOS Conference II* [dok.elektr.] http://www.delos.info/index.php?option=com_content&task=view&id=606&Itemid=337;http://www.delos.info/index.php?option=com_content&task=view&id=524&Itemid=271 [odczyt: 20.01.2012].
19. *DELOS AVIVDiLib. DELOS Newsletter Issue 2* [dok. elektr.] http://www.delos.info/index.php?option=com_content&task=view&id=405&Itemid=192 [odczyt: 29.12.2011].

20. *Digital Libraries 1996, DBLP Computer Science Bibliography* (1996). Proceedings of the 1st ACM International Conference on Digital Libraries, March 20-23, 1996, Bethesda, Maryland, USA. ACM 1996 [dok. elektr.] http://www.informatik.uni-trier.de/~ley/db/conf/dl/dl96.html [odczyt: 12.01.2012].
21. *DLMC 2011. Digital Library Multi-Conference, ISAST International Society for Advanced Science and Technology* [dok.elektr.]. http://www.isast.org/qqml2011/dlmcconference.html [odczyt: 23.01.2012].
22. Fox, Edward A. (1993). *Source Book on Digital Libraries. Technical Report TR-93-35, Computer Science, Virginia Polytechnic Institute and State University* [dok. elektr.] http://fox.cs.vt.edu/DigitalLibrary/DLSB.pdf [odczyt: 12.01.2012].
23. Gross, Ben (2011). *Digital Libraries. Previous Digital Library Conferences* [dok. elektr.] http://bengross.com/dl/#PreviousConferences [odczyt: 12.01.2012].
24. Fujita, Takehisa, University of Tsukuba (1995). *ISDL'95: International Symposium on Digital Libraries 1995* [dok. elektr.] http://www.dl.slis.tsukuba.ac.jp/ISDL95/ [odczyt: 28.12.2011].
25. *ICSD – The International Conferences on Digital Libraries and the Semantic Web 2009* (ICSD2009) [dok. elektr.] http://www.icsd-conference.org/ [odczyt: 17.01.2012].
26. *IJCAI 97* (1997). Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence, IJCAI 97, Nagoya, Japan, August 23-29, 1997. Morgan Kaufmann, 1997, 2 Volumes [dok. elektr.] http://www.informatik.uni-trier.de/~ley/db/conf/ijcai/ijcai97.html [odczyt: 21.01.2012].
27. *ISAST 2011. International Society for Advanced Science and Technology. QQML International Conference on Qualitative and Quantitative Methods in Libraries* [dok. elektr.] http://www.isast.org/qqml2011.html [odczyt: 23.01.2012].
28. *IWWW Conference. International World Wide Web Conferences Committee. Conference history* [dok. elektr.] http://www2010.org/www/about/history/index.html [odczyt: 20.11.2012].
29. *JCDL. Joint Conference on Digital Libraries. Past Events and Conferences* [dok. elektr.] http://www.jcdl.org/past-event-conf.shtml [odczyt: 20.11.2012].
30. *JCDL. Joint Conference on Digital Libraries, George Washington University* (2011). JCDL2012 [dok.elektr.] http://www.jcdl2012.info/ [odczyt: 20.01.2012].
31. *LIDA. Libraries in the Digital Age* (2012). Previous LIDAs [dok. elektr.] http://ozk.unizd.hr/lida/ [odczyt: 20.12.2011].
32. Nikolaou, Christos, Stephanidis, Constantine red. (2011): *Research and Advanced Technology for Digital Libraries, Second European Conference, ECDL ,98, Heraklion*, Crete, Greece, September 21-23, 1998, Proceedings. *Lecture Notes in Computer Science* 1513 Springer 1998 [dok. elektr.] http://www.informatik.uni-trier.de/~ley/db/conf/ercimdl/ercimdl98.html [odczyt: 20.01.2012].
33. Peters, Carol, Thanos, Costantino red. (1997): Research and Advanced Technology for Digital Libraries. First European Conference, ECDL ,97, Pisa, Italy, 1-3 September, Proceedings. *Lecture Notes in Computer Science* 1324 Springer 1997 [dok.elektr.] http://www.informatik.uni-trier.de/~ley/db/conf/ercimdl/ercimdl97.html [odczyt: 20.01.2012].
34. *RCDL – Digital Libraries: Advanced Methods and Technologies, National Russian Research Conference* [dok. elektr.] http://rcdl.ru/section.php?id=13&plang=e [odczyt: 17.01.2012].
35. *RCDL 2003. Digital Libraries: Advanced Methods and Technologies, National Russian Research Conference. RCDL 2003.* [dok. elektr.] http://rcdl.ru/section.php?id=9 [odczyt: 17.01.2012].

36. *RCDL 2005. Digital Libraries: Advanced Methods and Technologies, National Russian Research Conference. RCDL 2005.* [dok. elektr.] http://rcdl.ru/section.php?id=7 [odczyt: 17.01.2012].
37. *Royal School of LIS. Royal School of Library and Information Science. CoLIS 8* [dok.elektr.] http://www.iva.dk/english/colis8/ [odczyt: 11.01.2012].
38. *TPDL2011. International Conference on Theory and Practice of Digital Libraries 2011* [dok. elektr.] http://www.tpdl2011.org/ [odczyt: 11.01.2012].
39. Schnase, John L., Leggett ,John J., Furuta, Richard K., Metcalfe, Ted (1994). *Proceedings of Digital Libraries ,94. The First Annual Conference on the Theory and Practice of Digital Libraries* [dok. elektr.] http://jcdl.org/archived-conf-sites/dl94/ [odczyt: 20.01.2012].
40. Shipman, Frank (1995). *Digital Libraries ,95. The Second Annual Conference on the Theory and Practice of Digital Libraries* [dok. elektr.] http://www.csdl.tamu.edu/DL95/ [odczyt: 20.01.2012].
41. *UKOLN, The United Kingdom Office for Library and Information Networking* (1998). Beyond the Beginning: The Global Digital Library [dok.elektr.] http://old.cni.org/regconfs/1997/uk-content.html [odczyt: 20.12.2011].
42. University of Queensland (2010). *JOINT JCDL/ICADL International Digital Libraries Conference, Program schedule* [dok. elektr.] http://www.jcdl-icadl2010.org/schedule.php [odczyt: 29.12.2011].
43. University of Virginia Library, Library of Congress (1994). *Proceedings of the Seminar on Cataloging Digital Documents. Seminar on Cataloging Digital Documents University of Virginia Library and Library of Congress*, October 12-14, 1994 [dok.elektr.] http://lcweb.loc.gov/catdir/semdigdocs/seminar.html [odczyt: 2.10.2011].
44. Werla, Marcin (2010). *Polskie biblioteki cyfrowe, FBC i Europeana – etapy i bariery w przepływie informacji.* [W:] Bożena Bednarek-Michalska (red.) „Biuletyn EBIB" 1/2010 (110) luty. [Warszawa] : Stowarzyszenie Bibliotekarzy Polskich KWE, 2010. – [dok. elektr.] http://www.ebib.info/2010/110/a.php?werla [odczyt: 17.01.2012].
45. Vrana, Radovan, University of Zagreb Faculty of Philosophy Department of Information Sciences (2002). *CoLIS 3* [dok.elektr.] http://www.ffzg.unizg.hr/infoz/colis3/ [odczyt: 20.01.2012].
46. *VLDL 2010. Workshop on Very Large Digital Libraries* [dok. elektr.] http://www.delos.info/vldl2010/index.html [odczyt: 20.01.2012].

Biblioteki cyfrowe w pracy naukowca humanisty

Mirosław Górny

Biblioteki cyfrowe jako następczynie bibliotek tradycyjnych znacznie skróciły czas potrzebny badaczowi na dostęp do piśmiennictwa. Umożliwiły też wieloaspektową analizę tekstu. W artykule wspomniano o możliwościach lingwistyki korpusowej i technik typu 'content mining'. Podano propozycje sposobów wykorzystania zasobów bibliotek cyfrowych w różnych dziedzinach humanistyki i nauk społecznych.

Wstęp

Biblioteki cyfrowe są dzisiaj jednym z najbardziej docenianych narzędzi pracy tych naukowców, którzy ze zbiorów bibliotecznych muszą intensywnie korzystać.

Digitalizacja umożliwiła naukowcom dostęp do ogromnej liczby naukowych monografii i źródeł. Oczywiście część publikacji jest dostępna bez żadnych opłat, część udostępniana jest wyłącznie w systemach komercyjnych. Tak czy owak naukowcy dysponują obecnie milionami tekstów dostępnych **natychmiast i równocześnie.**

Tylko w polskich bibliotekach cyfrowych i repozytoriach dostępne jest obecnie ok. 700 tys. obiektów cyfrowych – zdecydowana większość tych pozycji może być czytana przez każdego w dowolnym czasie i w dowolnym miejscu gdzie dostępny jest Internet. Są to książki, czasopisma, gazety, księgi adresowe, starodruki, fotografie i pocztówki. Czyli wszystko to co można znaleźć w naukowej bibliotece.

Jest oczywiste, że zasób taki wzbudza ogromne zainteresowanie badaczy.

Obecnie ponad 17% wszystkich stałych (czyli takich którzy przynajmniej kilka razy w roku korzystają biblioteki) użytkowników polskich bibliotek cyfrowych to pracownicy naukowi. Szacuje się, że jest to ok. 5 tys. osób. Oczywiście w przeważającej większości są to przedstawiciele nauk humanistycznych.

Ok. 7% pracowników naukowych korzystających z bibliotek cyfrowych to osoby między 61 a 70 rokiem życia, 20% to osoby między 51 a 60 rokiem życia, 20% to

osoby między 41 a 50 rokiem życia, 28% to osoby między 31 a 40 rokiem życia, a 21% to osoby między 26 a 30 rokiem życia.

Ok. 13% z nich korzysta z bibliotek codziennie, 34% kilka razy w tygodniu, 40% kilka razy w miesiącu. Czyli prawie 90% pracowników naukowych korzystających z bibliotek cyfrowych wykorzystuje je stosunkowo intensywnie, traktując je jako jeden ze swoich warsztatów pracy[1] [Mazurek 2009].

W podobnych badaniach przeprowadzonych wśród pracowników naukowych hinduskiego uniwersytetu Manonmaniam Sundaranar w Tirunelveli uzyskano dane z których wynika, że codziennie korzysta z bibliotek cyfrowych prawie 18% badanych, 32% trzy razy w tygodniu, raz w tygodniu ponad 27% i ok. 13% raz w miesiącu. Dane te są zaskakująco zbieżne z wynikami polskimi [Thanuskodi; Ravi 2011].

Biblioteki cyfrowe znacznie ułatwiły pracę naukowców. Dostęp poprzez sieć do potrzebnych badaczowi publikacji oszczędza czas w stopniu niebywałym. I jest to obecnie główna i niezaprzeczalna zaleta masowej digitalizacji piśmiennictwa wykorzystywanego do celów naukowych.

Ale nie jest to jedyna zaleta. Możliwości biblioteki cyfrowej to również zautomatyzowane wyszukiwanie i odpowiednie grupowanie wybranych elementów, zarówno na poziomie jednego dokumentu jak i grupy dokumentów.

Oznacza to że badacz w stosunkowo krótkim czasie uzyskuje dostęp do zbioru potrzebnych publikacji, a nawet do zbioru potrzebnych mu fragmentów tekstów. W przypadku korzystania z naukowych opracowań w postaci monografii czy artykułów w czasopismach wydaje się być to rozwiązanie całkowicie wystarczające.

Należy jednak pamiętać że w niektórych dziedzinach liczba ukazujących się publikacji jest tak ogromna, że pojawiło się zapotrzebowanie na automatyczne analizowanie treści tychże.

Inaczej mówiąc antropolog czy historyk znajdujący kilkanaście książek i kilkadziesiąt artykułów na interesujący go temat przeczyta je czy choćby przejrzy bez pomocy dodatkowych narzędzi. Ważne że może zrobić to nie ruszając się zza biurka i mając wszystkie publikacje równocześnie dostępne. Ale np. biochemik, który uzyskuje natychmiastowy dostęp do kilkuset artykułów będzie oczekiwał programu, który przynajmniej po części wykona za niego żmudną pracę przeglądania.

Także ci naukowcy dla których teksty to nie tylko publikacje naukowe, ale także źródła do badań, będą poszukiwali narzędzi do szybkiego analizowania zasobów cyfrowych.

Takim rozwiązaniem jest oprogramowanie typu „text mining" (coraz częściej określane jako „content mining") - stosowane obecnie głównie w naukach biologicznych i medycznych. Jednak w pozostałych dziedzinach nauki zainteresowanie tymi technikami jest coraz większe [Smit; van der Graaf 2011].

[1] Dane na podstawie wyników z badań ankietowych przeprowadzonych w 2009 r. wśród użytkowników polskich bibliotek cyfrowych.

Biblioteka cyfrowa – ulepszona wersja biblioteki tradycyjnej

Bibliotekę cyfrową można postrzegać jako zbiór wyrazów i fraz rozmieszczonych w różnych dokumentach i w różnych ich miejscach. Publikacje są zwykle datowane (np. gazety bardzo dokładnie) i opatrywane informacją dotyczącą miejsca wydania.

Dzięki temu dla każdego wybranego obiektu (wyrazu lub frazy) uzyskujemy następujące informacje:

- data wystąpienia
- kontekst wystąpienia
- miejsce publikowania

Połączenie tych informacji daje swego rodzaju mapę, która dla danego obiektu pozwala obserwować skupiska jego wystąpień w powiązaniu z datą wystąpienia i miejscem. Mierzalną cechą skupiska jest oczywiście jego liczebność. Mierzalną cechą pojedynczego obiektu jest jego odległość od innych obiektów.

W rezultacie badacz otrzymuje dla dowolnego wyrazu, zbitki wyrazów bądź frazy, ciąg skupisk o ustalonej liczebności i zwykle precyzyjnie datowanych. Na podstawie miejsca wydania może (choć nie zawsze) przypisać badany obiekt do konkretnego regionu kulturowego. Pozostaje już teraz tylko interpretacja tego rodzaju danych.

Dla historyka dane tak uzyskane, mogą pozwolić na ustalenie kiedy pojawił się interesujący go wątek np. w prasie danego regionu i kiedy przestał się pojawiać. Pozawalają te dane również śledzić intensywność obecności danego wątku w publikacjach i ewentualną zmienność jego występowania w zależności od obszaru publikowania i czasu.

Naturalnie interpretacji podlegać może również stopień korelacji współwystępowania ciągów skupisk innych obiektów.

Podobne analizy można przeprowadzać dwojako – ręcznie albo wykorzystując wyspecjalizowane programy.

Ręczna metoda – rzecz jasna pracochłonna, choć nie wymagająca dodatkowych nakładów – w praktyce ma dość ograniczone znaczenie. A to z tego powodu, że badacz może przeanalizować w zasadzie tylko niewielki objętościowo materiał. Będzie ona stosowana zwykle wtedy kiedy rzecz dotyczy pojedynczych obiektów.

Jeśli np. historyka interesuje kontekst wypowiedzi dotyczącej jakiejś kwestii np. gospodarczej czy politycznej, to naturalnie posłuży się dostępnym dla czytelnika interfejsem, aby szybko odnaleźć odpowiednie fragmenty. Przygotowywanie odpowiedniego oprogramowania nie ma tu większego sensu.

Istnieją co prawda biblioteki cyfrowe umożliwiające badaczowi (dotyczy to głównie historyków) bardziej efektywne analizowanie treści niż to jest w przypadku wykorzystania zwykłego interfejsu użytkownika.

Np. projekt *Chronicling America* pozwala na bardzo szybkie odnajdywanie stron gazet, na których występują słowa i frazy interesujące badacza. Otrzymuje on w wy-

niku tylko te strony gazet na których znajdują się szukane obiekty. Wszystkie one są oczywiście odpowiednio oznaczone[2] [About Chronicling America].

To bardzo proste i wyjątkowo sprawne narzędzie w dużym stopniu ułatwiło pracę historykom.

Jednak należy pamiętać, że dla historyków sprawne narzędzia do przeszukiwania tekstu – aczkolwiek zawsze bardzo przydatne – ma często mniejsze znaczenie od możliwości przeczytania dokumentu w ogóle i to „obejrzenia" go w postaci możliwie najbardziej przypominającej oryginał. Stąd zapewne biblioteka cyfrowa bogata w zbiory i udostępniająca skany wysokiej jakości będzie cieszyła się wśród historyków powodzeniem, nawet jeśli nie będzie posiadała specjalnych narzędzi do analizy treści.

Choć należy się liczyć z tym że nadchodzące pokolenia historyków będą takich narzędzi się powszechnie domagać.

Tu warto przytoczyć wypowiedź historyka A. Maxwella z Nowej Zelandii (reprezentuje pokolenie między 40 a 50 rokiem życia) dotyczącą bibliotek cyfrowych.

Twierdzi on, że dla historyka najważniejszą funkcją biblioteki jest chronić oryginał. A jeśli udostępnia się jego skan to powinien być to w miarę możliwości najwierniejszy jego obraz.

Nie omieszkuje on jednak ponarzekać na interfejsy różnych bibliotek cyfrowych. M.in. krytycznie ocenia interfejsy Galliki i Europeany (w przypadku tej ostatniej wyraża się wręcz, komentując wypowiedzi jej twórców: „przechwałki bez pokrycia"). Ciekawe, że pozytywnie pisze o polskich bibliotekach cyfrowych, aczkolwiek krytykuje konieczność instalowania dodatkowego oprogramowania (chodzi o plugin DjVu). Z artykułu wynika jednak, że z bibliotek tych nie korzystał, a swoją wiedzę o nich uzyskał ze stron w internecie.

Natomiast interfejs Google Books ocenia jako znakomity i bezkonkurencyjny, choć zauważa że przeszukiwanie tekstu książek pisanych gotykiem bądź cyrylicą daje czasem błędne wyniki spowodowane problemami programów OCR z nietypowymi krojami czcionek. Jednakże uważa, że Google Books powinien stanowić swego rodzaju standard dla informacji cyfrowej[3].

I bardzo ważna konkluzja na koniec: wg Maxwella łatwość wyszukiwania w bibliotece cyfrowej nie jest specjalnie istotna, prawdziwym problemem jest łatwość przeglądania [Maxwell 2011].

[2] Chronicling America jest projektem realizowanym w ramach National Digital Newspaper Program przez National Endowment for the Humanities i Bibliotekę Kongresu, która ostatecznie będzie instytucją rozwijającą i utrzymującą to przedsięwzięcie.

[3] Trudno Autorowi nie przyznać racji. Google Books pozwala błyskawicznie wyodrębnić na podstawie wybranych przez czytelnika elementów wyszukiwawczych, względnie samodzielną bibliotekę poświęconą jakiejś dziedzinie (zakładamy że czytelnik wybiera tylko te pozycje, które udostępniono w całości). Treść książek daje się łatwo przeszukiwać i pobierać w formacie PDF. A wszystko działa niebywale szybko i sprawnie.

Biblioteka cyfrowa – nowe narzędzie badawcze

Jak już wspomniano w bibliotekach cyfrowych szuka się nie tylko opracowań naukowych, ale także źródeł do badań. I to badań wszelkiego rodzaju – językoznawczych, historycznych, filologicznych, politologicznych itp. Badania te związane są często z pracochłonną analizą dużych zasobów piśmiennictwa.

Przykładem takiej analizy zakrojonej na niespotykaną dotąd skalę jest projekt zrealizowany przez grupę naukowców z Uniwersytetu Harvarda i laboratoriów Google. Objął on ogromny – liczący prawie 5 mln 200 tys. książek – cyfrowy zasób. Wg szacunków autorów projektu jest to ok. 4% wszystkich dotąd opublikowanych na świecie książek. Książki wybrano spośród ok. 15 mln dotychczas zdigitalizowanych, kierując się przede wszystkim jakością procesu OCR. Najstarsze pozycje w tym zbiorze pochodzą z XVI wieku.

Zasób został przekształcony w korpus językowy liczący ok. 500 mld słów. 361 mld to słowa w języku angielskim, 45 mld to słowa w języku francuskim, tyle samo w hiszpańskim, 37 mld w niemieckim, 13 mld w chińskim, 35 mld w rosyjskim i 2 mld w języku hebrajskim.

Najstarsza część korpusu (XVI – XVIII w.) zawiera po kilkaset tysięcy słów przypadających na każdy rok, część z XIX w. to ok. 60 mln słów rocznie, część z XX w. to 1,4 mld słów rocznie i część z XXI w. to ok. 8 mld słów rocznie [Jean-Baptiste 2011].

Celem analizy było ustalenie częstotliwości występowania danego słowa w wybranym roku. W konsekwencji pozwoliło to uchwycić tendencje częstotliwości występowania i współwystępowania określonych słów. Oczywiście tak uzyskane wyniki same w sobie nic nie znaczą. Dopiero właściwa ich interpretacja pozwala odkrywać, potwierdzać czy uściślać przebieg określonych zjawisk kulturowych.

Ta technika, która związana jest z pojęciem „kulturonomii" (culturonomics) – czyli dziedziną stosującą analizowanie danych „o wysokiej wydajności" (high-throughput data) – daje się wykorzystać w wielu naukach społecznych. M.in. w językoznawstwie, historii, naukach politycznych, kulturoznawstwie itp.

Idea nie jest nowa (Bo przecież chociażby badania bibliometryczne prowadzi się od kilkudziesięciu lat). Nowe jest niezwykle wydajne narzędzie, którym do czasu digitalizacji ogromnych zasobów piśmienniczych nie dysponowano.

Na gruncie polskim prekursorem wykorzystania zasobów bibliotek cyfrowych do badań filologicznych jest Piotr Wierzchoń. Od kilku lat przeszukuje on teksty polskich gazet i czasopism wydawanych w okresie międzywojennym i w XIX wieku, aby ustalać najwcześniejsze wystąpienia określonych słów i konstrukcji leksykograficznych.

Istotne jest to, że Wierzchoń analizuje cały zasób polskiego czasopiśmiennictwa dostępnego w 60 polskich bibliotekach cyfrowych korzystając z narzędzi informatycznych. Dzięki temu jest w stanie przeszukać ogromny materiał w stosunkowo krótkim czasie i zaprezentować wyniki w postaci obrazów fragmentów tekstu zawierających odnalezione wystąpienia.

Taki sposób prezentacji wyników ma ogromne znaczenie dla badaczy. Pokazuje nie tylko samo wystąpienie słowa, precyzyjnie datowane (wszak są to periodyki, nierzadko gazety codzienne), nie tylko wskazuje na miejsce użycia danego terminu (gazety są przypisane często do określonego regionu), ale pozwala też wyciągać ewentualne wnioski z kontekstu. Nie mówiąc już o cechach typograficznych, które czasami dostarczają bardzo ważnych dla badacza informacji. Np. interesujące go słowo może być zapisane odmienną czcionką, może być wzięte w cudzysłów itp. co może wskazywać np. na to, że było jeszcze wtedy traktowane jako neologizm albo nie było uznawane za słowo pełnoprawne w języku polskim [Wierzchoń 2008a].

Przytoczone powyżej przykłady analizy języka polskiego stanowią zapowiedź nowej ery w badaniach filologicznych i językoznawczych. Nawet ten zebrany w stosunkowo prosty sposób materiał daje już ogromne możliwości interpretacyjne.

Ale przecież zarówno powiększające się stale zasoby bibliotek cyfrowych, jak i rozwój wyspecjalizowanych narzędzi informatycznych, pozwolą wkrótce na prowadzenie szerokich badań nad cyfrowymi tekstami. Na razie badania zasobu tekstów cyfrowych wymagają pewnych umiejętności informatycznych, co być może zniechęca wielu badaczy.

Sytuację radykalnie zmieni zapewne pojawienie się odpowiedniego oprogramowania. Niewykluczone też że i sam zasób cyfrowy zostanie udostępniony w formie ułatwiającej wszelkiego rodzaju analizy.

Wydaje się, że jedną z takich form zasobu cyfrowego, przydatnych naukowcom (szczególnie dotyczy to nauk filologicznych) są korpusy językowe tworzone na podstawie zasobów bibliotek cyfrowych.

Językoznawstwo korpusowe jest obecnie bardzo szybko rozwijającą się dziedziną. Doświadczenia tej dziedziny, stworzone zasoby w postaci korpusów językowych, a przede wszystkim techniki pracy z korpusami i programy do ich analizy pozwalają obecnie na bardzo szerokie wykorzystanie zasobów bibliotek cyfrowych. Wszak ich zawartość po odpowiednim przygotowaniu jest niczym innym jak właśnie korpusem językowym.

Pomijając oczywiście nieuniknione trudności wynikające z błędnej pracy programów OCR (niestety w bibliotekach cyfrowych znaczna część piśmiennictwa to gazety, dość często drukowane czcionką niskiej jakości, pełne reklam i ogłoszeń obfitujących w ramki, rysunki i nietypowe czcionki) możliwe jest uzyskanie ogromnego materiału przydatnego np. do badań nad językiem.

Bowiem czasopiśmiennictwo (a szczególnie gazety codzienne) pozwala na wyodrębnienie mikrokorpusów wytworzonych na określonym terytorium i w określonym czasie. To naturalnie umożliwia zastosowanie istniejących obecnie metod i narzędzi do badania np. ewolucji języka. Porównywanie tak utworzonych korpusów językowych pozwala stwierdzić nie tylko różnice między językiem używanym na poszczególnych obszarach i w danym czasie, ale przede wszystkim pozwala obserwować oddziaływanie języka danego regionu na inne obszary. Rzecz jasna same operacje statystyczne na korpusach dostarczają tylko danych, które wymagają interpretacji. Ale uzyskanie list częstotliwościowych wybranych słów, wyrażeń i fraz występujących np. na obszarze

Wielkopolski w poszczególnych latach i porównanie ich z podobnymi listami uzyskanymi na podstawie analizy czasopiśmiennictwa np. kujawsko-pomorskiego może stanowić niezwykle ciekawy materiał do językoznawczych analiz.

Warunkiem jest tutaj dysponowanie odpowiednio licznym zasobem zdigitalizowanych gazet (w jakimś stopniu już to zostało uzyskane, a zasób jest nadal przecież uzupełniany). Wszak kilkadziesiąt tytułów gazet z jednego obszaru kulturowego, wydawanych w ciągu powiedzmy 10 lat daje całkiem solidne podstawy do twierdzenia, że dość wiarygodnie został „odtworzony" sam język i proces jego zmian zachodzący na tym obszarze. Co prawda jest to tylko „język gazetowy", ale jest on zwykle lepszym odbiciem języka mówionego, niż język występujący w książkach. Poza tym gazety w związku ze swoją specyfiką traktują często o tematyce silniej związanej z życiem codziennym niż książki.

Porównanie takich zasobów językowych dla różnych obszarów w różnym czasie daje szanse na obserwowanie kulturowego oddziaływanie obszarów między sobą.

Mówiliśmy tutaj o problematyce językoznawczej – z oczywistych względów materiał pozyskany z bibliotek cyfrowych najlepiej się do badań filologicznych i językoznawczych nadaje.

Ale podobne metody mogą być wykorzystane np. przez historyków. Nawet analiza występujących w prasie nazwisk, ogłoszeń dotyczących poszukiwania pracy, pracowników, kupna i sprzedaży itp. może dać zaskakująco ciekawe wyniki.

Konkluzje

Biblioteki cyfrowe nie zastąpiły bibliotek tradycyjnych, jeśli idzie o pracę naukowca humanisty. Można jednak powiedzieć, że „w bardzo dużym stopniu zastąpiły". Co więcej. ułatwiły znacznie, a w niektórych przypadkach wręcz umożliwiły pewnego typu badania.

Ich głównym zadaniem będzie nadal umożliwianie klasycznej analizy źródeł, udostępnianie opracowań i monografii. Tyle tylko że o wiele mniej czasochłonne niż w przypadku bibliotek tradycyjnych.

Nowością jeśli idzie o wykorzystania bibliotek cyfrowych są wysoce efektywne, zautomatyzowane techniki analizy treści. Tutaj coraz większą rolę odgrywa lingwistyka korpusowa i techniki określane jako 'text mining'. Te ostatnie stosowane już szeroko w naukach przyrodniczych, zapewne już niedługo będą powszechnie wykorzystywane w naukach humanistycznych [Fox 2010]. Choć oczywiście natura obiektów którymi zajmują się te nauki jest inna. I stąd problem jest tu znacznie bardziej złożony niż np. w biologii czy chemii.

Można też przypuszczać, że naukowcy humaniści zwrócą uwagę na inną jeszcze możliwość którą dają biblioteki cyfrowe. Chodzi o badanie 'właściwości tekstu'. Pomijając tu badania językoznawcze gdzie takie badania są oczywiste (np. w lingwistyce tekstu), to wydaje się że możliwe jest na tej podstawie uzyskanie przydatnych wyników również w historii czy antropologii.

Obecnie analizuje się tekst głównie pod kątem odnajdywania faktów, opisów zdarzeń, idei, koncepcji itd. To jest poziom najniższy jeśli idzie o stopień trudności. Tekst jest tutaj jedynie swego rodzaju nośnikiem wspomnianych obiektów. Można powiedzieć że wykorzystuje się tu zbiór oznaczonych (datowanych, przypisanych do danego regionu i do określonych grup społecznych i instytucji) źródeł jako nośników informacji. Ewentualne powiązania między odnalezionymi obiektami są ustalane przez badacza.

Ale analiza tekstów daje już możliwość automatycznego ustalania pewnych konfiguracji faktów czy zdarzeń. Tutaj powiązania mogą być wykrywane przez odpowiednie oprogramowanie. Tekst pozostaje tu jednak nadal tylko nośnikiem sieci informacji.

Ale nie można wykluczyć w przyszłości badań nad samym tekstem, ściślej nad jego właściwościami. Struktura tekstu, jego objętość, występowanie podobieństw, powiązania tematyczne, zakres i dynamika zmian stosowanego słownictwa, konstrukcji gramatycznych itp. mogą być może stanowić cenny materiał do kompleksowych badań nad historią i kulturą danego okresu i danego regionu.

Dzisiaj biblioteki cyfrowe w pewnym stopniu nadal przypominają swoje poprzedniczki – biblioteki gromadzące piśmiennictwo drukowane. Ale stopniowo będą w coraz większym stopniu przygotowywane do zadań wykraczających poza zwykłe przeglądanie i czytanie poszczególnych publikacji.

Bibliografia

1. *About Chronicling America* [dok. elektr.] http://chroniclingamerica.loc.gov/about/ [odczyt: 1.06.2011].
2. Fox, Robert (2010). *Mining the digital library.* „OCLC Systems and Services" vol. 26 nr 4, s. 232-238.
3. Jean-Baptiste, Michel [i in.] (2010). *Quantitative analysis of culture using millions of digitized books.* [dok. elektr.] http://www.librarian.net/wp-content/uploads/science-google-labs.pdf [odczyt: 1.06.2011].
4. Wierzchoń, Piotr (2008). *Kotuś „verba Polona Ascondita..." w fotodokumentacji: szkic lingwochronologizacyjny.* Poznań: Instytut Językoznawstwa UAM, 136 s.
5. Maxwell, Alexander (2010). *Digital archives and history research: feedback from an end-user.* „Library Review" vol. 59 nr 1, s. 24-39.
6. Mazurek, Jolanta (2009). *Drogi czytelniku, jak korzystasz z biblioteki cyfrowej? Próba odpowiedzi* [dok. elektr.] „EBIB – Elektroniczny Biuletyn Bibliotekarzy" nr 5/105 http://www.ebib.info/2009/105/a.php?mazurek [odczyt: 2.12.2011].
7. Smit, Efke; van der Graaf, Maurits (2011). *Journal article mining: a research study into practices, policies, plans... and promises* [dok. elektr.] http://www.publishingresearch.net/documents/PRCSmitJAMreport20June2011VersionofRecord.pdf [odczyt: 21.06.2011].
8. Thanuskodi, S.; Ravi, S. (2011). *Use of digital resources by faculty and research scholars of Manonmaniam Sundaranar University*, Tirunelveli. „DESIDOC Journal of Library & Information Technology" vol. 31 nr 1, s. 25-30.

9. Wierzchoń, Piotr (2008a). *Fotodokumentacja, chronologia, emendacja: teoria i praktyka weryfikacji materiału leksykalnego w badaniach lingwistycznych*. Poznań: Instytut Językoznawstwa, 494 s.
10. Wierzchoń, Piotr (2008b). *Jaskółki przejawów internacjonalizacji w słowotwórstwie współczesnej polszczyzny w materiałach z lat 1894-1984: tylko sto przykładów*. Łask: Oficyna Wydawnicza Leksem, 100 s.

Aneks

1. Przekształcenie piśmiennictwa z danego obszaru i z danego okresu w uporządkowaną listę wyrazów
2. Porównanie list z kolejnych okresów
3. Poddanie interpretacji części list, które nie pokrywają się (nie są częścią wspólną)

1. Ogólny schemat analizy mikrokorpusów językowych pomagający wykrywać neologizmy

1. Przekształcenie piśmiennictwa z danego obszaru i z danego okresu w uporządkowaną listę wyrazów z podaniem częstotliwości ich występowania
2. Porównanie list z różnych obszarów kulturowych i kolejnych okresów
3. Interpretacja części wspólnej list z uwzględnieniem częstotliwości występowania wybranych wyrazów

2. Ogólny schemat analizy mikrokorpusów językowych pomagający ustalić stopień i zakres oddziaływań językowych między poszczególnymi regionami kulturowymi

Komentarz: Przedstawione tutaj schematy nie obejmują oczywiście licznych szczegółów procesu badawczego. W trakcie opisanego postępowania badacz otrzyma bardzo „surowy" materiał, który zapewne będzie musiał być poddany dodatkowej obróbce. Niemniej taka stosunkowo prosta – jeśli idzie o koncepcję – metoda „masowego odsiewania tekstu " pozwala na dosyć efektywne ustalenie zarówno czasu jak i miejsca pojawienia się wielu neologizmów formalnych. Pozwala też uzyskać dane mogące posłużyć do wnioskowania o wpływie określonych ośrodków kulturalnych na poszczególne regiony. Tymi danymi są wybrane obiekty pojawiające się z różną częstotliwością w poszczególnych mikrokorpusach. Trzeba jednak pamiętać, że występowanie jakiegoś słowa na liście utworzonej na podstawie piśmiennictwa z regionu A z częstotliwością większą niż na liście utworzonej na podstawie piśmiennictwa z regionu B z nieco późniejszego okresu (o ile późniejszego to już jest problem badacza) nie musi automatycznie oznaczać że region A oddziaływał na region B. Stale podkreślamy, że to są tylko dane, które muszą być odpowiednio przez badaczy interpretowane.

Ocena jakości bibliotek cyfrowych

Małgorzata Janiak, Monika Krakowska

1. Wstęp

W opracowaniu przedstawione zostały kryteria oceny jakości bibliotek cyfrowych, omówione na podstawie literatury polskiej i zagranicznej dotyczącej tego zagadnienia. Problematykę oceny jakości zaprezentowano poprzez analizę pól badawczych (elementów branych pod uwagę przez autorów omawianych publikacji), podstawowe cele analiz oraz ich zakresy. Autorki przedstawiły wiele propozycji kryteriów oceny jakości, które zaproponowane zostały w piśmiennictwie polskim i zagranicznym. Probierze te odnoszą się do ewaluacji systemu informacyjnego, jego zawartości, funkcjonalności, lub też wykorzystywania go przez użytkowników. Głównym powodem podjęcia decyzji o przedstawieniu wielu zestawów było przekonanie, iż nie trzeba wypracowywać jednej metody badań jakościowych. Metody i techniki badawcze zależą bowiem od podstawowych elementów: 1) co tak naprawdę badamy, 2) jak, 3) dlaczego, 4) kiedy, 5) komu i czemu mają służyć wyniki, 6) czy, a jeśli tak, to co nowego mogą wnieść badania oraz 6) w jakim paradygmacie funkcjonują badacze, a w jakim twórcy oraz użytkownicy.

Biblioteka cyfrowa rozumiana jest w tej publikacji jako rodzaj systemu informacyjnego, składający się z 4 elementów: samej kolekcji (stanowiącej zbiór danych, o bardzo różnorodnej tematyce, jak i formie), komputerowego systemu informatycznego, oferującego przechowywanie, modyfikację oraz dostęp do danych (infrastruktura sprzętowo-programowa), ludzi (zarówno twórców, jak i odbiorców, chociaż to użytkownik jest tu najważniejszym ogniwem) oraz projektów, a więc powodów, celów, całej infrastruktury dla jakiej te zasoby zostały stworzone [Fuhr 2007; Janiak; Krakowska 2010b]. Popatrzono więc na system informacyjny, jak na system logiczny, ale też techniczny oraz społeczny. Jest to najbardziej holistyczne spojrzenie, które powoduje, iż analiza jakościowa zawierać powinna całą *eksplikację informacyjną, socjologiczną oraz psychologiczną: zapotrzebowania na takie, a nie inne zbiory, celowości ich*

powstania, oczekiwań użytkowników, powodów wykorzystania oraz niewykorzystywania takich [jak biblioteki cyfrowe] *źródeł informacji, dynamiki ich rozwoju, planów na przyszłość itp.* [Janiak; Krakowska 2010a].

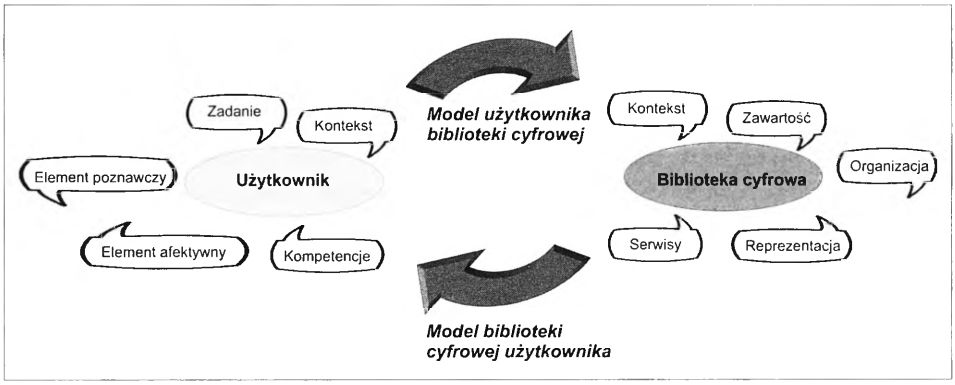
W piśmiennictwie polskim oraz zagranicznym odnaleźć można po pierwsze: pozycje które proponują koncepcje oceny, modele, podejścia oraz ogólną metodologię, po drugie: publikacje przedstawiające dokonane oceny konkretnych bibliotek cyfrowych (dane, wyniki, impresje, itp.) [Saracevic 2004].

W przypadku analiz metodologicznych, najbardziej holistyczny opis typologiczny zaproponowano w trakcie prac dla sieci doskonałości dla bibliotek cyfrowych (*Network of Excellence on Digital Libraries*) DELOS. Wydzielono cztery metody oceny:

1) ewaluację formatywną (kształtującą), którą rozpocząć należy w początkowej fazie rozwoju projektu prowadzącego do stworzenia biblioteki cyfrowej (implementacja systemu i zasobu), aby ustalić bieżącą strategię działań i zadań; ocena taka wyznacza cele, pomaga określić rezultaty. Prowadzona jest na zasadzie monitoringu, wyznaczana przez kontekst i cele danego projektu;
2) ocenę podsumowującą (końcową), która ma ustalić, czy zamierzone intencje zostały efektywnie spełnione; w jej trakcie porównuje się założenia początkowe ze zrealizowanym projektem, stąd zwykle jest ona poprzedzona ewaluacją formatywną;
3) ewaluację iteracyjną (wielokrotną, powtarzalną), która odbywa się w trakcie trwania całego projektu. Kolejne stadia przejściowe oceniane są w odniesieniu do planowanych celów i efektów, a wyniki tej ewaluacji decydują o realizacji następnych etapów implementowania projektu;
4) ewaluację porównawczą (komparatywną), która wymaga ustandaryzowanych kryteriów pomiaru. Według nich oceniane są procesy, systemy, zasoby, zachowania użytkowników itp. [Fourth DELOS Workshop 2002].

Tefko Sarcevic, współpracujący także przy tworzeniu analiz dla DELOS, podkreślił jeszcze, iż poza podziałem na typy ewaluacji, ważne jest określenie rodzaju podejścia do przeprowadzenia ewaluacji biblioteki cyfrowej: czy jest ono np. ukierunkowane na system, jego użyteczność, czy też jest to spojrzenie antropologiczne, socjologiczne, etnograficzne, lub ekonomiczne. Można oczywiście przeprowadzić analizy jakościowe, które łączą kilka sposobów pojmowania i oceniania biblioteki cyfrowej. Należy jednak podkreślić, iż jedno wybrane, konkretne podejście nie może prowadzić do ogólnych wniosków. Każda ocena jest zazwyczaj czasowa, a porównywanie ocen dokonanych na kilku bibliotekach powinno spełnić warunki: podobieństwa kryteriów, podejść, czasu itp.

Najbardziej podstawowe, wymienione przez Tefko Saracevica oraz Ying Zhang [2007], dwa paradygmaty odnoszące się do analiz jakości, to dwa 'centryzmy': skierowanie na użytkownika i na bibliotekę cyfrową. Najważniejsze aspekty brane pod uwagę w trakcie badań zorientowanych na czytelnika to: jego kompetencje, kontekst, w którym funkcjonuje, zadanie, które wykonuje oraz elementy poznawczy i afektywny. Dla biblioteki cyfrowej wymiarami są: zawartość, organizacja, reprezentacja, serwisy oraz kontekst. Wizualizacja takich modeli została przedstawiona na rysunku 1:



Rys. 1. Analiza i wymiar głównych elementów badań nad bibliotekami cyfrowymi oraz użytkownikami. Rysunek przetworzony za: *A Framework*...2007.

Natomiast określenie, jak blisko znajdują się użytkownik oraz system przedstawione zostało na rysunku 2:



Rys. 2. Modele biblioteki cyfrowej dla użytkownika oraz model użytkownika dla systemu biblioteki cyfrowej [Za: Zhang; Saracevic 2007a].

Tefko Saracevic pogrupował także poszczególne aspekty ewaluacji wyznaczając jej elementy:

- strukturę (*construct*) oceny – gdzie najczęściej określa się, co konkretnie ma być oceniane, jak rozumiana jest przez badacza 'biblioteka cyfrowa', jakie elementy (komponenty, części, procesy) mają zostać przeanalizowane w trakcie badań;
- kontekst oceny (*context*) – tu określić należy cel ewaluacji, jej ramy, paradygmat, w którym funkcjonuje badacz, podstawowe podejście lub perspektywę ewaluacji, poziomy oceny oraz jej cel;

- określone kryteria oceny (*criteria*) – pytania, parametry, wyznaczone elementy oceny według przyjętych i określonych wcześniej: struktury i kontekstu;
- metodologia oceny (*methodology*) – określająca jakie formy, środki i narzędzia pomiaru zostały wykorzystane, jakie istnieją przykłady ocen, jak wyznacza się próbę do badania, jakie procedury zostały zastosowane w trakcie zbierania danych oraz jakie procedury zastosowano do analizy danych.

Metodologia wskazująca na same narzędzia i formy zastosowane w ewaluacji biblioteki cyfrowej jest rozbudowana. Techniki badawcze stosowane w trakcie ewaluacji, odnoszą się do badań ludzi zaangażowanych w procesy tworzenia i wykorzystywania bibliotek cyfrowych, stąd są to: ankiety, także online, wywiady, zorganizowane grupy dyskusyjne, obserwacja, spotkania, analizy komunikacji (aspekt antropologiczny, socjologiczny), studia przypadku (*case study*) itp. Dla analiz systemu wykorzystywać można analizy dziennika transakcji, eksperymenty, analizę dokumentacji, analizy ekonomiczne itp.

W pracy Thomasa C. Reevesa, Xornama Apedoe i Young Hee Woo z 2005 r. zatytułowanej *Evaluating digital libraries: a user–friendly guide* pojawiła się tabela wskazująca, które zagadnienia powinny być badane i jakimi metodami [s. 17]. Przedstawia się ona następująco:

Tabela 1. Wskazane metody badań poszczególnych problemów i zagadnień w zakresie bibliotek cyfrowych. [Za: Reeves; Apedoe; Hee Woo 2003].

Metody / Zagadnienia	metody heurystyczne	testy użyteczności	opinie ekspertów	nastawienia na użytkownika	testy informatyczne	kwestionariusze dla użytkownika
Ekspercka perspektywa dla BC			X			
Ekspercka perspektywa dla interfejsu (Graphical User Interface)	X		X			
Perspektywa użytkownika		X		X	X	X
Rekomendowane rozszerzenia	X	X	X	X	X	X
Koszty wzmocnienia / rozszerzenia BC	X		X			

W pracy *Metodologia badań przez Internet* Dominika Batorskiego, Marty Olcoń i Jana Mateusza Zająca (2006) wymieniono natomiast standardowy podział technik badawczych, właśnie w odniesieniu do analiz sieciowych systemów informacyjnych. Kategoryzacja ta dotyczy: 1) badań ilościowych / jakościowych, 2) badań ze świadomym udziałem / bez konieczności angażowania osób badanych, 3) badań

realizowanych w czasie rzeczywistym (synchroniczne) / asynchroniczne. Natomiast główne techniki badań, które można zastosować przy analizie bibliotek cyfrowych to: 1) ankiety (różnorodne: dla ocen zapotrzebowania na informację, zasobów, analizy użytkowników itp. najczęściej dostępne online, ale także papierowe), 2) eksperymenty (np. na wymiernych grupach użytkowników – np. w trakcie zajęć dydaktycznych przy zadaniach typu: znajdź informację...), 3) testy psychologiczne (np. dla stworzenia modeli użytkowników oraz ich zachowań informacyjnych), 4) wywiad indywidualny (często pogłębiający np. ankietę), 5) wyselekcjonowanie grup fokusowych, wywiady grupowe, panele (stosowane np. dla grup użytkowników zaawansowanych lub współtworzących bibliotekę cyfrową), 6) obserwacja (np. zachowań użytkowników – może być także online), 7) analiza treści (ilościowa i jakościowa) (głównie dla oceny zasobów), 8) zbieranie danych (np. przez tworzone w systemach informatycznych statystyki), w tym: badania ruchu w internecie (*site–centric* i *user–centric*), analiza logów, analiza zachowań i dane z serwerów, analiza linków.

Ich przyporządkowanie do konkretnych elementów badanego systemu przedstawiono w tabeli 2 (przetworzonej przez autorki) [Janiak, Krakowska 2010a]:

Tabela 2. Techniki pomiaru oraz przyporządkowane im określone elementy systemu.

Techniki pomiaru	
Użytkownik – *dla wszystkich grup użytkowników: uzytkownicy (USR), badacze (RES), bibliotekarze (LIB), zarządzający biblioteką cyfrową (DEV) oraz administratorzy (ADM) – podział według Ying Zhang – zob. podrozdział poświęcony Ying Zhang*	użytkownicy – ankiety, eksperymenty, wywiad indywidualny, focusy, wywiady grupowe, panele, obserwacja, zbieranie danych, testy psychologiczne
	domena (obszar zainteresowań, działań) – ankiety, eksperymenty, wywiad indywidualny, focusy, wywiady grupowe, panele, obserwacja, zbieranie danych, testy psychologiczne
	sposób docierania do informacji – ankiety, eksperymenty, wywiad indywidualny, focusy, wywiady grupowe, panele, obserwacja, zbieranie danych, testy psychologiczne
	cel – ankiety, eksperymenty, wywiad indywidualny, zogniskowany wywiad grupowy (fokus), wywiady grupowe, panele, zbieranie danych (zwłaszcza: pytania zadawane przez użytkowników), testy psychologiczne
Dane / kolekcje	zawartość – zbieranie danych, analiza treści (ilościowa i jakościowa)
	meta–zawartość – analiza treści (ilościowa i jakościowa), także benchmarking czyli tu: zgodność z formatem
	zarządzanie – ankiety, obserwacja, zbieranie danych, zogniskowany wywiad grupowy (fokus), wywiady grupowe, panele, wywiad indywidualny (np. z administratorem, bibliotekarzem itp.)

Technologia	technologia użytkownika – ankiety, eksperymenty, wywiad indywidualny, zogniskowany wywiad grupowy (fokus), wywiady grupowe, panele, obserwacja, zbieranie danych, testy psychologiczne
	dostęp do informacji – przy analizie jakości dokonywanej przez użytkownika: ankiety, eksperymenty, wywiad indywidualny, zogniskowany wywiad grupowy (fokus), wywiady grupowe, panele, obserwacja, zbieranie danych, testy psychologiczne; metody informatyczne: zbieranie danych (np. sztucznie tworzone przez system komputerowy pytania – dla analizy szybkości reakcji itp.)
	technologia konstrukcji systemu – przy analizie jakości dokonywanej przez użytkownika: ankiety, eksperymenty, wywiad indywidualny, zogniskowany wywiad grupowy (fokus), wywiady grupowe, panele, obserwacja, zbieranie danych, testy psychologiczne; metody informatyczne: testy siłowe jak przy innych bazach danych itp.
	technologia dokumentu – analiza treści (ilościowa i jakościowa), także benchmarking czyli tu: zgodność z formatem

W związku z tym, iż biblioteki cyfrowe są złożonymi podmiotami, możliwe jest także integrowanie wielu metod i technik. Nie można też mówić o jednej, wypracowanej metodzie badań bibliotek cyfrowych, co już zostało zaznaczone.

2. Kryteria oceny systemów informacyjnych

Do oceny bibliotek cyfrowych można zastosować także różnorodne ogólne formularze i zestawy kryteriów ocen przygotowanych do analizy komputerowych systemów informacyjnych, które przygotowywane były już od połowy lat dziewięćdziesiątych XX wieku. Najbardziej znane opracowane zostały przez Carolyn Caywood (1995), Thaddeusa Matthew Ciołka (1994 – popr. 1997; 1996), Jane Alexander i Marshę Tate (1996), Alistaira Smitha (1996), Hope N. Tillman (1996, popr. 2003), Esther Grassian (1996) i Helge Clausena (1999).

W Polsce rozpowszechniony został zestaw kryteriów, przygotowany przez Alistaira Smitha, który wymienił następujące elementy: 1) zakres (szerokość; głębokość; czas; format); 2) zawartość (dokładność, ścisłość; autorytatywność; aktualność; unikalność; linki do innych źródeł; jakość tekstu); 3) wygląd grafiki i multimediów; 4) cel (tu: odbiorca); 5. przeglądy (recenzje); 6) wykonanie (przyjazność dla użytkownika; wymagany sprzęt, środowisko softwarowe; wyszukiwanie; przeglądalność (przeszukiwalność) i organizacja; interakcyjność; połączenia); 7) koszt.

Podobny zestaw, także popularny w naszym kraju to opisana przez Helge Clausena metoda jakościowo-heurystyczna, w której wymienione zostały komponenty: 1) struktura i wygląd stron (organizacja treści, racjonalne użycie grafiki), 2) jakość

informacji (instytucja sprawcza, aktualizacja, poprawność, ilość informacji), 3) linki i nawigacja (adnotacje, ilość odsyłaczy, łatwość wertowania), 4) wrażenie estetyczne (kolory, czcionki, format), 5) miscellanea (pomoc, użytkownik), 6) ogólne (oryginalność, przyjazność serwisu) [Bednarek-Michalska 2002].

Z instytucjonalnych zestawów kryteriów jakości, które także były wykorzystywane w Polsce, wymienić należy przede wszystkim ten przygotowany przez serwis SOSIG (*Social Science Information Gateway;* dziś Intute – od 2009) [*A list of quality...* 2007]. Podobne mierniki ewaluacji powstały np. pod auspicjami: UNESCO [*ICT for Library...* 2001], *American Library Association* [*Great Web Sites* 1997], Cornell University Library (2004) lub The Johns Hopkins University – The Sheridan Libraries (2010).

Inne, zdecydowanie bardziej specyficzne formularze ocen systemu informacyjnego to np. przygotowane przez Kathy Schrock, zestawy kryteriów dla nauczycieli i uczniów. Pomagają one spojrzeć na zawartość bibliotek cyfrowych od strony potrzeb edukacyjnych. Kathy Schrock przygotowała także np. formularze ocen blogów oraz portali, w których przedstawiono 'wirtualne wędrówki'. Mogą one dopomóc w analizie wirtualnych galerii i muzeów, które już są tworzone w podobnym do dLibry oprogramowaniu dMuseion (Poznańskie Centrum Superkomputerowo-Sieciowe – Cyfrowe Muzeum Narodowe w Warszawie http://cyfrowe.mnw.art.pl/dmuseion).

Z polskich twórców kryteriów jakościowych wymienić należy przede wszystkim Bożenę Bednarek-Michalską (2002, 2007) oraz Katarzynę Materską (2007), które przedstawiły ogólne zagadnienia oceny systemów informacyjnych oraz Remigiusza Sapę (2005), który przeanalizował strony bibliotek akademickich.

Bożena Bednarek-Michalska spośród cech określających jakość informacji wymieniła: relewantność; dokładność; aktualność; kompletność; spójność; odpowiedniość formy; dostępność; przystawalność; wiarygodność. Kryteria oceny jakości informacji elektronicznej dopracowała poprzez wymienienie cech specyficznych dla tego typu danych. I tak dla: 1) wiarygodności informacji ważna jest analiza url, identyfikacja instytucji sprawczej i autora oraz ich celów, odnalezienie opinii o portalach oraz analiza rzetelności opracowanej informacji; 2) obiektywizmu treści: sponsoring, linki do innych, także konkurencyjnych stron, obiektywny komentarz, prezentacja różnych punktów; 3) aktualizacji: jej częstotliwość, czy też np. podanie dat powstawania serwisu, 4) kompletności i dokładności zasobów: liczba danych, kryteria ich doboru – także dla grupy docelowych użytkowników, relewantność, związek między zasobami a celami serwisu, linki i ich wartość, 5) spójności i formy: design, struktura i jej logika, równowaga między formą a treścią, oryginalność, multimedialność, błędy, 6) ostępności i techniki: interfejs, wertowalność i nawigacja, interaktywność, jakość pracy serwera, liczniki, statystyka, software, najnowsze rozwiązania technologiczne do przekazywania informacji i poszerzania oferty, 7) relewancji i użyteczności: dobór treści do potrzeb docelowego (zidentyfikowanego przez twórców) użytkownika, pliki pomocowe, inne wersje językowe, komunikacja z użytkownikiem oraz dostosowanie serwisu dla słabo widzących lub niesłyszących (informacje dla plików audio i wideo).

Katarzyna Materska w książce *Informacja w organizacjach społeczeństwa wiedzy* wymieniła następujące pożądane cechy jakościowe: 1) dyspozycyjność, 2) aktualność,

3) prawdziwość, 4) poprawność, zgodność, 5) wierność, rzetelność, dokładność, precyzja, 6) użyteczność, przydatność, istotność, 7) wiarygodność, 8) selektywność (ilość informacji), 9) kompletność, pełność, 10) terminowość, 11) komunikatywność (zrozumiałość), 12) jednoznaczność, 13) dostępność, 14) porównywalność, kompatybilność, spójność, 15) 'nowość', 16) cenność, 17) ciągłość, 18) taniość, 19) obiektywność, neutralność, 20) wyłączność, 21) poufność, 22) atrakcyjność [s. 88-90]. Jak można zauważyć, badanie jakościowe wymienionych cech, odnosi się w największym stopniu do zawartości i funkcjonowania systemu informacyjnego. Są jednak także właściwości istotne dla użytkowników – np. przydatność.

Remigiusz Sapa podzielił badane cechy na odnoszące się do treści oraz funkcjonalności systemów. W związku z tematem badań: serwisy WWW bibliotek akademickich, analiza treści serwisów skupiona została wokół zagadnień: 1) biblioteki i zasoby tradycyjne w serwisie WWW, 2) katalog OPAC, 3) zasoby elektroniczne dostępne dla uprawnionych użytkowników, 4) zasoby elektroniczne dostępne publicznie, 5) usługi informacyjne, 6) obsługa procesu pozyskiwania zasobów niedostępnych w danej bibliotece, 7) edukacja i instrukcje, 8) aktywność naukowa i profesjonalna, 9) przestrzeń komunikacji, 10) promocja biblioteki serwisu, 11) obsługa interesantów, 12) element wirtualnego środowiska informacyjnego uczelni. Natomiast parametry odnoszące się do funkcjonalności to: 1) przejrzystość, 2) komunikatywność, 3) spójność i konsekwencja, 4) prewencyjność, 5) narzędzia nawigacji, 6) sprawność, 7) sekwencyjność, 8) elastyczność, 9) aktualność [s. 148-160].

Wszystkie wymienione w tym podrozdziale zestawy mogą oczywiście po pewnych przeróbkach stanowić także podstawę do oceny bibliotek cyfrowych, jako specyficznego typu systemów informacyjnych.

3. Ocena bibliotek cyfrowych

Zagadnienie oceny bibliotek cyfrowych pojawiło się w literaturze zachodniej na przełomie XX i XXI wieku. Wielu autorów przygotowało analizy na temat możliwości oceny modelowej biblioteki cyfrowej lub też przeprowadziło badania konkretnych istniejących obiektów (zob. bibliografia). Aspekty badań jakościowych przedstawione zostały np. w 2002 r. trakcie trwania 4. Forum DELOS [Fourth DELOS Workshop 2002]. Zakres badań ograniczony został do trzech podstawowych pytań: 1) *co* się znajduje w bibliotece cyfrowej, czyli jaka jest jakość *zasobu*, 2) *kto* korzysta z biblioteki cyfrowej, czyli jakie aspekty wpływają na oceny *użytkowników*, oraz 3) *jak* jest wykorzystywana biblioteka cyfrowa czyli a) jaka została zastosowana *technologia,* która umożliwia zastosowanie konkretnych funkcji oraz b) *cel* wchodzenia w relacje z systemem informacyjnym (potrzeby informacyjne). Elementy analizowane w takim rozumieniu biblioteki cyfrowej przedstawić można następująco:

Tabela 3. Analizowane elementy biblioteki cyfrowej [Za: Fourth DELOS Workshop 2002].

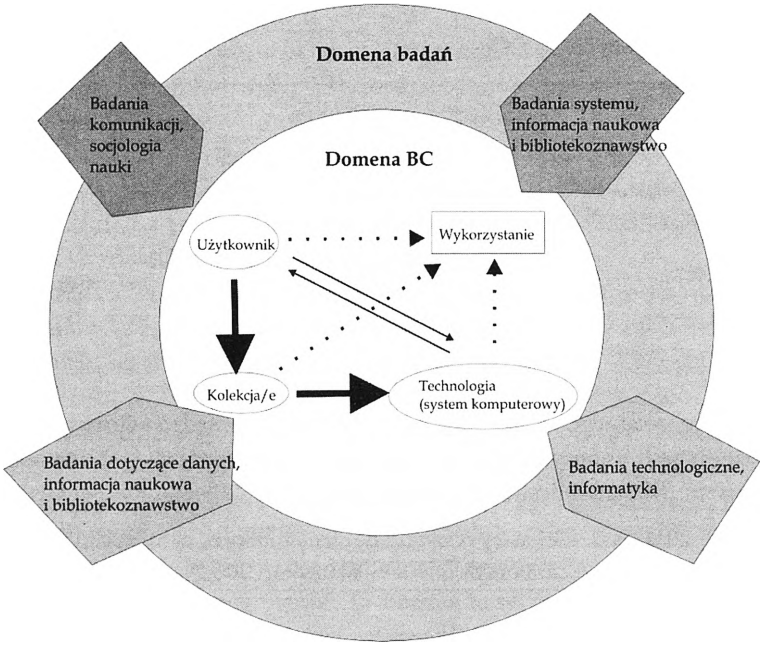
KTO **używa**	**CO** **i**	**JAK [Mabe 2002]**
aspekty demograficzny społeczno–ekonomiczny lojalność wobec instytucji (marki) stały / tymczasowy użytkownik styl życia	produkt miejsce cena rynek zbytu usługi	korzyści postawy percepcja preferencje

W trakcie tego samego forum (4. Forum DELOS), przedstawiono także elementy tworzące domenę bibliotek cyfrowych oraz określono jak mogą być one analizowane [Fuhr; Hansen; Mabe; Micsik; Sølvbergh 2001]:

Tabela 4. Elementy tworzące domenę bibliotek cyfrowych [Za: Fourth DELOS Workshop 2002].

Element domeny	**Badania**
użytkownicy	pod względem demografii, domen ich zainteresowań i metod, które zostały zastosowane w trak cie docierania do informacji;
dane / kolekcje	pod względem opisu, sposobów selekcji i gromadzenia oraz zarządzania;
technologie	pod względem stosowanych technologii (interfejs użytkownika, systemy informatyczne, formaty dokumentów itp.) oraz narzędzi informatycznych;
wykorzystanie (cel)	w naturalny sposób przyporządkowane do użytkownika, odnoszące się do charakteru działań informacyjnych. Głównym pytaniem jest określenie, czy informacje są jedynie wykorzystane, czy też użyte do analizy lub syntezy.

W trakcie konferencji zaprezentowano także obszary badań nad bibliotekami cyfrowymi w odniesieniu do elementów domeny. Wymienione zostały badania komunikacji, socjologiczne, systemowe, informatologiczne, bibliotekoznawcze, technologiczne, informatyczne. Wizualizacja wpływu różnych dyscyplin, wiedzy i doświadczenia naukowców, a także metod, które są wykorzystywane w badaniach interdyscyplinarnych oraz w określonych domenach została przedstawiona na rysunku 3.



Rys. 3. Paradygmat naukoznawczy domeny badań nad biblioteką cyfrową, rysunek przetworzony [Za: Fourth DELOS Workshop 2002].

Wyselekcjonowane propozycje klasyfikowania oraz parametryzacji bibliotek cyfrowych, wymiary badań oraz sposoby ich pomiaru przedstawiono także w formie tabeli 5.

Tabela 5. Propozycja klasyfikacji i parametryzacji bibliotek cyfrowych oraz wymiary badań i sposoby ich pomiaru według Fourth DELOS Workshop 2002 – przetworzone przez autorki.

PROPOZYCJE KLASYFIKACJI I PARAMETRYZACJI BIBLIOTEK CYFROWYCH według *Fourth DELOS Workshop. Evaluation of Digital Libraries: Testbeds, Measurements, and Metrics* – przetworzone przez autorki [Janiak; Krakowska 2010a]			
Klasyfikacja Wymiar pierwszy Wymiar drugi			Sposoby pomiaru – systemy wskaźników
Użytkownik	użytkownicy	wewnętrzny, ogólny, edukacyjny, profesjonalny, badawczy	liczba, dystrybucja (rozkład danych w wymiarach drugich – np. rozkład demograficzny)
	domena (obszar zainteresowań, działań)		dystrybucja
	sposób docierania do informacji	szukanie, przeglądanie	dystrybucja

	cel	konsumpcja, analiza, synteza	dystrybucja
Dane / kolekcje	zawartość	brak (*none*) / częściowy / całościowy, audio, wideo, tekst 2D/3D	różnorodność, wiek, wielkość, jakość („biała" – ogólnodostępna / szara literatura)
	meta–zawartość	opis bibliograficzny, indeksowanie / tezaurus / klasyfikacja	media (np. ogólny i łatwy dostęp), poziom szczegółowości
	zarządzanie	prawa, workflow, zarządzanie użytkownikami, utrzymywanie kolekcji	wiek dokumentów (prawa autorskie i wydawnicze), tempo wzrostu, bliskość (w stosunku do użytkowników) – bycie w kontakcie z użytkownikami, kompletność, utrzymywanie kolekcji badane w pewnych odstępach czasu (dynamika zmian)
Technologia	technologia użytkownika	tworzenie dokumentów, docieranie do informacji, interfejs, przeglądanie, wyszukiwanie, drukowanie, grupa / osoba	[brak jest jasnej metodologii tworzenia pomiarów, można próbować wykorzystywać praktykę benchmarkingu]
	dostęp do informacji	wyszukiwanie, nawigacja, filtrowanie, wyekstrahowanie, *text mining* (przeszukiwanie tekstów)	efektywność, skuteczność
	technologia konstrukcji systemu	repozytorium, modele dostarczania i przechowywania	[brak jest jasnej metodologii tworzenia pomiarów, można próbować wykorzystywać praktykę benchmarkingu]
	technologia dokumentu	model dokumentu (dla różnych ich typów), format	[brak jest jasnej metodologii tworzenia pomiarów, można próbować wykorzystywać praktykę benchmarkingu]

Obiekty badań ewaluacyjnych zostały także przedstawione w pracy *Digital Library Evaluation, User Studies. Draft, 2007* autorstwa Barbary Wildemuth, która powstała m.in. dla projektu *Digital Library Curriculum Project.* Są to: same biblioteki cyfrowe i ich procesy, funkcje oraz studia nad użytkownikiem. Autorka odniosła się więc do trzyczłonowego modelu biblioteki cyfrowej: zasoby, system (m.in. oprogramowanie) oraz ludzie. Główne aspekty badań (elementy poddawane badaniom), to: 1) informacje o reprezentacji, metadanych i substytutach wykorzystywanych w bibliotekach cyfrowych (adekwatność zadań, użyteczność, zadowolenie użytkowników); 2) poszczególne narzędzia (indeks, wyszukiwanie i cechy (właściwości) wyjścia, nawigacji, przeglądanie, błędy w funkcjonalności i użyteczności, zadowolenie użytkowników), 3) usługi (jakość kolekcji, jakość wykonania (pamięć, dokładność), niezawodność, usługi dla ludzi-pośredników (np. usługi informacyjne), zadowolenie użytkowników z poszczególnych usług lub zestawu usług), 4) zachowania użytkownika podczas interakcji – może być lub nie być ewaluacji bezpośredniej (szukanie informacji / zachowania informacyjne, korzystanie z informacji odnalezionej, formy (wzorce) pracy). Ponadto wyodrębniła ona badania nad: szczególnymi aspektami BC, indywidualnymi bibliotekami cyfrowymi oraz wieloma bibliotekami cyfrowymi.

4. Wybrane kryteria ocen bibliotek cyfrowych

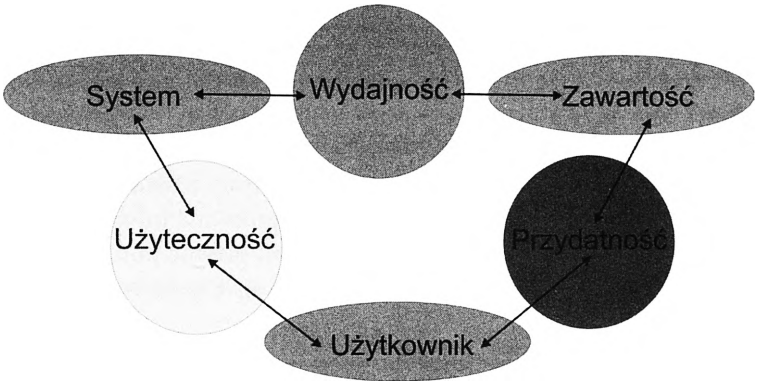
Do przedstawienia poszczególnych kryteriów oceny jakości bibliotek cyfrowych zostały wybrane najbardziej ogólne, zawierające szerokie spojrzenie na to zagadnienie oraz akceptowane w środowisku prace, takie jak: 1) analiza Norberta Fuhra i jego zespołu, 2) analizy jakościowe przygotowywane dla projektu DELOS, 3) holistyczne spojrzenie na ocenę jakości Ying Zhang, 4) model 5S, 5) ewaluacja z perspektywy użytkownika Hong Iris Xie oraz 6) lista atrybutów koniecznych do porównywania bibliotek cyfrowych Roberta J. Sanduskiego.

A framework of guidance for building good digital collections (2007), który także może stanowić podstawę oceny bibliotek cyfrowych został pominięty, gdyż jego analiza znalazła się w tej książce w pracy Bożeny Bednarek-Michalskiej.

*

Pierwszą szerzej omówioną analizą są prace autorstwa zespołu: **Norbert Fuhr, Giannis Tsakonas, Trond Adalberg, Maristella Agosti, Preben Hansen, Sarantos Kapidakis, Claus-Peter Klas, László Kovács, Monica Landowi, András Micsik, Christos Papatheodorou, Carol Peters, Ingeborg Sølvberg,** wykorzystane także dla projektu DELOS.

Autorzy ci dla celów badań jakościowych przygotowali trzyczłonowy model biblioteki cyfrowej, której poszczególne elementy połączone są relacjami. Najważniejsze komponenty to: 1) użytkownik, 2) zawartość, 3) system. Natomiast relacje między nimi opisane zostały jako: 1) wydajność (system-zawartość), 2) przydatność (użytkownik-zawartość), 3) użyteczność (użytkownik-system). Wizualizacja tego modelu przedstawiona została na rysunku 4:



Rys. 4. Interaktywny model trójczłonowy biblioteki cyfrowej, rysunek przetworzony za Fuhr 2007.

Dwie relacje związane z użytkownikami: użyteczność i przydatność wymagają poza analizą systemu oraz jego zawartości, także badań nad ludźmi zaangażowanymi w procesy informacyjne. Można to robić także bez konieczności angażowania samych czytelników np. przy badaniu użyteczności można wykorzystywać automatyczne techniki, opierając się na ilościowych danych, analizując dzienniki transakcji, czy też obserwując i rejestrując ich zachowania w trakcie interakcji z biblioteką cyfrową. Innymi technikami są techniki komparatywne, które porównują system z określonymi standardami i wytycznymi lub innymi systemami operacyjnymi (benchmarking). Część tych analiz także może być wykonywana automatycznie. Analityczne techniki, które można tu jeszcze wykorzystać wymagają udziału ekspertów, którzy na podstawie własnej wiedzy zbadają system, samodzielnie lub we współpracy z innymi ekspertami. Odmienne badania użyteczności opierać można na badaniach empirycznych, wspartych wywiadami, dyskusjami w grupach, ankietami, czy kwestionariuszami. Dzięki nim zgromadzić można takie dane jakościowe, jak np. preferencje użytkownika, opinie objaśniające ich zachowania informacyjne czy poziom satysfakcji.

Badania użyteczności skupiają się także na ocenianiu interfejsu, jego funkcjonalności dla pojedynczych użytkowników oraz różnych ich grup (wiekowych, połączonych wspólnymi zainteresowaniami, podobnym wykształceniem itp.).

Możliwe jest również prowadzenie np. studiów nad zadaniami oraz kontekstem określonych sytuacji informacyjnych. Aspekty, które można ocenić to m.in.: 1) rodzaje zadań i etapy poszczególnych zadań, 2) wykonywanie zadań i procedury zadaniowe, 3) złożoność zadań i zmiany zachodzące w trakcie zadań, 4) różnice między domenami, 5) wpływy (impact) istniejące na równych poziomach organizacyjnych, 6) rodzaje systemów dostępu i przetwarzania informacji.

Pozostałe analizy odnoszą się do badań wykorzystania zdobytych informacji. Nie kończy się więc procesu badawczego w sytuacji opuszczenia biblioteki cyfrowej, lecz poszukuje się odpowiedzi na pytania: 1) jaka część źródeł informacji jest faktycznie wykorzystywana i w jaki sposób, 2) kiedy i jak nabyte / dostępne / odnalezione źródło

informacji jest wykorzystane, 3) jakie elementy źródła informacyjnego są wykorzystywane i jaki rodzaj wyników jest tworzony przez użytkowników.

Kolejne kryterium – przydatność powstaje w trakcie realizacji potrzeb informacyjnych użytkowników przy dostępie do zawartości bibliotek cyfrowych. Odzwierciedla ono sposób, w jaki użytkownicy widzą znaczenie konkretnego zasobu. Ocena przydatności powstaje więc w dość subiektywny sposób, który badacze starają się później uogólnić. Analizowane tu cechy to: szerokość, szczegółowość informacji, jakość, oraz ważność (znaczenie, unikatowość itp.) zbiorów. Dwa główne podejścia, które stosowane są przy badaniu tej relacji to: badania zachowań użytkowników oraz badania samej informacji.

W zakresie badań nad użytkownikami mieszczą się analizy nad zrozumieniem takich zjawisk jak zachowania w trakcie poszukiwania informacji, obsługi systemów oraz „udoskonalania" własnej wiedzy, a więc wykorzystywania zdobytej informacji. Dla zbadania zachowań informacyjnych należy skoncentrować się na wzorcach zachowań, które są generowane poprzez różne klasy użytkowników oraz różne systemy informatyczne. Analizy takie są bardzo interdyscyplinarne, odnoszą się do wielu dyscyplin, np. wykorzystują metody psychologii i nauk kognitywnych, a z drugiej strony odnoszą się do informatyki oraz informacji naukowej. Badany jest bowiem cały proces (zachowanie w trakcie) poszukiwania informacji i jej pobierania, chociaż często dzieli się go na różne etapy i analizuje w odniesieniu do działań, aspektów poznawczych oraz emocji. Interesujące tematy do badań to np.: 1) motywacje użytkowników, 2) afektywne stany procesu wyszukiwania informacji, 3) różnice między niektórymi grupami użytkowników, 4) zorganizowany lub niestrukturalny charakter dostępu do informacji, czy 5) nieoczekiwane wybory w trakcie poszukiwania informacji. Techniki gromadzenia danych to: techniki surveyowe, kwestionariusze, wywiady, a także badania grup dyskusyjnych. Stosowane są także obserwacje, w tym nagrywanie sesji interakcyjnej z użytkownikiem.

Wśród rodzajów badań nad kryterium przydatności wymienić można: 1) typy użytkowników i ich cechy, takich jak różny poziom wiedzy i doświadczenia, 2) potrzeby informacyjne użytkowników, 3) różne wskaźniki satysfakcji (np. funkcje i realizacji zadań), 4) strategie przetwarzania informacji, 5) zadania i procedury zadaniowe, 6) reprezentacja domen pracy i środowiska, 7) współpraca między użytkownikami i grupami użytkowników.

Do zakresu badań nad zawartością należą przede wszystkim zagadnienia: jakiego rodzaju treści dostępne i jak uzyskać do nich dostęp. Badane elementy, związane z naturą treści to np.: 1) charakter treści, 2) forma informacji (tekst, wideo, obraz, dźwięk), 3) język (język (–i) i czy zawartość jest jedno–, dwujęzyczna, czy wielojęzyczna), 4) sposób tworzenia (zdigitalizowane, powstałe w formacie cyfrowym – *digital born*), 5) typ (biała, szara literatura). Zagadnienia związane ze strukturą zawartości to: 1) poziom (obiekty, czy metadane), 2) rozmiar / wielkość, 3) spójność (homogeniczne lub heterogeniczne atrybuty – związane z normami i interoperacyjnością technologii), 4) tryb dostarczania zawartości (protokoły hipertekstowe, technologie typu *streaming*). Natomiast spośród zagadnień związanych z administrowaniem wymienić

należy: 1) dynamika rozwoju kolekcji (rosnąca, brak wzrostu), 2) prawa (związane ze sposobem i treścią licencji autorskich, wydawniczych, ogólnoprawnych), 3) tryb licencji (otwarty dostęp, jednorazowy zakup, abonament w wielu odmianach).

Ostatnia już relacja wydajności odnosi się do systemu i jego zawartości. Badania nad tym kryterium są najbardziej rozpowszechnione, obejmując również ocenę skuteczności, funkcjonalności oraz analizy określonej konstrukcji systemu. Badacze koncentrują się w nich na ocenie dostępu do informacji i jej ekstrakcji z całego otoczenia. Najczęściej analizowane elementy to: architektura systemu, sposoby przechowywania danych, punkty dostępu do informacji, czy zarządzanie treściami multimedialnymi.

*

Kolejne szerzej omówione w tej pracy probierze jakości przygotowane zostały w ramach Sieci Doskonałości **DELOS**.

Cała biblioteka cyfrowa może być analizowana jako obiekt / twór, wytwór / idea (*constructs*). Tefko Saracevic wyróżnił dwa odmienne podejścia do wartościowania: 1) ewaluację określonej biblioteki cyfrowej, traktowanej jako podmiot oraz 2) ewaluację określonego procesu, który jest związany z funkcjonowaniem biblioteki cyfrowej, ale nie stanowi integralnej części danego podmiotu. I tak podmioty ocen (*entitites*) to: 1) biblioteki cyfrowe, które zostały specjalnie skonstruowane do realizacji projektów badawczo–rozwojowych (np. projekt *Perseus*), 2) niektóre aspekty działania bibliotek cyfrowych, 3) oceny złożonych bibliotek cyfrowych [Saracevic 2000; 2004]. Natomiast oceny procesów odnoszą się do: 1) oceny różnych reprezentacji cyfrowych bibliotek, w tym formułowania określeń (słów kluczowych, rzeczowników) w medycznych kolekcjach cyfrowych, kontekstowej reprezentacji literatury medycznej lub ekstrakcji fraz w bibliotekach cyfrowych, 2) oceny różnych narzędzi, w tym ewaluacji wytwarzania związków, wyszukiwania dokumentów np. cyfrowego wideo, interfejs użytkownika, metodologii pobierania obrazu lub równoważenia obciążenia na serwerach, 3) oceny różnych usług, w tym systemu rekomendacyjnego bibliotek cyfrowych, serwisów referencyjnych, 4) ewaluacji samego systemu oceny, 5) badań zachowań użytkowników w związku z interakcją użytkownik – biblioteka cyfrowa lub w zakresie oceny usług, np. sposobu korzystania z usług lub ocena logowań, oceny jakości bibliotek cyfrowych przez nauczycieli, oceny modeli aktywności profesjonalistów w zakresie rozwoju narzędzi biblioteki cyfrowej, preferencji użytkowników w procesie wyszukiwania w pełnotekstowych bazach danych, poszukiwania informacji.

Proces ewaluacji zakotwiczony został w ogólnych ramach, jakim jest kontekst badań, podejście, orientacja, poziom oraz cele oceny. Biblioteki cyfrowe stanowią bowiem kompleksowy system, który jest złożony zarówno społecznie, instytucjonalnie, jak i technicznie (co zostało wcześniej kilkakrotnie zaznaczone). Ewaluacja może być więc różnorodna z uwagi na wybrany aspekt oceny systemu oraz ze względu na orientacje stosowane w trakcie badań. Podstawowe, zidentyfikowane orientacje to: 1) podejścia skoncentrowanego na systemie – najbardziej rozpowszechnione, odnoszące się do wydajności biblioteki, skuteczności całego mechanizmu, i/lub skuteczności niektórych funkcji, czy też wybranych elementów technicznych systemu, 2) podejście

zorientowane na interakcję człowiek–komputer, które odnosi się do badań zachowań informacyjnych, które często analizuje się w odniesieniu do potrzeb informacyjnych (poszukiwanie informacji, przeglądanie, wyszukiwanie, efektywność realizacji powierzonych zadań), 3) podejście skupione na użyteczności, w którym ocena odnosi się analizy różnorodnych funkcji, często określonej przez użytkowników (kompilacja podejścia systemowego i skoncentrowanego na użytkowniku), 4) podejście etnograficzne, które polega na badaniu sposobów istnienia, kultury i obyczajów w środowisku cyfrowym biblioteki oraz wpływu bibliotek cyfrowych na dane społeczności, 5) podejście antropologiczne, które obejmuje różnych interesariuszy (współtwórców i użytkowników końcowych) systemu oraz społeczności i ich kulturę w odniesieniu do danej biblioteki cyfrowej, 6) podejście socjologiczne, skupione na ocenie sytuacyjnych działań lub społeczności użytkowników w otoczeniu biblioteki cyfrowej, 7) podejście ekonomiczne, odnoszące się do oceny kosztów, korzyści finansowych, wartości ekonomicznych oraz oddziaływania względem siebie parametrów ekonomicznych [Saracevic 2004].

Jednak za najtrudniejsze zagadnienie związane z procesem ewaluacji zaklasyfikowane zostało opracowywanie kryteriów ocen, odnoszących się do wybranych przez badaczy standardów, wykorzystanych później do opracowania wskaźników ilościowych. Jako podstawowa określona została relewancja, już od lat 50. XX w. stosowana w ocenie systemów informacyjno-wyszukiwawczych. Pozostałe elementy to: same kolekcje, technologia, czy usługi.

Do wielokrotnie stosowanych kryteriów Saracevic zaliczył użyteczność, pomimo braku jednoznacznej zgody co do tego, czego dotyczy funkcjonalność czy przydatność w kontekście oceny biblioteki cyfrowej. Najczęściej stosowana jest definicja użyteczności *jako zakresu, w którym dany obiekt, produkt, system może być wykorzystywany przez użytkowników w celu osiągnięcia określonych celów, z uwzględnieniem jego skuteczności (dokładności i kompletności uzyskanej informacji), efektywności oraz satysfakcji (czyli pozytywnego nastawienia do przyszłego użytkowania obiektu, produktu, systemu) w wymiernym kontekście użycia* [ISO http://www.userfocus.co.uk/resources/iso9241/part11.html]. Można więc potraktować ten wskaźnik jako *meta–kryterium*, dla którego wybrać należy takie probierze, jak: zawartość, proces (zachowania informacyjne), format, a także inne, bardziej ogólne wskaźniki.

Sama zawartość, treść może być ewaluowana przez różne cechy, a zwłaszcza przystępność, osiągalność, klarowność, poziom złożoności (organizacji, struktury), informacyjność, zrozumiałość, przejrzystość, wysiłek włożony przez użytkownika w zrozumienie treści, adekwatność, relacyjność, współmierność z oczekiwaniami, jakość, aktualność, dokładność, rzetelność oraz autorytatywność. Kolejny probierz, czyli proces obejmuje wszystkie działania, jakie użytkownik podejmuje w systemie, a więc: wyszukiwanie, nawigowanie, przeglądanie, ocenianie lub pozyskiwanie zasobów. Ocenianie go odbywa się poprzez mierzenie łatwości: uczenia się systemu oraz realizowania procesów informacyjnych (*learnability*), a także stopnia wysiłku i czasu koniecznego do wykonania zadań Dochodzą do tego analizy przygotowanych udogodnień, elementów wspomagania aktywności, komfortu i łatwości korzystania, po-

czucia niepewności i zagubienia podczas wykonywania zadań, pewności wyników po realizacji zadania, interpretacji trudności oraz wskaźnik błędu. Jeśli chodzi o ocenę formatu, to istotne jest utrzymywanie danej reprezentacji obiektu, kolekcji, spójności etykiet oraz komunikatywności zawartej w formacie informacji. Kryterium stanowiące zestaw ogólnych elementów oceny odnosi się do satysfakcji użytkownika, osiągnięcia przez niego celu wyszukiwawczego, relewancji oraz przydatności wyników, ich wartości i jakości, barier, wpływu całego procesu na użytkownika i jego dalsze zachowania informacyjne, a także uczenie się, zdobytych doświadczeń podczas korzystania z obiektu, preferencji, ale także negatywnych afektów [Saracevic 2000].

Inne kryteria, to te odnoszące się do oceny systemów informatycznych. Są to więc parametry związane z technologią, wydajnością procesów oraz zastosowanych algorytmów, takie jak: czas reakcji oprogramowania, systemu, czas przetwarzania, pojemność, prędkość oraz poziom obciążenia serwerów i sieci. Wydajność procesu lub algorytmu mierzona jest przez relewancję, grupowanie podobieństw, funkcjonalność, porównanie systemu, wskaźnik błędu, optymalizację, logiczne decyzje systemowe, liczbę kliknięć, długość ścieżki, czas pobierania danych. Ogólnie system ewaluuje się w oparciu o wskaźnik obsługiwalności (*maintability, performance*) biblioteki cyfrowej, skalowalności, ilości i zakresu udostępnianej informacji, a także kosztów. Ocena według tych kryteriów odnosi się także do parametru użyteczności, omówionego wcześniej [Saracevic 2004].

Można także ewaluować biblioteki cyfrowe, rozpatrując je jak organizacje posiadające zasoby oraz wyspecjalizowane grupy pracowników, do zadań których należy wybór, tworzenie struktury, intelektualnych narzędzi dostępu i reprezentacji cyfrowych obiektów. Dodatkowymi pracami są: rozpowszechnianie oraz dbanie o zachowanie integralności i trwałości kolekcji, które mają być łatwo i po niewielkich kosztach dostępne dla określonej społeczności czy grup społecznych [Digital Library Federation 1998]. Na tej podstawie zbudować można zestaw kryteriów odnoszący się do: profesjonalnego personelu, zbioru dzieł cyfrowych, zasad wyboru, struktury i dostępu, interpretacji (zasobów) i dystrybucji, ochrony, a także użyteczności, ekonomiczności oraz dostępności dla określonych społeczności użytkowników.

Kolejny ważny aspekt badań to integralność biblioteki cyfrowej, dla oceny której istotne są następujące elementy: zasoby cyfrowe, zbiory; ich wybór, gromadzenie, utrzymywanie, rozpowszechnianie; dystrybucja, relacje pomiędzy nimi, linki; organizacja, struktura, przechowywanie; interpretacja; reprezentacja, metadane; zarządzanie kolekcją; trwałość; dostęp; istniejące sieci fizyczne; interfejsy, interakcje; wyszukiwanie, przeszukiwanie; usługi, serwisy; dostępność, zakres dostępnych usług – np. rozpowszechnianie informacji, dostarczanie zbiorów cyfrowych; integracja i współpraca z innymi bibliotekami cyfrowymi, zasobami; wsparcie, pomoc (skierowywanie do odpowiednich zasobów, informacji); użytkowanie, użytkownicy, społeczności; bezpieczeństwo, prywatność, polityka, aspekty prawne, licencje; zarządzanie, ekonomika; integracja, współpraca z innymi zasobami, źródłami, bibliotekami, serwisami [Saracevic 2000].

Jak widać z powyższych kryteriów ocena bibliotek cyfrowych odbywać się może na różnych poziomach. Tefko Saracevic podzielił podejścia do ewaluacji na siedem głównych klas, poziomów, etapów, które nie muszą wykluczać się wzajemnie. Pierwsze trzy mieszczą się w ogólnym zorientowaniu badań na człowieka (np. badanie zachowań w odniesieniu do danych potrzeb informacyjnych) i są to:

1) poziom społeczny, z zagadnieniami prawidłowego wspierania i odpowiadania biblioteki cyfrowej na potrzeby i wymagania społeczeństwa lub społeczności,
2) poziom instytucjonalny, gdzie problematyka skupiona jest na misji i celach instytucji, organizacji, integracji zawartości biblioteki cyfrowej z innymi zasobami instytucjonalnymi,
3) poziom indywidualny, w którym główne kwestie związane są z prawidłowym wspieraniem potrzeb informacyjnych, realizacji zadań oraz działań indywidualnych użytkowników lub grup użytkowników.

Kolejna klasa skupiona jest na relacji człowiek-system i odnosi się do:

4) interfejsu. Przede wszystkim analizowany jest tutaj ogół problemów związanych z ty, czy dany interfejs zapewnia dostęp i wsparcie, wyszukiwanie, nawigowanie, przeglądanie i współdziałanie użytkownika z biblioteką cyfrową. Pytania można zadawać zarówno w odniesieniu do użytkownika, jak i systemu, a także w obu kierunkach.

Ostatnia już orientacja skoncentrowana wokół systemu dotyczy trzech podstawowych elementów takich, jak:

5) technologia, dla której najważniejsze są kwestie: jak dobrze sprzęt, sieci oraz elementy związane z konfiguracją działają,
6) proces, gdzie podstawowe elementy oceny to: prawidłowość procedur, technik, algorytmów działań oraz ich realizacja,
7) zawartość, dla której najważniejsze pytania to: poprawność zbioru wybranych zasobów, jego reprezentacja, organizacja, uporządkowanie, zarządzanie [Saracevic 2000].

Podsumowując analizy przygotowywane dla DELOS wymienić należy główne elementy oceny: 1) cyfrowych zbiorów, źródeł; 2) wyboru/selekcji danych, środków przekazu; 3) rozmieszczenia danych, dystrybucji, związków między elementami, odsyłaczami (linkami); 4) organizacji, struktur, przechowywania danych; 5) interpretacji, przedstawienia danych, metadanych; 6) zarządzania; 7) ochrony i trwałości; 8) dostępu; 9) sieci; 10) rozmieszczenia; 11) interfejsów, interakcji; 12) wyszukiwania; 13) usług; 14) rozpowszechniania, dostarczania; 15) pomocy, skierowywania do danych; 16) wykorzystania źródeł przez użytkowników, tworzenia społeczności wokół bibliotek cyfrowych; 17) bezpieczeństwa, prywatności, polityki dostępu, prawnych aspektów, licencji; 18) personelu, pracowników bibliotek cyfrowych; 19) kosztów, ekonomii; 20) integracji, współpracy z innymi zasobami, bibliotekami, albo usługami.

Wizualizację w układzie hierarchicznym uwzględnionych przez Saracevica mierników oceny jakości biblioteki cyfrowej w różnorodnych paradygmatach, stosując się

do odmiennych konsytuacyjnych stanów, w których znajduje się biblioteka cyfrowa i jej użytkownicy przedstawiono na rysunku 5. Należy zauważyć, że wyselekcjonowane kryteria odnoszą się do nurtu badań pragmatycznych i systemowych, chociaż coraz istotniejszy staje się dla Saracevica także kierunek badań kognitywnych, uwzględniających emocjonalne, kulturowe i antropologiczne podejście do analizowania tych obiektów.

Kryteria ewaluacji bibliotek cyfrowych Ying Zhang

Rozwój badań nad bibliotekami cyfrowymi, wypracowanie bardziej komplementarnego spojrzenia na te obiekty oraz opracowanie nowych kryteriów oceny jakości w oparciu o paradygmaty kognitywne, socjologiczne, naukowe, systematyczne znalazło swoje uwieńczenie w pracy Ying Zhang (2010), która stworzyła holistyczny model analizy tych swoistych systemów informacyjno-wyszukiwawczych. Zhang zwraca przede wszystkim uwagę na kategoryzację interesariuszy, wszystkich grup odbiorców oraz osób zaangażowanych w tworzenie bibliotek cyfrowych. Typizacja dotyczy tzw. użytkowników końcowych (*users*, USR), badaczy (*researchers*, RES), bibliotekarzy (*librarians*, LIB), zarządzających biblioteką cyfrową (*developers*, DEV) oraz administratorów (*administrators*, ADM). W odniesieniu do powyższych zbiorowości wyznaczono kluczowe probierze oceny w stosunku do obszarów. Kryteria oceny jakości dotyczą:

a) kontekstu, oznaczającego zrównoważony rozwój, współpracę i wsparcie w zakresie zarządzania zasobami,
b) użytkownika, dla którego najistotniejsza jest wydajność systemu, pomyślne wyszukiwanie informacji oraz satysfakcja,
c) usług charakteryzowanych poprzez takie mierniki jak: dostępność, przydatność, niezawodność, łatwość i prostota wyszukiwania, szybkość reakcji oraz integralność informacji,
d) interfejsu, który powinien być zrozumiały, skuteczny i spójny,
e) zasobów cyfrowych, czyli zawartości, wyrażonej przez dostępność, bezbłędność, przydatność dla użytkowników oraz relewancji w stosunku do ich potrzeb informacyjnych,
f) technologii, która powinna być niezawodna i łatwa w obsłudze.

Kryteria zostały zwizualizowane przez Zhang w postaci koncentrycznych kół. W centrum każdego z nich umieszczono konstytutywne kryteria dla wszystkich wyszczególnionych grup interesariuszy oraz mierniki mapowania dla poszczególnych grup odbiorców biblioteki cyfrowej. Każdy z pierścieni tworzących model ewaluacyjny dla określonego obszaru oceny jakości zawiera kryterium, które stanowi zasadniczy probierz dla co najmniej jednej grupy użytkowników. Liczba koncentrycznych okręgów wyraża stopień rozbieżności oceny pomiędzy grupami. Oznacza to, że im więcej pierścieni, tym większa dysharmonia pomiędzy grupami użytkowników i sprzeczność w zakresie ewaluacji jakości biblioteki na określonym poziomie. Nie wszystkie gru-

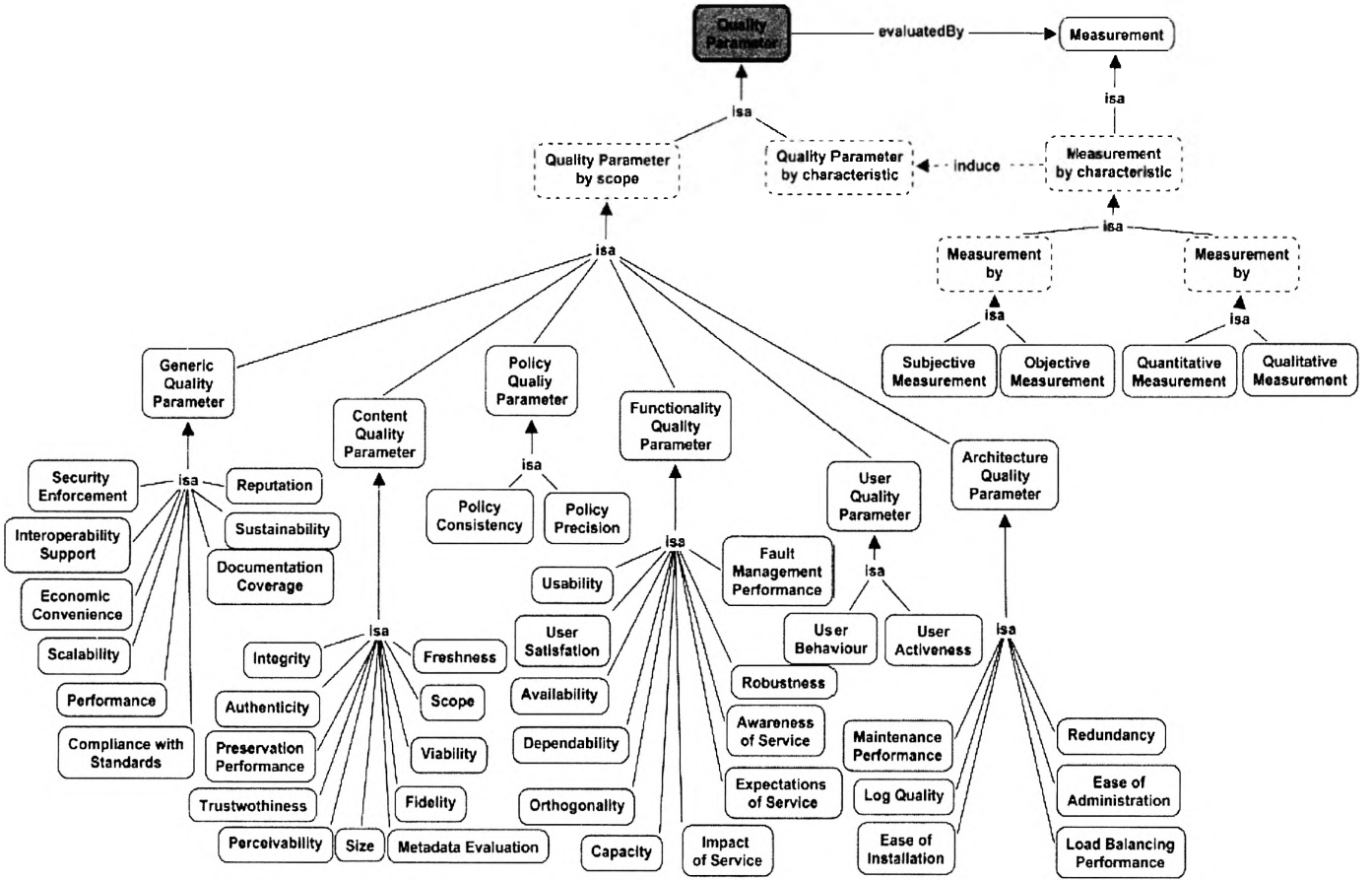
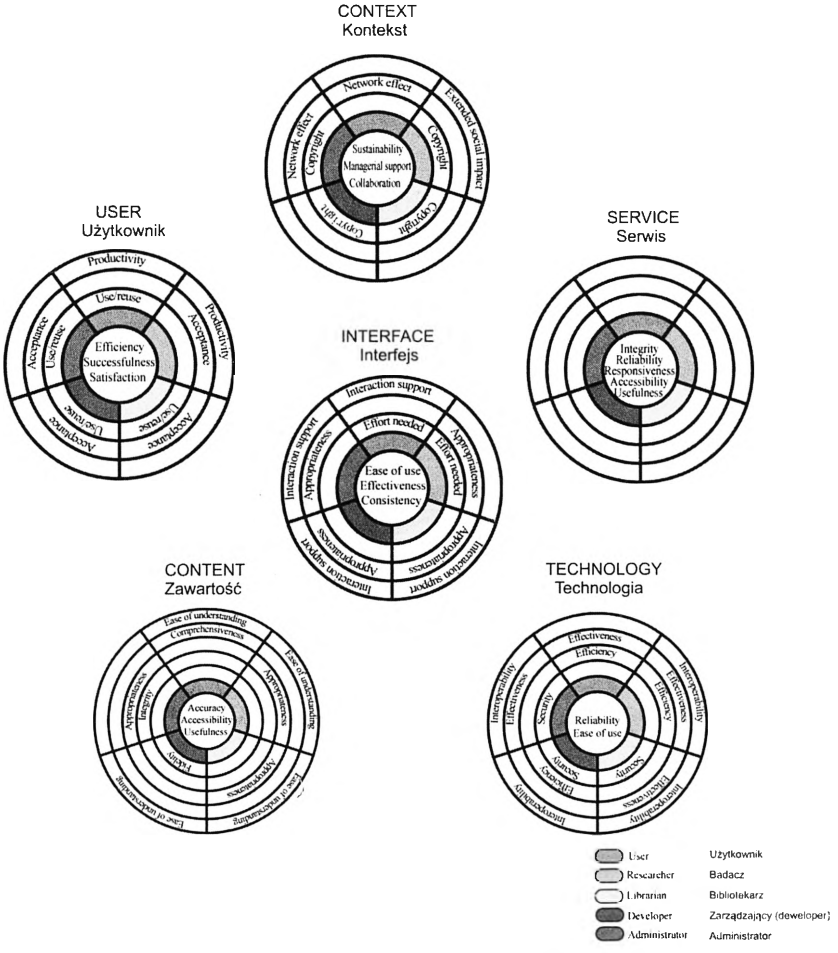


Figure II.2.18. Quality Domain Concept Map: Quality Parameters' Hierarchy

Rys. 5. Kryteria oceny jakości bibliotek cyfrowych w układzie hierarchicznym, uwzględniające różnorodne paradygmaty w badaniach podejmowanych w projekcie DELOS, uwzględniających obszar analiz Tefko Saracevica [Candela, Nardi, dok. elektr., s. 207]

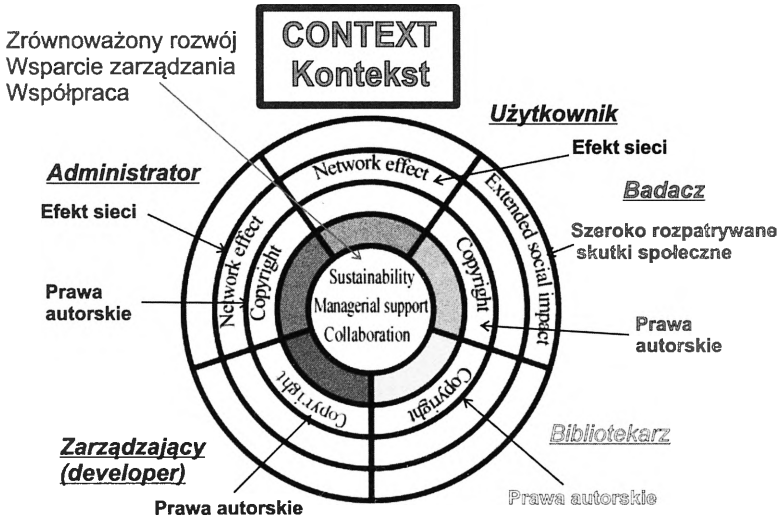
py odbiorców wskazały na ważność poszczególnych kryteriów w zakresie oceny kontekstu, usług, interfejsu, zawartości, technologii, lub społeczności użytkowników. Im bliżej centralnej części modelu, tym większa zgodność poszczególnych interesariuszy w zakresie poszczególnych kryteriów. Na dalszych pierścieniach znajdują się kolejne probierze, które mogą wyznaczać tylko pojedyncze grupy użytkowników. Na rysunku 6 przedstawiono wizualny model holistyczny dla ewaluacji bibliotek cyfrowych opracowany przez Zhang, uwzględniający wszystkie komponenty systemu podczas ewaluacji oraz wytypowane kryteria wraz z grupami interesariuszy.



Rys. 6. Model holistyczny ewaluacji bibliotek cyfrowych zaproponowany przez Ying Zhang [Źródło: Zhang 2010].

Komplementarna ocena jakości biblioteki cyfrowej w obszarze wpływu i konsytuacyjności, w której się znajduje system charakteryzuje szerokie tło, warunki jego organizacji i funkcjonowania. Ewaluacja ta wartościuje szerszy kontekst instytucjo-

nalny, społeczny, kulturowy, ekonomiczny, prawny biblioteki cyfrowej, jej wpływu, implementacji i udoskonalania w społeczeństwie wiedzy. W obszarze tym Zhang wyróżniła najważniejsze mierniki oceny jakości, do których zaliczyła: 1) dostępność i brak ograniczeń ekonomicznych, administracyjnych oraz społecznych w korzystaniu z biblioteki cyfrowej; 2) trwałości systemu, polityki wdrażania, rozwoju i jego utrzymywania; 3) kompatybilności z innymi bibliotekami cyfrowymi, integralności w zakresie organizacji tego typu systemów informacyjno-wyszukiwawczych; 4) istnienia różnorodnych linków, odsyłaczy do innych bibliotek cyfrowych; 5) jasno określonych celów i zadań instytucjonalnych równoważnych z efektywnością organizacji zasobów cyfrowych; 6) efektywności zasobów cyfrowych oraz wydajności społeczności tworzącej bibliotekę cyfrową, rozwoju naukowego i skuteczności działań twórców, bibliotekarzy, administratorów zaangażowanych w tworzenie, doskonalenie i wdrażanie systemu (np. poprzez sprawdzanie liczby publikacji oraz przyznawanych dotacji, grantów, udziału w projektach). Kryteria te w układzie rangowym dla poszczególnych grup interesariuszy, przedstawionym w koncentrycznym modelu oceny jakości kontekstu biblioteki cyfrowej zaprezentowano na rysunku 7.

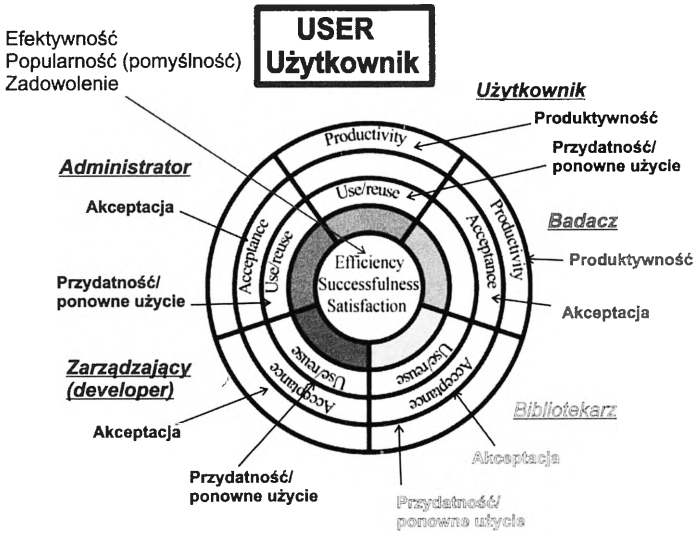


Rys. 7. Kryteria oceny jakości skupione wokół kontekstu biblioteki cyfrowej [Źródło: Zhang 2010].

Podobnie jak Tefko Saracevic lub Iris Xie, której charakterystyka probierzy ewaluacyjnych znajduje się w dalszej części pracy, również Zhang zwróciła szczególną uwagę na konsytuacyjność roli użytkownika w tworzeniu, implementacji i doskonaleniu systemu biblioteki cyfrowej. Ocena jakości zaproponowana z perspektywy użytkownika zawiera się w efektach i implikacjach związanych z zakresem wykorzystywania zasobów cyfrowych, ewentualnych zmian w zachowaniach informacyjnych odbiorców, ich reakcjach poznawczych, afektywnych dysproporcjach oraz zdolności

rozwiązywania problemów i podejmowania działań wspieranych modelami mentalnymi bibliotek. Konsytuacyjność społeczno-kulturowa, analiza efektów korzystania z zasobów biblioteki cyfrowej powinna być także uwzględniona w opracowywanych probierzach jakości tego systemu. Wśród najważniejszych kryteriów ewaluacyjnych w obszarze użytkownika Zhang wymieniła: 1) stopień akceptacji biblioteki cyfrowej; 2) kompetencje informacyjne (*information literacy*) charakteryzujące użytkowników oraz ich rozwój i doskonalenie do zastosowania w analizie rzetelności informacji i zasobów cyfrowych; 3) implikacje w zakresie procesu uczenia się i krytycznego myślenia przez wykorzystywanie biblioteki cyfrowej; 4) efektywności zachowań informacyjnych; 5) satysfakcji użytkowników i pozytywnych afektywnych stanów podczas podejmowanych procesów informacyjnych; 6) efektywnych efektów zachowań informacyjnych, prawidłowego procesu wyszukiwawczego oraz konsytuacyjności czasu podczas podejmowanych zachowań informacyjnych w bibliotece cyfrowej; 7) stopnia eksploatacji systemu informacyjno-wyszukiwawczego, co oznacza m.in. liczbę logowań w sesji, ilość przeglądniętych i pobranych dokumentów oraz zwiększającej się liczby nowych użytkowników biblioteki cyfrowej [Zhang 2007b].

Probierze jakości biblioteki cyfrowej, które dotyczą obszaru użytkowników, a więc wszystkich osób zaangażowanych w tworzenie, wdrażanie, doskonalenie, organizowanie i korzystanie z zasobów, które zostały rankingowane w odniesieniu do określonej grupy tych interesariuszy zostały przedstawione na rysunku 8.

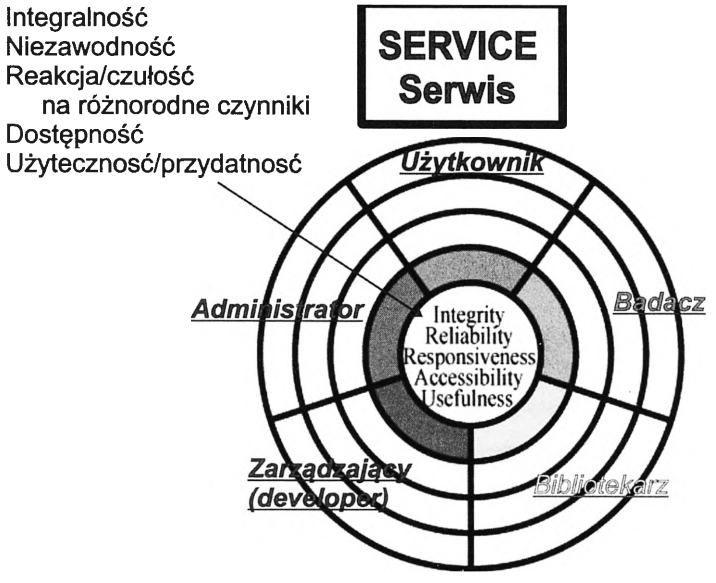


Rys. 8. Kryteria oceny jakości skupione wokół użytkowników biblioteki cyfrowej [Źródło: Zhang 2010].

W procesie parametryzacji oceny usług biblioteki cyfrowej określone zostają zarówno serwisy oferowane przez ten obiekt traktowany jako tradycyjna organizacja, jak

i wszystkie nowe i unikalne możliwości, serwisy realizowane w środowisku cyfrowym. Dotyczą one np. usług referencyjnych, organizacji i zaleceń w zakresie terminologii, SSD *Solid State Drive*–selektywnego upowszechniania dokumentu w pamięci urządzenia, czasem szkoleń oraz rozmaitych kursów standardowych oraz e-learningowych. Atrybuty wyselekcjonowane przez Zhang związane są zwykle z oceną jakości samego serwisu biblioteki cyfrowej i dotyczą: 1) dostępności do usług, niwelowania przeszkód i barier; 2) satysfakcji użytkowników z proponowanych usług oraz poniesionych ewentualnie kosztów dostępu; 3) empatii osób organizujących usługi – ich kultury osobistej i uprzejmości w stosunku do odbiorcy informacji cyfrowej; 4) wyłapywania różnic pomiędzy oczekiwaniami użytkowników a indywidualna percepcją i realizacją zadań; 5) informacji o ponownym korzystaniu z danej biblioteki cyfrowej oraz wiedzy o poziomie satysfakcji odbiorców; 6) niezawodności usług i serwisu oraz stopnia zaufania w zakresie funkcjonowania biblioteki cyfrowej; 7) czasu reakcji pracowników biblioteki cyfrowej – administratorów, personelu organizującego serwis na ewentualne zgłoszenia od użytkowników; 8) stopnia wykorzystania zasobów cyfrowych w określonym czasie.

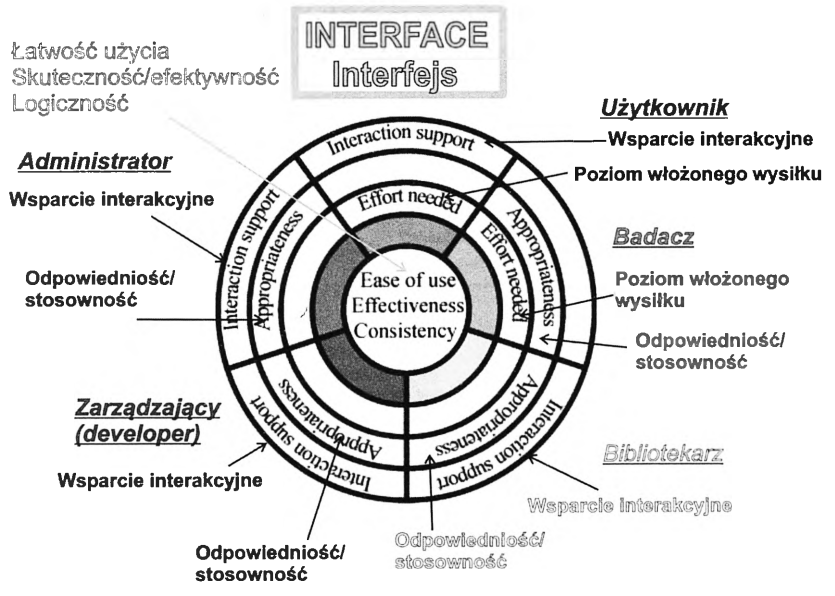
Na rysunku 9 przedstawiono najważniejsze kryteria oceny jakości usług biblioteki cyfrowej. Wszyscy wskazani przez Zhang interesariusze wytypowali probierze takie jak integralność, niezawodność, reakcja na zapytania i zachowania użytkowników, dostępność oraz użyteczność za najbardziej reprezentatywne i istotne dla wszystkich grup osób zaangażowanych w organizację biblioteki cyfrowej. Stąd kryteria te znajdują się w samym centrum koncentrycznego modelu.



Rys. 9. Kryteria oceny jakości skupione wokół usług oferowanych w ramach organizacji biblioteki cyfrowej [Źródło: Zhang 2010].

Wśród istotnych modułów współtworzących system biblioteki cyfrowej wielu autorów, w tym także Ying Zhang wymienia interfejs użytkownika. Jako najważniejsze mierniki oceny jego jakości wyszczególnia ona: 1) efektywność i skuteczność podczas procesu wyszukiwania informacji w zasobach cyfrowych; 2) relewantność w stosunku do wiedzy i doświadczenia użytkownika, jego modelu mentalnego biblioteki cyfrowej, z uwzględnieniem zsynchronizowania z potrzebami i zachowaniami informacyjnymi interesariuszy; 3) zgodności z zasadami tworzenia graficznego interfejsu użytkownika. Zhang wyeksponowała ponadto probierze oceny jakości konkretnych interfejsów użytkowników, do których zaliczyła: a) estetykę i atrakcyjność oraz poziom satysfakcji odbiorców; b) stosowność i adekwatność do wiedzy, zachowań i potrzeb informacyjnych użytkowników; c) spójności, ujednolicenia w zakresie zastosowanej wizualizacji, czyli grafiki, koloru, układu, rozmieszczenia informacji na stronie, czcionki, tła, terminologii; d) intuicyjności interfejsu, łatwości i użyteczności; e) skuteczności w realizacji potrzeb i zadań użytkowników; f) wydajności i oszczędności czasu odbiorców; g) zaangażowania użytkownika w procesie interakcji z systemem oraz utworzenia modelu mentalnego stosowanej biblioteki cyfrowej; h) zdolności obsługi interfejsu i umiejętności niwelowania błędów; i) możliwości personalizacji systemu przez użytkownika oraz j) możliwości podtrzymywania procesu interakcji człowiek-komputer podczas sesji wykorzystywania biblioteki cyfrowej.

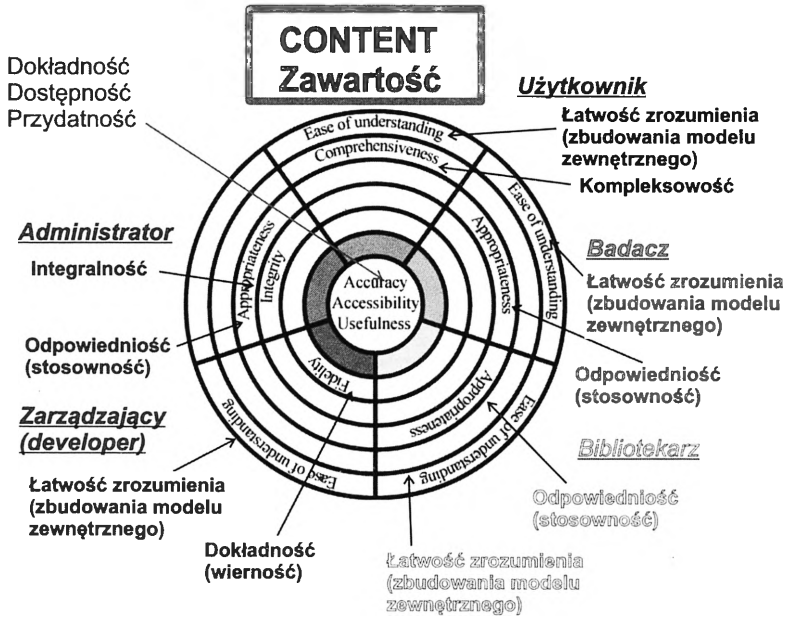
Na rysunku 10 zaprezentowano najważniejsze probierze oceny jakości dla domeny interfejsu wraz z hierarchia wskazań wśród użytkowników, zarządzających deweloperów, administratorów, bibliotekarzy i badaczy.



Rys. 10. Kryteria oceny jakości skupione wokół interfesju biblioteki cyfrowej [Źródło: Zhang 2010].

Ocena jakości w obszarze obiektu dotyczy przede wszystkim komponentu zawartości, opracowania i tworzenia treści zasobów cyfrowych, indeksowania, organizacji i prezentacji informacji w bibliotece cyfrowej. Zhang wyselekcjonowała tu probierze dotyczące m.in.: 1) precyzji, dokładności, czyli zakresu w jakim niwelowane są ewentualne błędy, 2) relewancji w stosunku do potrzeb informacyjnych określonej grupy użytkowników, stopnia wykorzystania w danej dyscyplinie oraz doboru zasobów dziedzinowych, 3) wiarygodności i fachowości zasobów, które powinny być organizowane przez specjalistów w określonej dziedzinie, 4) transparentności, przejrzystości informacji, jej sumaryczności i esencjonalności, 5) wszechstronności kolekcji cyfrowej i jej kompletności, 6) klarowności informacji, 7) dokładności i wierności w stosunku do oryginalnych zasobów posadowionych w wersji cyfrowej, 8) zakresu i informacyjności zasobów, 9) skalowalności, swoistości i możliwości komplementarnego analizowania zasobów w odniesieniu do potrzeb użytkowników, 10) aktualizacji, 11) przydatności.

Poszczególne kryteria jakości w odniesieniu do zawartości biblioteki cyfrowej, wraz z przyporządkowaniem ich do określonych grup użytkowników, wraz ze wskazaniem na najistotniejsze dla tychże społeczności mierników przedstawiono na rysunku 11.

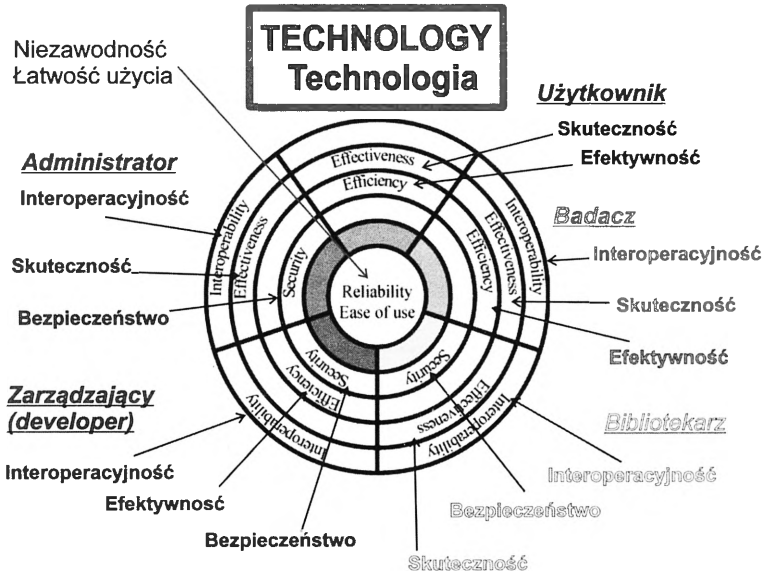


Rys.11. Kryteria oceny jakości skupione wokół zawartości, zasobów biblioteki cyfrowej [Źródło: Zhang 2010].

Obszar ewaluacji skupiony wokół technologii zastosowanej do stworzenia biblioteki cyfrowej jest skoncentrowany przede wszystkim wokół oceny sprzętu i oprogramowania. Analizie poddana została także metodologia selekcji zasobów oraz ekspansja systemu technologicznego. Wśród probierzy stosowanych w tej domenie znajdują się elementy oceny: 1) adekwatnego sprzętu i oprogramowania do zapisu, przetwarzania

i prezentacji informacji w środowisku cyfrowym; 2) wygody korzystania z zastosowanego oprogramowania i systemu technologicznego; 3) kosztów oraz relacji w stosunku do liczby zasobów ludzkich oraz finansowych koniecznych do zaimplementowania nowego środowiska biblioteki cyfrowej; 4) poziomu i jakości informacji cyfrowej; 5) skuteczności zarówno technologii, oprogramowania i sprzętu; 6) wydajności i oszczędności (czasu); 7) równowagi sprzętowo-programowej, elastyczności w stosunku do różnych, zmiennych środowisk cyfrowych; 8) kompatybilności różnorodności systemu technicznego, jego interoperacyjności z odmiennymi technologiami; 9) niezawodności systemu oraz 10) bezpieczeństwa danych [Zhang 2007b].

Wyeksponowane kryteria oceny jakości w zakresie technologii zaaplikowanej do utworzenia biblioteki cyfrowej oraz ich intensywność użycia w różnych grupach użytkowników i ich rankingowaniem zostały przedstawione na rysunku 12.



Rys. 12. Kryteria oceny jakości skupione wokół technologii, systemu, sprzętu i oprogramowania biblioteki cyfrowej [Źródło: Zhang 2010].

*

Podsumowanie dla kryteriów Ying Zhang i Tefko Saracevica

W 2007 r. pojawiła się praca Ying Zhang i Tefko Saracevica [2007a], w której autorzy dokonali scalenia kryteriów wybieranych przez siebie, w odniesieniu do 6 głównych obszarów badań, zidentyfikowanych przez Zhang, a opisanych i wizualizowanych poprzednio. Stąd analiza tych obszarów oraz opis elementów zostaną w tym miejscu pominięte, zwłaszcza, że w pracy Ying Zhang zostały one jeszcze rozbudowane, a przede wszystkim poklasyfikowane hierarchicznie (w wizualizacji – poprzez odległość od centrum koła). Natomiast sama tabela wraz z wykazem przedstawia się następująco:

Tabela 6. Wykaz kryteriów wyselekcjonowanych przez Tefko Saracevica i Ying Zhang.

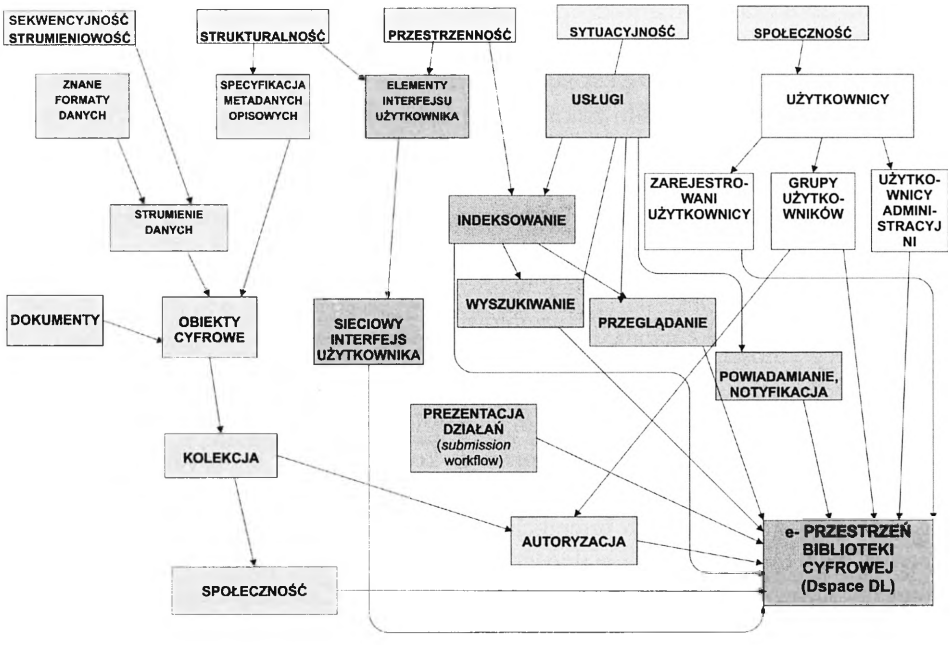
Zawartość	Technologia	Interfejs	Proces / serwis	Użytkownik	Kontekst
kompletność, wielkość	czas reakcji	atrakcyjność	uczenie się, wysiłek/czas, wsparcie, wygoda	zadowolenie, sukces	dopasowanie instytucjonalne, użyteczność
sprawozdanie, częściowe pokrywanie się	czas przetwarzania	logiczność	łatwość wykorzystania	relewancja, użyteczność rezultatów	wydajność & wpływ na członków społeczności
sprawdzalność, dokładność	prędkość	przedstawienie pojęć – metki wiadomości	zagubienie (zamieszanie)	wpływ (impact), wartość	zrównoważoność
ważność, autorytatywność	pojemność	obrazowość, atrakcyjność	ukończenie (realizacja zadania)	jakość doświadczenia	interoperacyjność
dostateczność, rozmaitość	ładowalnosc	stosowność	trudność interpretacji	bariery, drażliwość	prawa (zarządzania), prawo autorskie przestrzeganie przepisów
'informacyjność' (*informativeness*)	dostępność	konsekwencja	pewność rezultatów	preferencje	użyteczność organizacyjna…
świeżość	skuteczność	łatwość wykorzystania	współczynnik błędu	efekt uczenia się	
dostępność, osiągalność	wydajność	wysiłek	wrażliwość	wydajność	
złożoność struktury organizacyjnej	zgodność	wykrywanie błędów, personalizacja…	niezawodność…	użycie/ powtórne użycie…	
przezroczystość, jasność	jakość				
osiągnięcie zrozumienia …	solidność				
	odporność…				

*

Model biblioteki cyfrowej 5S

Model 5S jest specyficznym modelem biblioteki cyfrowej, wykorzystanym także dla analiz przygotowanych dla DELOS, stąd opisany zostanie tutaj pokrótce. Ten komplementarny model opiera się na wizualizacji uwzględniającej kryteria ewaluacji oraz mentalne wyobrażenia interesariuszy: twórców, bibliotekarzy, administratorów, odbiorców, analizując bibliotekę cyfrową z różnorodnych perspektyw. Struktura ta zamyka się w pięciu najważniejszych obszarach: *Streams, Structures, Spaces, Scenarios, Societies*, określając poszczególne elementy tego systemu: strumieniową, strukturalną, przestrzenną, sytuacyjną oraz społecznościową płaszczyznę oceny jakości biblioteki cyfrowej. Wszystkie te ramy mogą stanowić osobne probierze oceny jakości, według których można kompleksowo wartościować bibliotekę cyfrową, uwzględniając szerokie spektrum interesariuszy, opracowując, doskonaląc lub implementując nową strukturę. Model strumieniowy określa zawartość systemu. Wspólną cechą różnych form plików w kolekcji jest to, że wszystkich mogą być reprezentowane jako sekwencje bardziej podstawowych komponentów (np. znaków). Model strukturalny charakteryzuje architekturę, wszelkie rodzaje konfiguracji obiektów, standardów metadanych, właściwości kolekcji i katalogów, a także narzędzi zarządzania i organizacji wiedzy. W mo-

delu przestrzennym zawarte zostały kryteria oceny jakości charakteryzujące interfejs użytkownika, paradygmaty wyszukiwania oraz procesy indeksowania i katalogowania zasobów. Model sytuacyjny dotyczy przede wszystkim oceny wszystkich usług oferowanych przez bibliotekę cyfrową, wszelkiego rodzaju działania, interakcje, politykę oraz warunki funkcjonowania. W modelu społecznościowym wyróżnić można mierniki oceny jakości dotyczące zarówno technologicznego aspektu (zbioru komputerów, jak i komponentu ludzkiego, czyli społeczności użytkowników oraz wszelkich relacji zachodzących pomiędzy nimi. Relatywne pola w modelu, charakterystyczne dla ewaluacji jakości zostały mogą być zawarte czterech domenach: funkcjonalności (*functionality*), zawartości (*content*), struktury (*structures*) oraz użytkowników (*user*) i wyznaczać określone kryteria jakości dla poszczególnych zakresów [Gonçalves i in. 2004].



Rys. 13. Model biblioteki cyfrowej uwzględniający strukturę 5S
[Źródło: Murthy i in. 2007].

*

Rozwój badań nad bibliotekami cyfrowymi, ich kompleksowość, uwzględnienie różnorodnych paradygmatów, a szczególnie zwrócenie uwagi na użytkowników spowodowało wypracowanie probierzy ewaluacyjnych z perspektywy odbiorców tych systemów. Kryteria te zostały zaproponowane przez m.in. **Hong Iris Xie** (2008). Zestaw mierników uwzględnia aspekty rozpatrywania z jednej strony dostępu oraz wyszukiwania w zasobach cyfrowych, zaś z drugiej oceny kolekcji, organizacji i serwisów udostępnionych w bibliotekach cyfrowych [Chowdhury 2010]. Analiza struktury biblioteki cyfrowej, uwzględniającej komponent ludzki, rolę użytkownika w tworze-

niu modelu projektowego systemu, skupia się często na użyteczności i funkcjonalności obiektu[1]. Wskaźnikowi funkcjonalności przypisywane są następujące atrybuty: przyswajalność (łatwość nauki), efektywność, możliwość zapamiętania, błędy, satysfakcja. Mają one silny związek z modelem mentalnym biblioteki cyfrowej jaki tworzy użytkownik na podstawie wiedzy oraz doświadczenia, co zostało szerzej wyjaśnione w innej części tej książki, pracy Moniki Krakowskiej dotyczącej reprezentacji umysłowych zasobów cyfrowych, systemów informacyjno-wyszukiwawczych. Wśród analizowanych probierzy jakości bibliotek cyfrowych uwzględniających optykę odbiorcy oraz konsytuacyjność technologiczną często wymienia się kryteria: zastosowania bibliotek cyfrowych w porównaniu do innych systemów, przeznaczenia, celu i ważności wykorzystywania, charakteru zastosowania, a także satysfakcji użytkownika [Bishop; Neumann; Star; Merkel; Ignacjo; Sandusky 2000], edukacyjnej specyfiki zasobów i sposobu ich użycia, wydajności oprogramowania, estetyki i prostoty interfejsu [Marchionini 1998; Hill 2000], a wreszcie potrzeb i oczekiwań w stosunku do biblioteki cyfrowej. Komponenty uwzględniające afektywne i kognitywne interakcje, jakie podejmuje użytkownik w środowisku informacyjno-wyszukiwawczym oraz konsytuacyjność infrastruktury skupionej wokół interesariuszy zostały wyselekcjonowane przez Hong Iris Xie w drodze analizowania bibliotek cyfrowych z perspektywy indywidualnego odbiorcy i jego zachowań informacyjnych. Jak zauważa Xie, mierniki ewaluacji bibliotek cyfrowych często odnoszą się do typowych probierzy, które są stosowane do oceny jakości tradycyjnych bibliotek w obszarze: kolekcji (jej celu, zakresu, wartości, relewancji, odbiorców, kosztów, formatów, zarządzania zasobami, ochrony); informacji (precyzyjności, trafności, odpowiedniości, odsyłaczy, reprezentacji, unikalności, porównywalności i sposobu prezentacji); zastosowania (dostępności, udostępniania, możliwości wyszukiwania, zastosowania); standardów. W obrębie tradycyjnych procesów wyszukiwania (*information retrieval*) konieczne jest uwzględnienie podczas ewaluacji kryteriów relewancji (precyzji i niwelowania błędów), satysfakcji i wyników, zaś w ramach standardowych interakcji człowiek-komputer oraz interfejsu probierze dotyczą użyteczności, funkcjonalności, poprawności i ewentualnych niepowodzeń. Postulaty zarówno Hong Iris Xie, Tefko Saracevica lub Gobindy Chowdhury dotyczą jednocześnie uwzględnienia ogólnego kontekstu i wpływu bibliotek cyfrowych na użytkowników i całą społeczność [Chowdhury 2010; Saracevic 2004].

Specyfika kategorii wskaźników opracowanych przez Xie, polega przede wszystkim na wskazaniu tych elementów, które są najistotniejsze dla użytkowników. Zostały one oparte przede wszystkim na badaniu subiektywnych stymulantów, indywidualnych potrzeb informacyjnych oraz doświadczenia odbiorców, co przede wszystkim decyduje o całościowym odbiorze biblioteki cyfrowej. Estetyka strony serwisu, możliwość łatwej nawigacji, kompletność zasobów, znaczenie kulturowo-społeczne, stanowią recepcję atrybutów oceny projektu i decydują o pozytywnym odbiorze samego systemu, ale także o odpowiednich zachowaniach informacyjnych, bez elementów frustrujących. Zarówno satysfakcja, jak i emocjonalne podejście do systemu, kryte-

[1] Co pozwala na rekomendowanie zasad projektowania biblioteki cyfrowej, uwzględniając to kryterium, lub udoskonalanie już istniejących repozytoriów [Xie 2006, s. 435].

rium interpretacji modeli mentalnych, zadowolenie użytkowników, konieczność pozyskiwania ich opinii pojawiają się również wśród mierników jakości bibliotek cyfrowych, a które jak twierdzi Xie [Xie 2008, s. 1356-1357], należy uwzględniać podczas kompleksowej oceny tych obiektów.

Wykaz kryteriów stosowanych przez grupę użytkowników końcowych eksplorowanych przez Xie oraz porównanie ich z miernikami wyeksponowanymi przez badaczy[2], a także te, które zdefiniowane przez innych autorów, zostały zaprezentowane w poniższej tabeli. Dla odbiorców szczególnie istotne są kryteria dotyczące zawartości, funkcjonalności interfejsu oraz jakości samej kolekcji. Niezbędne okazują się wskaźniki jakościowe dla serwisu, wydajności systemu oraz ogólna użyteczność biblioteki cyfrowej.

Należy zwrócić uwagę, że nie wszystkie mierniki są ważne zarówno dla samych odbiorców, jak i badaczy ewaluujących biblioteki cyfrowe. W przypadku funkcjonalności naukowcy nie wyeksponowali m.in. kryteriów estetyki i wydajności, wyszukiwania, nawigacji, ale także wyłącznych usług proponowanych przez biblioteki cyfrowe, a także zwrotnej informacji pozyskiwanej od odbiorców. Z kolei sami użytkownicy końcowi nie wymienili probierzy dotyczących formatów, standardów, kosztów organizacji biblioteki cyfrowej, zastosowania zasobów cyfrowych lub paradygmatu społecznego, czyli kompleksowego oddziaływania systemu informacyjnego na zbiorowość i wymiar kulturowo-społeczny.

Tabela 7. Kryteria oraz subkryteria oceny jakość poszczególnych elementów biblioteki cyfrowej wyszczególnione podczas badań Hong Iris Xie. [Źródło: Xie, 2006, s.443].

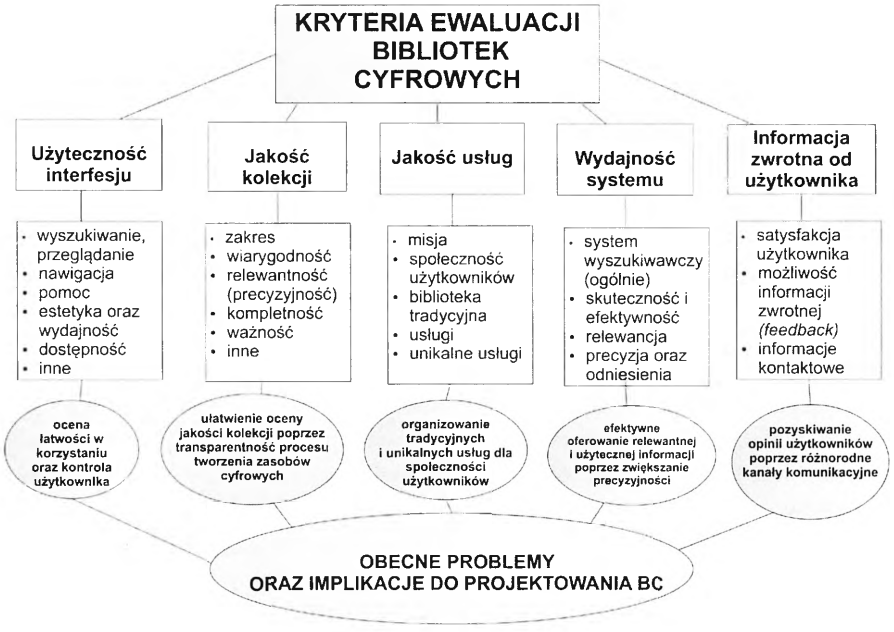
Kryteria (*criteria*)	**Subkryteria (*subcriteria*)**	**Użytkownicy końcowi (*users*)**	**Badacze (*researchers*)**	**Kryteria i subkryteria wskazane w innych badaniach niż analizy Iris Xie**
Użyteczność, Funkcjonalność (*usability*)	Funkcjonalność (ogólnie)	X	X	–
	Funkcjonalność interfejsu	X	–	–
	– użyteczność	–	X	X
	– wyszukiwanie i przeglądanie;	X	–	–
	– nawigacja;	X	–	–
	– pomoc;	X	–	X
	– estetyka i wydajność;	X	–	X
	– dostępność;	X	–	–
	– błędy	X	X	–
Jakość kolekcji	Zawartość (ogólnie)	X	X	X

[2] Mierniki jakości bibliotek cyfrowych z różnych perspektyw, zaproponowane przez badaczy zostały przedstawione w publikacjach m.in. Tefko Saracevica: *Digital library evaluation: toward an evolution of concepts*; Gobindę Chowdhury *From digital libraries to digital preservation research: the importance of users and context*; Garego Marchioniniego: *Evaluation Digital libraries: a longitudinal and multifacet view*; Tefko Saracevica i Lisę Covi: *Challenges for Digital library evaluation.*

	Jakość – zakres – wartość – dokładność – kompletność – powszechność – koszty – format – zarządzanie kolekcją – ochrona – wyjątkowość	X X X X X X – – – – –	X X X X X X X X X X X	X X – X – – X – – – –
Jakość serwisu	Misja, zadania BC	X	–	–
	Określona społeczność użytkowników (przeznaczenie)	X	X	X
	Serwisy BC o charakterze biblioteki tradycyjnej	X	X	X
	Unikalne serwisy BC	X	–	–
Efektywność systemu	Wyszukiwanie informacji (*information retrieval*) – ogólnie	X	X	X
	Efektywność i wydajność	X	X	X
	Relewancja	X	X	X
	Precyzja i usuwanie błędów	X	X	X
	Zastosowanie	–	X	X
	Wynik (wyszukiwania)	–	X	–
Sprzężenie zwrotne, informacja zwrotna od użytkowników	Satysfakcja użytkownika	X	X	X
	Informacja zwrotna od użytkownika	X	–	–
	Informacje kontaktowe	X	–	–
Oddziaływanie, wpływ (*impact*)	Społeczne oddziaływanie (*social impact*), szerokie oddziaływanie	–	X	–
	Wpływ na społeczność określonych użytkowników	–	–	X
ne	Standardy	–	X	–

Jednym z najistotniejszych, wręcz kluczowych kryteriów oceny jakości z punktu widzenia użytkowników jest funkcjonalność, użyteczność biblioteki cyfrowej. Odbiorcy wskazują na istotę prostoty nawigacji, intuicyjność i praktyczność interfejsu, jako swoistego łącznika pomiędzy systemem a klientem [Xie 2006, s. 440]. Podobnie ważną miarą jakości i użyteczności biblioteki cyfrowej jest funkcja przeglądania i wyszukiwania informacji. Łatwość opanowania, nauczenia się systemu (*learnability*), satysfakcja, efektywność, klarowność, przejrzystość w zapamiętywaniu zadań oraz ocena błędów wpływa intensywnie na powstające modele mentalne użytkowników, co wiążę się również z zastosowaniem samego systemu informacyjno-wyszukiwawczego i przenoszeniem reprezentacji umysłowej na coraz nowsze struktury bibliotek [Zhang 2010]. Wśród kluczowych kryteriów Hong Xie wymienia dodatkowo probierze ewaluacji jakości kolekcji, usług wynikających zarówno z działalności i funkcji tradycyjnej biblioteki, jak i specyficznych, wyłącznych dla cyfrowego jej odpowiednika. Istotne dla udoskonalenia systemów jest pozyskiwanie opinii użytkowników, co wpływa na niwelowanie problemów, precyzję w implikowaniu zasobów cyfrowych, a także zwiększenie społeczności odbiorców.

Wszelkie nieścisłości, na które wskazywali użytkownicy w badaniach Xie, jak np. brak prostoty w korzystaniu z biblioteki cyfrowej, zbytnia powierzchowność w organizacji informacji i zasobów cyfrowych, brak prowadzenia dokładnego i szczegółowego wyszukiwania, brak możliwości większej kontroli i personalizacji systemu przez bardziej zaawansowanych odbiorców, znikoma informacja dotycząca aspektów legislacji i praw autorskich, a także częsty brak mechanizmów świadczących o miarodajnej i precyzyjnie zorganizowanej kolekcji, zmniejszają zarówno jakość obiektu, jak i dość istotny poziom satysfakcji i zadowolenia interesariuszy. Kryteria uwzględnione przez Xie wydają się być znaczące dla dynamiki implementacji bibliotek cyfrowych skoncentrowanych na użytkownikach, personalizacji systemów, rozwiniętej komunikacji pomiędzy interesariuszami oraz zrozumienia zachowań odbiorców, ich preferencji, potrzeb, doświadczenia oraz uwzględnienia reprezentacji mentalnych wpływających na odbiór tych bibliotek. Probierze oceny jakości bibliotek cyfrowych Hong Xie, wskazane przez użytkowników oraz konsekwencje wynikające z ewaluacji, wpływające na projektowanie oraz niwelowanie wszelkich barier i problemów w korzystaniu z tych systemów zostały przedstawione na rysunku.



Rys.14. Ewaluacja bibliotek cyfrowych: kryteria oraz implikacje dla projektowania [Za: Xie 2006].

5. Ocenianie / zestawianie bibliotek cyfrowych (porównywanie)

Próbę wyznaczenia kryteriów do zestawiania ze sobą bibliotek cyfrowych podjął Robert J. Sandusky (2002). Porównywanie wielu obiektów nie jest łatwe, a określenia 'najlepsza biblioteka cyfrowa' w ustach specjalistów brzmią niejednoznacznie. Nie wiadomo bowiem, jakie było kryterium oceny: wielkość zbiorów, ich unikatowość, prostota i intuicyjność interfejsu, logiczność systemu, a może nastawienie na konkretny typ użytkownika. Stąd Sandusky opracował szeroką listę atrybutów, które pomogą zidentyfikować podobieństwa, różnice oraz wzorce istniejące w bibliotekach cyfrowych. Podstawowe parametry to:

1. odbiorcy – analiza docelowych i faktycznych użytkowników,
2. instytucja – jako instytucja sprawcza, odpowiedzialna za profil biblioteki, jej zasoby, itp. – tu w szczególności uczelnie wyższe, firmy lub inne podmioty – sponsorzy,
3. dostęp – kto może korzystać z BC, na jakich warunkach, oraz jakie jest wsparcie w dostępie do treści,
4. zawartość – jakie informacje(np. dokumenty), zawarte są w bibliotece,
5. usługi – dotyczą możliwości ludzi lub technologii linkujących (systemów komputerowych), pozwalających na do korzystanie z kolekcji oraz odwoływanie się do nich,

6. projekt (design) i rozwój – związane z procesem budowy i utrzymania biblioteki cyfrowej

Szczegółowe objaśnienia poszczególnych atrybutów (parametrów) oraz ich aspektów dla lepszego zobrazowania przedstawione zostały w poniższej tabeli [Janiak; Krakowska 2010a]:

Tabela 8. Atrybuty (parametry) i aspekty oceny jakości zaproponowane przez Roberta J. Sanduskiego [2004].

Atrybut	Aspekt	objaśnienia, hasła główne
odbiorcy	zakres (społeczności)	społeczność użytkowników ograniczona czy nieograniczona Np. dostęp dla wszystkich; dla wybranej/wybranych instytucji; dystrybucja zawartości odbywa się za opłatą chociażby w ograniczonym zakresie; tylko część dostępna jest za darmo (np. dotowana przez reklamę lub rząd)
	spójność (podobieństwo użytkowników)	głównymi użytkownikami są osoby o podobnych zainteresowaniach lub też społeczność użytkowników tworzą różnorodni odbiorcy, w przekroju społecznym podobni do ogółu społeczeństwa Hasła główne: – spójna publiczność składająca się z zamierzonych lub rzeczywistych użytkowników, którzy mają wspólne cechy kluczowe – odbiorcy rozproszeni, posiadający po kilka cech wspólnych (lub tylko przez przypadek wydający się do siebie podobni). Biblioteka cyfrowa przeznaczona dla ogółu społeczeństwa posiada rozproszonych odbiorców.
	dopasowanie	Rzeczywiści użytkownicy są tacy jak docelowi (założeni przez twórców) Hasła główne: – nieprzystosowanie (*mis–fit*) – użytkownicy rzeczywiści posiadają profile, które są niespójne założonym targetem odbiorców – ścisłe przyleganie – profile użytkowników są zgodne z założonymi członkami grupy docelowej
instytucja	zarządzanie i kontrola	kto i co wpływa na podejmowania decyzji w odniesieniu do BC, jej ogólnego projektu, działalności oraz ciągłego rozwoju Hasła główne: – ścisłe zarządzanie i kontrola – w konkurencyjnych, komercyjnych BC, w celu zminimalizowania przecieku planów i przekazywania informacji konkurentom. – otwarte zarządzanie i kontrola – w publicznych bibliotekach cyfrowych
	typ	sponsoring BC przynależy do instytucji publicznej (rządu, uczelni) czy prywatnej (osób prawnych) Hasła główne: – instytucja publiczna – instytucja prywatna
	model ekonomiczny	jak biblioteka cyfrowa jest finansowana i czy będzie podtrzymywana Biblioteka cyfrowa może być obsługiwana tylko z dotacji przekazanych przez instytucje publiczne, fundacje itp., lub też utrzymywana jest z opłat bezpośrednich (płatne) lub pośrednich (reklama) Hasła główne: – granty (dotacje publiczne) – sprzedaż (np. z reklam)

	misja	co jest deklarowanym celem biblioteki – eksperyment opracowany głównie w celu zbadania nowych technik, czy jest produktem przygotowanym w celu zapewnienia niezbędnych informacji i usług na rzecz społeczności użytkowników Hasła główne: – eksperyment / badanie – twór ostateczny
dostęp	model opłat	użytkownicy muszą zapłacić bezpośrednio lub pośrednio za dostęp Opłata może być oparta o abonament miesięczny lub dotyczyć poszczególnych tekstów (opłata za dokument); model opłat może też być mniej bezpośredni np. dostęp jest częścią czesnego lub też wspierany jest przez środki z podatków lub finansowany z przychodów reklamowych. Nie ma jednak prostej korelacji między częstszym wykorzystywaniem a bezpłatnością źródła. Niektóre płatne biblioteki cyfrowe ze względu na swoją unikatowość są najchętniej wykorzystywaną kolekcją przez użytkowników np. specjalistów. Hasła główne: – płatne – bezpłatne
	widzialność	biblioteki cyfrowe są widoczne w Internecie – np. dla wyszukiwarek (czy część lub wszystkie materiały są dostępne za darmo) lub też ich istnienie jest ukryte (przez firewall lub inne technologie zabezpieczeń) Prywatne biblioteki są hostowane w kontrolowanych intranetach, chyba że niektórym użytkownikom przyznawane są szczególne uprawnienia. Publiczne dostępne za pośrednictwem sieci są na ogół widoczne. Hasła główne: – widoczne – niewidoczne / ukryte
	trwałość	trwanie biblioteki cyfrowej jest zagwarantowane na przyszłość lub zapisane w jej misji Hasła główne: – zachowane (*preserved*) – ulotne – nietrwałe – czasowe
	spójność (konsekwencja)	dostęp do zawartości jest wzmocniony przez zastosowanie zasad dobrej organizacji informacji Minimalnie zorganizowane lub niezorganizowane biblioteki mogą zapewnić dostęp użytkownikom jedynie przez wyszukiwanie pełnotekstowe – ich struktura jest często płaska (jeden poziom hierarchii). Dobrze zorganizowane biblioteki zapewniają wiele rodzajów wyszukiwania oraz dostęp przez np. hierarchicznie przygotowane struktury tematyczne. Hasła główne: – zorganizowane – niezorganizowane
zawartość	zakres	zakres treści zawartych w bibliotece cyfrowej Biblioteki o ograniczonej zawartości będą miały informacje na jeden temat lub też zawierać kompleksową informację, odnoszącą się do dowolnego działu ludzkiej wiedzy. Hasła główne: – wszechstronna (ogólna) – ograniczona – tematycznie lub formalnie (*limited*)
	dopasowanie (do potrzeb użytkowników)	treść jest odpowiednia dla zamierzonych lub rzeczywistych użytkowników i ich potrzeb Hasła główne: – nieprzystosowanie (*mis–fit*) – dostosowanie

	organizacja	zawartość biblioteki jest zorganizowana lub nie w sposób ułatwiający wyszukanie i dotarcie do informacji Hasła główne: – zorganizowane – niezorganizowane
	specjalizacja	zawartość jest przygotowana dla specjalistów, konkretnych odbiorców lub też jest ogólna, zakrojona na szeroką skalę Hasła główne: – wyspecjalizowana – ogólna
	zdigitalizowana zawartość	dokumenty powstały w trakcie digitalizacji istniejących materiałów nie–cyfrowych lub też informacje w nich zawarte od początku były cyfrowe (*digital born*) Hasła główne: – *digital born* – dokumenty zdigitalizowane
	źródła	zawartość pochodzi z jednego lub z wielu źródeł Jednoźródłowe biblioteki cyfrowe przedstawiają treści tworzone przez pojedynczy podmiot lub pochodzą z jednego, kontrolowanego procesu (np. dziennik lub czasopismo), natomiast wieloźródłowe mają w swoich zbiorach treści tworzone przez wiele podmiotów. Hasła główne: – jedno źródło – wiele źródeł
	restrykcje	treści znajdują się w domenie publicznej (nie podlegają ograniczeniom w prawach autorskich czy wydawniczych) lub też obwarowane są ograniczonymi (np. CC) lub pełnymi prawami autorskimi Hasła główne: – wolny dostęp – ograniczony dostęp
usługi (serwisy)	natura interakcji	wymiar ten określa sposób, w jaki sposób biblioteka cyfrowa zapewnia lub nie zapewnia wsparcie dla aktywnej pracy Biblioteki, które zapewniają pojedynczy dostęp do informacji (podobny do podstawowego modelu realnej biblioteki) są bierne, natomiast te, które dostarczają narzędzi do analizy lub narzędzi interakcyjnych: użytkownik–użytkownik (lub użytkownik–ekspert) są aktywne. Hasła główne: – pasywny – aktywny
	analityczność	zapewnienie użytkownikowi wsparcia dla oceny lub asymilacji znalezionych informacji W analitycznych bibliotekach znajdują się interaktywne narzędzia wspierające analizę danych lub pozwalające na przechowywanie danych generowanych przez użytkowników. Hasła główne: – analityczne – nieanalityczne
	współpraca	wsparcie dla komunikacji pomiędzy biblioteką cyfrową a użytkownikami Biblioteki oparte na współpracy z użytkownikami często zawierają wbudowany chat lub forum, przygotowane dla wszystkich lub tylko określonych grup interesów (np. dla użytkowników zalogowanych); pozwalają na kontakt użytkownik–użytkownik, pozwalają dookreślać opis obiektów itp. Hasła główne: – współpraca – brak współpracy

	kolekcja	dobór treści i usługi organizacji informacji Biblioteka cyfrowa, która w jasny sposób przedstawia politykę rozwoju kolekcji oraz sposoby indeksowania danych może być kontrolowana, natomiast ta umożliwiająca dodawanie treści bez ich ewaluacji jest biblioteką niekontrolowaną. Hasła główne: – kontrolowana – niekontrolowana
	odnośniki (rekomendacje)	specjalny typ narzędzi (usługi współpracy): możliwość kontaktu z domeną i / lub z ekspertami biblioteki cyfrowej Użytkownicy mogą skontaktować się z kimś za pośrednictwem poczty elektronicznej lub w czasie rzeczywistym np. za pośrednictwem czatu. Hasła główne: – rekomendacje – brak rekomendacji
projekt (design) i rozwój	podejście do *designu*	biblioteka została zaprojektowana w odniesieniu do użytkowników, instytucji, dostępu oraz swojej zawartości Biblioteki cyfrowe eksperymentalne stosują nowe systemy zgodne z najnowszą myślą techniczną: innowacje w samej bazie danych (np. strukturze), sieci, interfejsie, oprogramowaniu, natomiast te społecznie ugruntowane biorą pod uwagę niektóre lub wszystkie kwestie związane z przyzwyczajeniami użytkowników, wymaganiami instytucji, klasycznymi sposobami dostępu oraz prezentacji treści. Hasła główne: – eksperymentalne – społecznie zakotwiczone
	cykl projektowania	Stała, niezmienna – nieiterowana biblioteka cyfrowa jest jednym tworem, który został opracowany np. w trakcie trwania jednego grantu, generuje środków utrzymania się, a więc się nie zmienia, natomiast iterowana podlega ciągłemu przeprojektowaniu w oparciu o metodyczne podejście, które obejmuje opinie z analizy wykorzystania zasobów. Hasła główne: – iterowane – nieiterowane
	konsultacje wśród użytkowników	prowadzenie zamierzonych konsultacji z użytkownikami, które mogą być pomocne w dalszym jej projektowania i rozwoju Wiele BC jest tworzonych i rozwijanych bez bezpośredniego kontaktu z użytkownikami, co zwiększa ryzyko niedopasowania (mis–fit). W innych przypadkach, pełnomocnicy użytkowników (wybrani użytkownicy) mogą konsultować się z projektantami i programistami. Na drugim biegunie jest stała reprezentacja użytkowników, czynnie pracująca nad projektem – pełnoprawni członkowie grupy (np. uczestnictwo w projekcie). Hasła główne: – aktualni użytkownicy – nikt

Jak widać do porównania bibliotek cyfrowych potrzebne są głębokie analizy, gdyż ich różnorodność jest duża, a więc nie da się określić jakości na podstawie samych danych liczbowych (zasoby, użytkownicy itp.). Trudno też jednoznacznie wyznaczyć na podstawie przytoczonej tabeli, czy np. eksperymentalny *design* jest jakościowo lepszy od tradycyjnego. Stąd wymienione aspekty wskazują całą gamę modeli bibliotek, wyznaczając ich najważniejsze parametry, ale nie proponują jednej metody ewaluacji.

Podsumowanie

W opracowaniu tym przedstawiono najważniejsze zestawy kryteriów oceny jakości bibliotek cyfrowych wraz z podstawowymi atrybutami oraz aspektami, które mogą być badane. Zostały one zaprezentowane w odniesieniu do celów badań oraz ich zakresów. Dodatkowo zaprezentowano główne pytania zadawane przy ocenianiu i opisywaniu tych systemów informacyjnych. Należy zwrócić uwagę, że przeprowadzone badania i analizy bibliotek cyfrowych, realizowane w różnorodnych paradygmatach stosowanych w informatologii: pragmatycznym, naukowym, systemowym, cybernetycznym, psychologicznym: kognitywnym, czy afektywnym, powoduje wyznaczenie, wyselekcjonowanie różnorodnych kryteriów, które relacyjnie wspomagają proces ewaluacji systemów informacyjno-wyszukiwawczych. Dostosowywanie probierzy do istniejących bibliotek cyfrowych jest niezwykle istotne i w głównej mierze zależy od przyjętego modelu tego systemu, składającego się zwykle z zasobów, użytkowników (interesariuszy), infrastruktury technicznej oraz projektu z jego intencją, planem, posłannictwem. Wieloaspektowość i możliwość komplementarnego zbadania i oceny jakość biblioteki cyfrowej opiera się na zastosowaniu wielu, odmiennych perspektyw, uwzględnienia konsytuacyjności, w której ona operuje, jest implementowana oraz zastosowania wielostronnych probierzy i mierników ewaluacyjnych. Ma to oczywiście wpływ także na powstawanie prototypów tych systemów, modeli referencyjnych i pierwotnych struktur jak np. przytoczony przykład modelu 5S, od których należy rozpocząć proces tworzenia biblioteki cyfrowej.

Na zakończenie powtórzyć można za Tefko Saracevicem i Ying Zhang, iż ewaluacja bibliotek cyfrowych nie jest niemożliwa, ale jest trudna. Nie ma bowiem do końca przygotowanych generalizacji, nie mamy jeszcze sprawdzonych wielokrotnie w praktyce teorii, jest za to wiele zaproponowanych modeli. W piśmiennictwie, zwłaszcza polskim, poświęconym bibliotekom cyfrowym natomiast problem jakości nie jest tematem często podejmowanym [Zhang; Saracevic 2007a], stąd autorki mają nadzieję, iż niniejsza praca, przedstawiająca tak wiele różnych spojrzeń na jakość bibliotek cyfrowych zapełni trochę tę lukę, a badaczom pozwoli wybrać najbardziej adekwatny do ich celów, badanych obiektów oraz całego otoczenia (zaplecze kulturowe, projektowe itp.) zestaw kryteriów ocen.

Bibliografia

1. *A Framework of Guidance for Building Good Digital Collections – the 3rd edition* (2007) [dok. elektr.] Institute for Museum and Library Services http://framework.niso.org/ [odczyt: 2.02.2011].
2. *A list of quality selection criteria: a reference tool for Internet subject gateways* [dok. elektr.] UKOLN Metadata Group http://www.ukoln.ac.uk/metadata/desire/quality/report-2.html [odczyt: 2.02.2009].
3. American Library Association (1997-2012). *Great Web Sites for Kids Selection Criteria.* [dok. elektr.] http://www.ala.org/ala/mgrps/divs/alsc/greatwebsites/greatwebsitesforkids/greatwebsites.cfm [odczyt: 2.02.2009].
4. Batorski, Dominik; Olcoń, Marta; Zając, Jan M. (2006). *Metodologia badań przez Internet: warsztat.* Ogólnopolska Konferencja Naukowa „Klikasz i znikasz? Psychologia i internet". Kraków, 17 marca 2006 [dok. elektr.] http://sna.pl/dbatorski/Warsztat-metodologia.pdf [odczyt: 10.11.2011].
5. Beck, Susan E. (2009). *The Good, The Bad & The Ugly: or, why it's a good idea to evaluate web sources* [dok. elektr.] Collection Development, New Mexico State University Library http://lib.nmsu.edu/instruction/evalcrit.html [odczyt: 2.02.2009].
6. Bednarek-Michalska, Bożena (2002). *Ocena jakości bibliotekarskich serwisów informacyjnych udostępnianych w Internecie* [dok. elektr.] „EBIB – Elektroniczny Biuletyn Bibliotekarzy" nr 31 http://ebib.oss.wroc.pl/2002/31/michalska.php [odczyt: 2.02.2009].
7. Bednarek-Michalska, Bożena (2007a). *Informacja w Internecie: podręcznik studenta.* [dok. elektr.] Kujawsko-Pomorska Biblioteka Cyfrowa http://kpbc.umk.pl/dlibra/docmetadata?id=32425&from=pubindex&dirids=79&lp=39 [odczyt: 2.02.2009].
8. Bednarek-Michalska, Bożena (2007b). *Ocena jakości informacji elektronicznej: pułapki sieci* [dok. elektr.] „EBIB – Elektroniczny Biuletyn Bibliotekarzy" nr 5(86) http://www.ebib.info/2007/86/a.php?bednarek [odczyt: 2.02.2009].
9. Bednarek-Michalska, Bożena (2007c). *Ocena jakości serwisu internetowego* [dok. elektr.] http://moodle.weinoe.us.edu.pl/file.php?file=%2F1%2FTechnologia_informacyjna_w_edukacji_wczesnoszkolnej%2FKryteria_Oceny_Stron_internetowych-1.doc [odczyt: 2.02.2009].
10. Bednarek-Michalska, Bożena (2007d). *Oceń jakość informacji elektronicznej!* [dok. elektr.] http://www.biblos.pk.edu.pl/files/File/BazTOL-Ocen%20jakosc%20informacji.ppt oraz http://www.bu.uni.torun.pl/BE/ocenjakosc.html [odczyt: 2.02.2009].
11. Bishop, Ann Peterson; Neumann, Laura J.; Star, Susan Leigh; Merkel, Cecelia; Ignacio, Emily; Sandusky, Robert J. (2000). *Digital libraries: Situating use in changing information infrastructure.* „Journal of the American Society for Information Science" 51, s. 394-413.
12. Candela, Leonardo; Nardi, A. red. (2011). *The Digital Library Reference Model* [dok. elektr.] http://www.dlorg.eu/index.php/outcomes/reference-model [odczyt: 2.06.2011].
13. Caywood, Carolyn (1995). *Library Selection Criteria for WWW Resources* [dok. elektr.] http://www.keele.ac.uk/depts/aa/landt/lt/Internet/criteria.htm [odczyt: 2.02.2009].
14. Chowdhury, Gobinda (2010). *From digital libraries to digital preservation research: the importance of users and context.* „Journal of Documentation" vol. 66 nr 2, s. 207-223.
15. Ciołek, Matthew T. (1994). *Information Quality – catalogue of potent truisms* [dok. elektr.]3http://www.ciolek.com/WWWVLPages/QltyPages/QltyTruisms.html [odczyt: 2.02.2009].

16. Ciołek, Matthew T. (1996a). *Information Quality – some definitions* [dok. elektr.] http://www.ciolek.com/WWWVLPages/QltyPages/QltyDefinitions.html [odczyt: 2.02.2009].
17. Ciołek, Matthew T. (1996b). *The six questions for the electronic grail: current approaches to information quality in WWW resources* [dok. elektr.] http://www.ciolek.com/PAPERS/six-quests1996.html [odczyt: 2.02.2009].
18. Ciołek, Matthew T. (1997). *Information Quality WWW Virtual Library: the Internet guide to construction of quality online resources.* [dok. elektr.] http://www.ciolek.com/WWWVL-InfoQuality.html [odczyt: 2.02.2009].
19. Ciolek, Matthiew T..; Goltz, Irene (1994; popr. 1997). *WWW Virtual Library: Information Quality.* [dok. elektr.] http://www.ciolek.com/WWWVL–InfoQuality.html ; http://www.ciolek.com/WWWVLPages/QltyPages/QltyTruisms.html [odczyt: 10.12.2011].
20. Cornell University Library (2004). *Critically analyzing information sources.* [dok. elektr.] http://www.library.cornell.edu/olinuris/ref/research/skill26.htm [odczyt: 2.02.2009].
21. Digital Library Federation (1998). *A working definition of digital library* [dok. elektr.] http://www.diglib.org/about/dldefinition.htm [odczyt: 10.12.2011].
22. Fourth DELOS Workshop (2002). *Evaluation of Digital Libraries: testbeds, measurements, and metrics.* Hungarian Academy of Sciences Computer and Automation Research Institute (MTA SZTAKI). Budapest, Hungary 6–7 June 2002 [dok. elektr.] http://www.ercim.eu/publication/ws–proceedings/DelNoe04.pdf [odczyt: 10.12.2011].
23. Fuhr, Norbert [i in.] (2007). *Evaluation of digital libraries* [dok. elektr.] „The International Journal on Digital Libraries" http://www.scribd.com/doc/185523/Evaluation-of-Digital-Libraries-Fulltext [odczyt: 3.03.2011].
24. Fuhr, Norbert; Hansen, Preben; Mabe, Michael; Micsik, Andras; Sølvbergh, Ingeborg (2001). *Digital Libraries: a generic classification and evaluation scheme* [dok. elektr.] http://www.is.informatik.uni-duisburg.de/bib/pdf/ir/Fuhr_etal:01.pdf [odczyt: 10.12.2011].
25. Gonçalves, Marcos André, Fox, Edward A.; Watson, Layne T.; Kipp, Neill A. (2004). *Streams, Structures, Spaces, Scenarios, Societies (5S): A formal model for digital libraries* [dok. elektr.] http://www.dlib.vt.edu/projects/5S-Model/5s6.pdf [odczyt: 2.10.2011].
26. Grassian, Esther (1996). *Thinking critically about world wide web resources* [dok. elektr.] http://www.wartburg.edu/library/infolit/Handouts/ThinkingCritically.doc oraz http://www.uwm.edu/~maryb/evalnet.htm [odczyt: 2.02.2009].
27. Griffiths, Jillian R. (2002). *A quality management approach to the evaluation of digital library services.* [W:] Fourth DELOS Workshop. Evaluation of Digital Libraries: Testbeds, Measurements, and Metrics. *Hungarian Academy of Sciences Computer and Automation Research Institute (MTA SZTAKI). Budapest, Hungary 6-7 June 2002* [dok. elektr.] http://www.docstoc.com/docs/95288604/A-Quality-Management-approach-to-the-evaluation-of-digital-library-_1_ [odczyt: 10.12.2011].
28. *How to evaluate a web page* (modyf. 2011) [dok. elektr.] http://lib.colostate.edu/howto/evalweb.html [odczyt: 10.12.2011].
29. *ICT for library and information professionals: a training package for developing countries* (2001) [dok. elektr.] Unesco. Module 4: *Database design, and information storage and retrieval.* http://dlist.sir.arizona.edu/1537/01/127825e.pdf [odczyt: 2.02.2009]; Module 5: *The Internet as the Information resource.* http://dlist.sir.arizona.edu/1538/01/127901e.pdf [odczyt: 2.02.2009] oraz http://greenstonesupport.iimk.ac.in/greenstone2010-nepal/pdf/UNESCO-Publications/Introduction%20to%20information%20and%20communication%20technologies.pdf [odczyt: 12.12.2011].
30. Janiak, Małgorzata (2008). *Bazy danych – kryteria estetyczne.* [W:] Nowakowski, Antoni red. (2008). *Infobazy'2008 „Systemy – Aplikacje – Usługi"*, Materiały V Konferencji Nau-

kowej, Sopot 15-17 października 2008 r. Gdańsk: Wydawnictwo Centrum Informatyczne TASK, Politechnika Gdańska, s. 174-179.

31. Janiak, Małgorzata (2009a). *Kryteria estetyczne w ocenie jakości bibliotek cyfrowych.* Referat wygłoszony na konferencji: 04-05.04.2011 – Warszawa „Nauka o informacji (informacja naukowa) w okresie zmian". Organizator: Instytut Informacji Naukowej i Studiów Bibliologicznych Uniwersytetu Warszawskiego.
32. Janiak, Małgorzata (2009b). *Kryteria estetyczne w ocenie jakości systemów informacyjnych.* [W:] Kamińska, Joanna; Żołędowska-Król, Beata red. (2009). *Jakość usług bibliotecznych w społeczeństwie informacyjnym: praca zbiorowa.* Warszawa: Wydaw. SBP, s. 136-144.
33. Janiak, Małgorzata; Krakowska, Monika (2010a, maszynopis Biblioteka Narodowa). *Kryteria oceny jakości bibliotek cyfrowych wraz z propozycją modelowej konstrukcji kryteriów oceny jakości dla środowiska polskich bibliotek cyfrowych.*
34. Janiak, Małgorzata; Krakowska, Monika (2010b). *Ocena bibliotek cyfrowych – kryteria jakości.* [W:] Mazurek, Cezary, Stroiński, Maciej; Węglarz, Jan red. (2010). *Polskie biblioteki cyfrowe 2010: materiały z konferencji zorganizowanej w dniach 20-21 października 2010 roku przez: Bibliotekę Kórnicką PAN, Poznańską Fundację Bibliotek Narodowych, Poznańskie centrum Superkomputerowo-Sieciowe.* Poznań: Ośrodek Wydawnictw Naukowych, s. 101-108.
35. Johns Hopkins University, The Sheridian Libraries (2010). *Evaluating Information Found on the Internet.* [dok. elektr.] http://guides.library.jhu.edu/evaluatinginformation [odczyt: 2.02.2011].
36. Kluszczyński, Ryszard W. (2001). *Społeczeństwo informacyjne. Cyberkultura. Sztuka multimediów.* Warszawa: Rabid, 219 s.
37. Mabe, Michael (2002). *DL Classification & Evaluation: A Publisher's View of the Work of the DELOS Evaluation Forum.* [dok. elektr.] http://www.sztaki.hu/conferences/deval/presentations/mabe.ppt [odczyt: 10.12.2011].
38. Marchionini, Gary (2000). *Evaluation Digital libraries: a longitudinal and multifacet view* [dok. elektr.] http://ils.unc.edu/~march/perseus/lib-trends-final.pdf [odczyt: 10.12.2011].
39. Materska, Katarzyna (2007). *Informacja w organizacjach społeczeństwa wiedzy.* Warszawa: Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich, 405, [1] s.
40. Murthy, Uma, Gorton, Douglas, Torres, Ricardo, Gonçalves, Marcos André; Fox, Edward A., Delcambre, Lois (2007). *Extending the 5S Digital Library (DL) Framework: From a Minimal DL towards a DL Reference Model: paper presented at the First Digital Library Foundations Workshop – held in conjunction with the 2007 Joint Conference on Digital Libraries (Vancouver, Canada, June 18-24, 2007).* [dok. elektr.] http://www.dlib.vt.edu/projects/5S-Model/extending5S.pdf
41. *Quality Selection Criteria for Subject Gateways* [dok. elektr.] Produced by the European Union-funded DESIRE Project http://sosig.ac.uk/desire/qindex.html [odczyt: 2.02.2009].
42. Reeves, Thomas; Apedoe, Xornam; Hee Woo, Young. (2003). *Evaluating digital libraries: a user–friendly guide.* University Corporation for Atmospheric Research; National Science Digital Library. [dok. elektr.] http://www.dpc.ucar.edu/projects/evalbook/EvaluatingDigitalLibraries.pdf [odczyt: 10.12.2011].
43. Reeves, Thomas C.; Apedoe, Xornam; Hee Woo, Young. (2005). *Evaluating digital libraries: a user–friendly guide. University Corporation for Atmospheric Research; National Science Digital Library.* [dok. elektr.] http://www.dpc.ucar.edu/projects/evalbook/EvaluatingDigitalLibraries.pdf [odczyt: 12.11.2011].

44. Sandusky, Robert J. (2002). *Digital library attributes: framing research and results.* [W:] *Fourth DELOS Workshop. Evaluation of Digital Libraries: Testbeds, Measurements, and Metrics. Hungarian Academy of Sciences Computer and Automation Research Institute (MTA SZTAKI).* Budapest, Hungary 6–7 June 2002 [dok. elektr.] http://www.ercim.eu/publication/ws–proceedings/DelNoe04.pdf [odczyt: 5.03.2011].
45. Sandusky, Robert J. (2004). *Digital library attributes: framing usability research.* [W:] *Fourth DELOS Workshop. Evaluation of Digital Libraries:* Testbeds, Measurements, and Metrics. *Hungarian Academy of Sciences Computer and Automation Research Institute (MTA SZTAKI). Budapest, Hungary 6–7 June 2002* [dok. elektr.] http://www.uclic.ucl.ac.uk/annb/docs/Sandusky35.pdf oraz *http://www.ercim.eu/publication/ws–proceedings/DelNoe04.pdf* [odczyt: 5.03.2011].
46. Sapa, Remigiusz (2005). *Benchmarking w doskonaleniu serwisów WWW bibliotek akademickich.* Kraków: Wydaw. Uniw. Jagiel., 254 s.
47. Saracevic, Tefko (2000). *Digital library evaluation: toward an evolution of concepts – 1 – evaluation criteria for design and management of digital libraries.* Library Trends, Fall, 2000 [dok. elektr.] http://findarticles.com/p/articles/mi_m1387/is_2_49/ai_72274400/pg_18/?tag=content;col1 [odczyt: 10.12.2011].
48. Saracevic, Tefko (2000). *Digital library evaluation: toward an evolution of concepts – 1 – evaluation criteria for design and management of digital libraries* [dok. elektr.] „Library Trends" http://findarticles.com/p/articles/mi_m1387/is_2_49/ai_72274400/pg_18/?tag=content;col1 [odczyt: 3.03.2011].
49. Saracevic, Tefko (2004). *Evaluation of digital libraries: an overview, paper presented at DELOS Workshop on the Evaluation of Digital Libraries* [dok. elektr.] *DELOS* http://dlib.ionio.gr/wp7/WS2004_Saracevic.pdf [odczyt: 2.02.2009].
50. Saracevic, Tefko (2004). *How were digital libraries evaluated.* [dok. elektr.] http://comminfo.rutgers.edu/~tefko/DL_evaluation_LIDA.pdf [odczyt: 25.11.2011].
51. Saracevic, Tefko.; Covi, Lisa (2000). *Challenges for digital library evaluation.* [W:] Kraft, D.H. red. (2000). *Knowledge innovations: celebrating our heritage, designing our future: proceedings of the 63rd Annual Meeting, November 11–16, 2000, Chicago, IL. Washington*: American Society for Information Science, s. 341-350. [Wersja elektroniczna: http://www.is.informatik.uni-duisburg.de/courses/dl_ss04/folien/saracevic00.pdf].
52. Schrock, Kathy (2008). [dok. elektr.] www.kathyschrock.net [odczyt: 2.02.2009].
53. Smith, Alistair (1996). *Criteria for Evaluation of Internet Information Resources.* [dok. elektr.] http://www.vuw.ac.nz/~agsmith/evaln/index.htm [odczyt: 2.02.2009].
54. Smith, Alistair (1996; modyf. 2011). *Evaluation of Information Resources.* [dok. elektr.] http://www.vuw.ac.nz/~agsmith/evaln/evaln.htm [odczyt: 2.02.2009].
55. Tate, Marsha; Alexander, Jane (1996). *Teaching critical evaluation skills for world wide web resources.* „Computers in Libraries" nr 16(10), s. 49-54.
56. Tillman, Hope N. (1996). *Evaluating quality on the net.* [dok. elektr.] http://www.hope-tillman.com/findqual.html [odczyt: 2.02.2009].
57. UC Berkeley – Teaching Library Internet Workshops (2008). *Evaluating web pages: techniques to apply & questions to ask* [dok. elektr.] http://www.lib.berkeley.edu/TeachingLib/Guides/Internet/Evaluate.html oraz http://www.lib.berkeley.edu/TeachingLib/Guides/Internet/evaluation_checklist_2008_spring.pdf [odczyt: 2.02.2009].
58. Wildemuth, Barbary (2007). *Digital library evaluation, user studies. Draft, 2* [dok. elektr.] http://curric.dlib.vt.edu/wiki/index.php?title=Module_9–c:_Digital_library_evaluation,_user_studies [odczyt: 2.11.2011].

59. Xie, Hong (2006). *Evaluation of digital libraries: criteria and problems from users' perspective*. „Library and Information Science Research" 28, s. 433-452.
60. Xie, Hong (2008). *Users' evaluation of digital libraries: their uses, their criteria, and their assessment.* „Information Processing & Management" 44(3), s. 1346-1373.
61. Zhang, Ying (2010). *Developing a holistic model for digital library evaluation* [dok. elektr.] „Wiley Online" *Library* http://onlinelibrary.wiley.com/doi/10.1002/asi.21220/pdf [odczyt: 2.02.2011].
62. Zhang, Ying; Saracevic, Tefko (2007a). *Criteria in evaluation of use & usability in digital libraries* [dok. elektr.] http://www.ffos.hr/lida/datoteke/LIDA2007-Zhang_Saracevic.ppt#256,1,criteria in evaluation of use & usability in digital libraries [odczyt: 10.12.2011].
63. Zhang, Ying (2007b). *Developing A holistic model for digital library evaluation.* A dissertation submitted to the Graduate School-New Brunswick Rutgers, The State University of New Jersey in partial fulfillment of the requirements for the degree of Doctor of Philosophy Graduate Program in Communication, Information, and Library Studies written under the direction of Dr. Tefko Saracevic. New Brunswick, New Jersey, 248 s.

Biblioteki cyfrowe a regulacje prawne

Sybilla Stanisławska-Kloc

W tekście zamieszczono podstawowe informacje o przepisach prawnych istotnych dla tworzenia i rozbudowywania bibliotek cyfrowych. Omówiono przede wszystkim regulacje autorskoprawne oraz te dotyczące ochrony baz danych, danych osobowych, partnerstwa publiczno-prywatnego, prasy. Przedstawiono problematykę związaną z domeną publiczną, dziełami osieroconymi, licencjami Creative Commons.

Wstęp

Jedną z dziedzin prawa, która ma obecnie bardzo istotne znaczenie dla funkcjonowania bibliotek cyfrowych, jak i samej działalności podejmowanej przez bibliotekarzy jest prawo autorskie*.

Wśród instytucji, które podejmują inicjatywy służące zachowaniu dorobku – dziedzictwa kulturalnego takich, jak muzea, archiwa to właśnie biblioteki są w największym stopniu uzależnione od regulacji autorskoprawnych. Wynika to z kilku powodów. Co prawda, wszystkie te trzy kategorie instytucji w coraz większym stopniu udostępniają swoje zbiory w wersji cyfrowej (na miejscu lub online), stąd tworzone są cyfrowe: biblioteki, archiwa i muzea. Jednak jak wydaje się to biblioteki posiadają najliczniejsze zbiory i są zainteresowane jak najszerszym ich udostępnianiem, szczególnie właśnie w wersji cyfrowej. Ale to co najistotniejsze to okoliczność, iż wśród zbiorów bibliotecznych znajdują się w dużej mierze utwory, i to takie w stosunku do których nie wygasły jeszcze autorskie prawa majątkowe. Zatem określenie zasad korzystania z nich (udostępniania) wymaga znajomości podstawowych założeń prawa autorskiego. Natomiast archiwa (poza niektórymi archiwami współczesnymi) oraz muzea (poza muzeami sztuki współczesnej) głównie dysponują i udostępniają utwory, które bardzo często nie podlegają już ochronie w zakresie autorskich praw majątkowych. A co za tym idzie w mniejszym stopniu aniżeli biblioteki instytucje te są uzależnione od regulacji autorskoprawnych.

Ponadto wśród bibliotekarzy zajmujących się informacją naukową, brokerów informacji, przygotowujących opracowania cudzych utworów, zapewniających do-

* Opracowanie zawiera poglądy wyrażone przez autorkę we wcześniejszych publikacjach dotyczących omawianej problematyki.

stęp do informacji zawartych w utworach, znajomość podstawowych zasad prawa autorskiego wydaje się także nieodzowna. Sami bibliotekarze uświadomili sobie tę konieczność (jak i zarazem potrzebę) posiadania wiedzy, nabywania praktyki z zakresu stosowania przepisów prawa autorskiego, ochrony utworów stąd wykształciła się pośród nich grupa specjalistów określanych mianem „copyright librarian". Znajomość założeń prawa autorskiego pozwala bibliotekarzom na zabieranie głosu na forum międzynarodowym, aktywne uczestniczenie w inicjatywach, których celem jest wpływanie na kształt regulacji autorskoprawnych, które dotyczą sfery ich działalności, jak i służenie pomocą oraz edukowanie użytkowników (czytelników) w zakresie zgodnego z prawem korzystania z utworów.

Jak już wskazywałam biblioteki cyfrowe m.in. z uwagi na wielkość i charakter swoich zbiorów (jak żadne inne instytucje o charakterze naukowo-kulturalnym) stając przed coraz to nowymi wyzwaniami technologicznymi, chcąc korzystać z nowych rozwiązań technicznych (w zakresie udostępniania zbiorów) muszą uwzględniać zasady autorskoprawne. Znajomość prawa autorskiego pozwala – realizować zasadę „nowego renesansu" (szerokiego udostępniania utworów społeczeństwu) – o który zaapelowali twórcy raportu Sages. Przywołując wypowiedź Jeana Monneta: *jeżeli trzeba byłoby odbudowywać Europę, zacząłbym od kultury, nie od gospodarki,* wskazali, iż *dziedzictwo kulturowe Starego Kontynentu wzbogacało edukację oraz formowało postawy i ducha pokoleń żyjących przed nami; dlatego czujemy się zobowiązani przekazać to bogate (w istocie jedno z najbogatszych na świecie) dziedzictwo przyszłym pokoleniom i dopilnować, aby było ono chronione, wzbogacane i rozpowszechniane.* Ochrona, wzbogacanie oraz rozpowszechnianie tego dziedzictwa musi odbywać się z uwzględnieniem reguł prawa autorskiego. Twórcy raportu Sages zwrócili uwagę na to, że *nowe technologie informatyczne stworzyły niewiarygodne szanse, dzięki którym to wspólne dziedzictwo może stać się bardziej dostępne dla wszystkich. Kultura podąża ścieżką digitalizacji, a instytuty pamięci dostosowują sposób, w jaki komunikują się ze społeczeństwem. Digitalizacja pozwala tchnąć nowe życie w materiały z przeszłości, dzięki czemu stają się one cennymi zasobami dla indywidualnych użytkowników i ważną częścią składową gospodarki cyfrowej* [Report of the 'Comite des Sages' 2011]. Jak wskażę w dalszej części opracowania to właśnie proces digitalizacji materiałów bibliotecznych oraz udostępniania ich cyfrowych wersji rodzi wiele zagadnień autorskoprawnych. Poza regulacją autorskoprawną zostaną wskazane także inne, wybrane, regulacje prawne, które mają wpływ na funkcjonowanie bibliotek cyfrowych.

Należy podkreślić, iż ustawa o bibliotekach z 27 czerwca 1997 r. (opubl. w Dz. U. nr 85, poz. 539, dalej „ust. o bibl.") nie zawiera żadnych szczegółowych regulacji autorskoprawnych dotyczących korzystania z utworów, które znajdują się w zbiorach bibliotecznych. Zatem tylko i wyłącznie ustawa o prawie autorskim i prawach pokrewnych z 24 lutego 1994 r. (t.j. opubl. w Dz. U. z 2006 r., nr 90, poz. 630 ze zm., dalej „pr. aut.") stanowi, jedyny, najważniejszy akt prawny regulujący kwestię ochrony i korzystania z utworów, które znajdują się w zbiorach bibliotecznych.

Tytułem wprowadzenia należy wskazać, iż prawo autorskie należy do dziedziny prawa określanej mianem prawa własności intelektualnej, które chroni rezultaty dzia-

łalności intelektualnej człowieka. Inaczej prawa te są określane terminem praw na dobrach niematerialnych – chronią bowiem (niematerialne), intelektualne rezultaty działalności (np. utwory, wynalazki), bez względu na to, w jakim materialnym nośniku będą one ucieleśnione (np. płyta CD lub DVD, książka, towar (wytwór) wyprodukowany według wynalazku (metody) patentowanego).

Pierwszym, najważniejszym artykułem, który poruszał problematykę prawa autorskiego w kontekście funkcjonowania bibliotek cyfrowych był artykuł Janusza Barty i Ryszarda Markiewicza *Wirtualne biblioteki a prawo autorskie*, opublikowany w księdze pamiątkowej przygotowanej dla Wandy Pindlowej z 2004 r. [Kocójowa 2004][1]. W tym to roku nastąpiły istotne zmiany, w zakresie regulacji autorskoprawnej, które odnosiły się do bibliotek. Wtedy też Panowie Profesorowie postawili tezę, która do dzisiaj jest aktualna i która wymaga na bieżąco „sprawdzania" w zakresie realizacji: *Racjonalnym generalnym postulatem, także pod adresem ustawodawcy, byłoby zapewnienie swobody dostępu czytelników (użytkowników końcowych) do zgromadzonych dzieł na poziomie nie niższym niż gwarantowany jest obecnie bibliotekom tradycyjnym przez przepisy o dozwolonym użytku.*

Na koniec tych wstępnych rozważań zasygnalizuję jedno zagadnienie, które choć po części naświetli jak skomplikowane/szczególne mogą być relacje oraz jak bywają „zgrupowanie" racje pomiędzy podmiotami uczestniczącymi w procesie udostępniania utworów. W USA biblioteki, które współpracują z Google (udostępniły do digitalizacji swoje zbiory) zostały pozwane z tytułu naruszenia prawa autorskiego przez wydawców i autorów [Timmer 2011]. Natomiast w Europie, jak wydaje się większa część bibliotek nie przystąpiła do współpracy z Google, poniekąd zaważyły na tym nie tylko kwestie ideologiczne, ale i odmienne od amerykańskich założenia i szczegółowe rozwiązania w prawie autorskim.

Regulacje prawne – zarys ogólny

1. Prawo polskie – ustawa o prawie autorskim i prawach pokrewnych z 1994 r.

Prawo autorskie zostało tak skonstruowane, aby chronić przede wszystkim interesy osoby, która stworzyła utwór. Twórcy (autorowi) przysługują dwa rodzaje praw: autorskie prawa osobiste (w tym prawo do autorstwa, oznaczania utworu swoim nazwiskiem, nienaruszalności treści i formy utworu), jak i autorskie prawa majątkowe (w tym prawo do zwielokrotniania, rozpowszechniania). Prawa osobiste są nieograni-

[1] Omówienie wybranych zagadnień dotyczących stosowania prawa autorskiego w aspekcie bibliotek cyfrowych zawarłam w opinii *Prawo autorskie a biblioteka cyfrowa*, maj 2005, udostępnianej na stronie: http://www.kbnrkp.edu.pl/publikacje/Michalska2.rtf [odczyt: 2.01.2012]. Por. też. *Podręcznik prawa autorskiego dla bibliotekarzy*, E. FL. Handbook on Copyright – Polish, tł. i pol. oprac. Barbara Szczepańska [dok. eketr.] http://www.eifl.net/eifl-handbook-copyright-polish [odczyt: 1.05.2012].

czone w czasie (trwają wiecznie), natomiast prawa majątkowe trwają przez okres życia twórcy i 70 lat po jego śmierci.

Po wygaśnięciu autorskich praw majątkowych, każdy, zarówno w celach naukowych, edukacyjnych, rozrywkowych, jak i komercyjnych może korzystać z utworu (np. zwielokrotniać, umieszczać na stronie www), bez konieczności uzyskiwania zgody od autora (czy też jego spadkobierców), ale pod warunkiem przestrzegania autorskich praw osobistych [Papadopoulou 2011; Iglezakis 2011] (głównie chodzi tu o właściwe oznaczenie autorstwa, nieingerowanie w treść i formę utworu).

Natomiast korzystanie z utworów w stosunku do których trwają autorskie prawa majątkowe może odbywać się na podstawie umowy z autorem lub na podstawie przepisów o tzw. dozwolonym użytku. Jedna z form dozwolonego użytku, została przewidziana specjalnie dla bibliotek (por. art. 28 pr.aut.).

Źródła prawa

Zgodnie z art. 87 Konstytucji RP źródłami powszechnie obowiązującego prawa Rzeczypospolitej Polskiej są: Konstytucja, ustawy, ratyfikowane umowy międzynarodowe oraz rozporządzenia.

Prawo polskie ma charakter prawa stanowionego, obowiązują tylko te przepisy prawne, które zostały uchwalone przez sejm i senat, a następnie opublikowane w Dzienniku Ustaw[2]. Orzeczenia sądowe nie tworzą prawa; tj. nie są powszechnie obowiązujące. Natomiast orzeczenia służą do interpretacji prawa, zawarte w nich rozważania pomagają wyjaśnić, doprecyzować zakres stosowania przepisów, często są powoływane w zarówno w pismach procesowych, jak i komentarzach, opracowaniach tak naukowych, jak i popularnonaukowych. Sąd Najwyższy oraz Naczelny Sąd Administracyjny publikują orzeczenia na swoich stronach internetowych[3].

Pierwsza polska ustawa o prawie autorskich pochodziła z 1926 r., następnie w 1952 r. uchwalono kolejną ustawę. Obecnie obowiązująca od 1994 r. ustawa reguluje zasady ochrony i korzystania z utworów oraz przedmiotów praw pokrewnych. Najważniejsze rozporządzenia wykonawcze do ustawy to: a) rozporządzenie Ministra Kultury z dnia 24 lutego 2003 r. w sprawie wysokości procentu wpłat na Fundusz Promocji Twórczości (Dz. U. z 2003 r., nr 41, poz. 354); b) rozporządzenie Ministra Kultury z dnia 17 stycznia 2003 r. w sprawie podmiotów uprawnionych do występowania z wnioskiem o przyznanie środków z Funduszu Promocji Twórczości oraz wymogów formalnych, jakim powinien odpowiadać ten wniosek (Dz. U. z 2003, nr 13, poz. 134); c) rozporządzenie Ministra Kultury z dnia 2 czerwca 2003 r. w sprawie określenia kategorii urządzeń i nośników służących do utrwalania utworów oraz opłat od tych urządzeń

[2] Oficjalne publikatory oraz zasady promulgacji aktów prawnych określa ustawa z dnia 20 lipca 2000 r. o ogłaszaniu aktów prawnych i niektórych innych aktów (tj. Dz. U. z 2005 r., nr 190, poz. 1606). Obecnie wszystkie Dzienniki Ustaw są nieodpłatnie dostępne w wersji elektronicznej na stronie: http://dziennikustaw.gov.pl/DziennikiUrzedowe.aspx [odczyt: 3.01.2012].

[3] Baza orzeczeń SN dostępna jest pod adresem: http://www.sn.pl/orzecznictwo/index.html; natomiast baza orzeczeń NSA jest dostępna pod adresem: http://orzeczenia.nsa.gov.pl/cbo/query [odczyt: 14.01.2012].

i nośników z tytułu ich sprzedaży przez producentów i importerów (Dz. U. z 2003 r., nr 105, poz. 991, ze zm. Dz. U. z 2008 r., nr 235, poz. 1599; Dz. U. z 2011 r., nr 105, poz. 616); d) rozporządzenie Ministra Kultury z dnia 27 czerwca 2003 r., w sprawie opłat uiszczanych przez posiadaczy urządzeń reprograficznych (Dz. U. z 2003 r., nr 132, poz. 1232); e) rozporządzenie Ministra Kultury z dnia 30 kwietnia 2004 r. w sprawie rejestru informacji o produkcji nośników optycznych oraz rodzajów kodów identyfikacyjnych (Dz. U. z 2004 r., nr 124, poz. 1301).

Zakres przedmiotowy ustawy

Zakres przedmiotowy ustawy znajduje swoje „odbicie" w jej tytule („ustawa o prawie autorskim i prawach pokrewnych"). Ustawa obejmuje utwory chronione prawem autorskim (wszelkie kategorie utworów, zarówno te przykładowo wymienione w art. 1 i 3 pr. aut., jak i takie które nie zostały wskazane np. multimedialne; istniejące lub takie które dopiero powstaną w przyszłości). Definicja utworu jest szeroka, pozwala na objęcie tym terminem wszelkich rezultatów działalności twórczej o indywidualnym charakterze. Drugą kategorię, objętą zbiorczą nazwą przedmiotów praw pokrewnych stanowią: artystyczne wykonania, fonogramy, wideogramy, nadania programów stacji radiowych i telewizyjnych, pierwsze wydania oraz wydania krytyczne i naukowe. Katalog przedmiotów praw pokrewnych jest zamknięty. Jeżeli w przyszłości miałby on zostać rozszerzony to wiązać się to będzie z nowelizacją ustawy i wprowadzeniem specjalnej regulacji w odniesieniu do nowego przedmiotu praw pokrewnych (np. nietwórczych fotografii, tak jak ma to miejsce w prawie niemieckim; obecnie nie trwają jednak żadne prace nad nowelizacją polskiej ustawy w zakresie praw pokrewnych).

Zakres podmiotowy ustawy

Reguły zawarte w pr. aut. odnoszą się do utworów twórców polskich (posiadających obywatelstwo polskie), jak i cudzoziemców (twórców obcych państw), co wynika z art. 5 pr. aut. Zgodnie z tym przepisem polskie prawo autorskie stosuje się do utworów:

- których twórca lub współtwórca jest obywatelem polskim,
- których twórca jest obywatelem państwa UE lub EFTA,
- które zostały po raz pierwszy opublikowane na terytorium Polski (albo równocześnie na tym terytorium i za granicą),
- które zostały opublikowane po raz pierwszy w języku polskim,
- które są chronione na podstawie umów międzynarodowych, w zakresie, w jakim ich ochrona wynika z tych umów[4].

Korzystanie na terytorium Polski z utworów autorów polskich, jak i zagranicznych (w zakresie o którym mowa w art. 5 pr. aut.) odbywa się na zasadach określonych w polskiej ustawie. Ustawa polska ma zastosowanie do w/w utworów wykorzystywa-

[4] Na temat regulacji międzynarodowej por. pkt. II.3

nych w bibliotekach cyfrowych prowadzonych przez podmioty, które mają siedzibę w Polsce.

Treść praw

Na treść praw autorskich składają się uprawnienia osobiste, jak i majątkowe, zaś na treść praw pokrewnych zasadniczo tylko uprawnienia majątkowe (wyjątek stanowią tutaj prawa osobiste artystów wykonawców). Prawa autorskie i pokrewne podlegają swoistemu ograniczeniu, w związku z funkcjonowaniem instytucji dozwolonego użytku. Szczegółowo wybrane kwestie autorskoprawne dotyczące bibliotek omówione są w pkt III.

2. Ustawa o ochronie baz danych z 2001 r.

Szczegółowe zasady ochrony baz danych zawiera ustawa z dnia 27 lipca 2001 r., o ochronie baz danych (Dz. U. z 2003 , nr 128, poz. 1402, ze zm., dalej „ust. b. d."). Ustawa ta wprowadza odrębną, niejako uzupełniającą (istniejąca równolegle obok ochrony autorskoprawnej) ochronę dla tych baz danych, które wymagały istotnego nakładu inwestycyjnego w celu ich sporządzenia, weryfikacji, prezentacji. Ochrona ta określana jest mianem ochrony *sui generis*. System ochrony *sui generis* został wprowadzony dyrektywą 96/9 z dnia 11 marca 1996 r. o prawnej ochronie baz danych; nie istnieje on poza UE.

Zgodnie z art. 2 ust. 2 pkt 1 „ust. b. d.": baza danych oznacza zbiór danych lub jakichkolwiek innych materiałów i elementów, zgromadzonych według określonej systematyki lub metody, indywidualnie dostępnych w jakikolwiek sposób, w tym środkami elektronicznymi, wymagający istotnego nakładu inwestycyjnego w celu sporządzenia, weryfikacji lub prezentacji zawartości. [Barta; Markiewicz 2002; Stanisławska-Kloc 2002; Barta; Markiewicz 2010, s. 272].

Wydaje się uzasadnione twierdzenie, iż wiele kolekcji bibliotek cyfrowych, portali internetowych bibliotek będzie spełniać kryteria właściwe bazom danych (będą się składać z niezależnych materiałów, indywidualnie dostępnych).

Ochrona *sui generis* przysługuje producentom, w stosunku do baz danych, które wymagały poniesienia istotnego nakładu inwestycyjnego. Nakład ten może być związany z poczynionymi wydatkami na stworzenie bazy, zatrudnienie pracowników, którzy ją opracowali – zebrali potrzebne materiały, jak i własną pracą (poświęconym czasem). Pomimo, że brak jednoznacznych wytycznych co do minimalnej wysokości nakładu warunkującego przyznanie ochrony to wydaje się, że wydatki rzędu kilkudziesięciu tysięcy złotych będą uznane za istotny nakład inwestycyjny [por. Barta; Markiewicz 2010, s. 276]. W kontekście tworzenia kolekcji cyfrowych wydatki te mogłyby dotyczyć np. digitalizacji i obróbki cyfrowej materiałów pierwotnie dostępnych w wersji papierowej. Ochrona *sui generis* bazy danych ma coraz większe znaczenia dla zabezpieczenia interesów podmiotów tworzących różnego rodzaju zbiory,

zestawienia, kompilacje – w tym dla bibliotek. Ochrona trwa zasadniczo przez okres 15 lat od stworzenia bazy.

3. Rozwiązania i propozycje prawne w UE

Działania podejmowane na forum unijnym odgrywają istotną rolę przy określaniu zasad korzystania z utworów udostępnianych w bibliotekach cyfrowych, jak i wypracowywaniu kierunku unijnej polityki w tym zakresie. Znajduje to potwierdzenie we wskazanych poniżej dokumentach. Uchwalone już unijne regulacje prawne wpływają na kształt polskiego prawa; zarówno dotyczy to sfery stanowienia prawa (w związku z uchwalaniem nowych polskich aktów prawnych, lub nowelizacją istniejących już aktów), jak i interpretacji obowiązujących przepisów (tzw. pro unijna interpretacja polskiego prawa).

Prawo unijne ma charakter prawa stanowionego, zgodnie z art. 288 TFUE (dawny art. 249 TWE): *W celu wykonania kompetencji Unii instytucje przyjmują rozporządzenia, dyrektywy, decyzje, zalecenia i opinie. Rozporządzenie ma zasięg ogólny. Wiąże w całości i jest bezpośrednio stosowane we wszystkich Państwach Członkowskich. Dyrektywa wiąże każde Państwo Członkowskie, do którego jest kierowana, w odniesieniu do rezultatu, który ma być osiągnięty, pozostawia jednak organom krajowym swobodę wyboru formy i środków. Decyzja wiąże w całości. Decyzja, która wskazuje adresatów, wiąże tylko tych adresatów. Zalecenia i opinie nie mają mocy wiążącej*[5]. Podstawowym publikatorem unijnych aktów prawnych jest Dziennik Urzędowy Unii Europejskiej[6].

W zakresie prawa autorskiego i praw pokrewnych nie uchwalono dotąd rozporządzeń, natomiast przyjęto 8 dyrektyw[7], które stanowią narzędzie harmonizacji regulacji krajowych, a odbywa się to poprzez implementację rozwiązań zawartych w dyrektywie do ustaw krajowych. Uchwalone dotychczas dyrektywy (poza ubiegłoroczną nowelizacją dyrektywy o czasie trwania praw pokrewnych) zostały w zasadzie implementowane do polskiego systemu prawnego [Barta; Markiewicz 1995, 2010].

[5] Traktat o funkcjonowaniu Unii Europejskiej, opubl. w *Dzienniku Urzędowym Unii Europejskiej*, seria C, nr 83 z dnia 30 marca 2010 r.

[6] Dostęp do unijnych aktów prawnych odbywa się za pośrednictwem strony *EUR-LEX*: http://eur-lex.europa.eu/JOIndex.do?ihmlang=pl [odczyt: 2.02.2012].

[7] Są to następujące dyrektywy: 1) dyrektywa 2009/24/WE Parlamentu Europejskiego i Rady z 23 kwietnia 2009 r. (wersja ujednolicona wcześniejszej dyrektywy 91/250/WE Rady z 14 maja 1991 r. o ochronie programów komputerowych); 2) dyrektywa 93/83/WE Rady z 27 września 1993 r., w sprawie koordynacji określonych przepisów prawa autorskiego oraz praw pokrewnych w odniesieniu do przekazu satelitarnego i rozpowszechniania kablowego; 3) dyrektywa 2006/116/WE Parlamentu Europejskiego i Rady z 12 grudnia 2006 r. w prawie czasu ochrony prawa autorskiego i niektórych praw pokrewnych (wersja ujednolicona wcześniejszej dyrektywy 93/98/WE Rady z 29 października 1993 r. w sprawie ujednolicenia czasu ochrony prawa autorskiego i niektórych praw pokrewnych), znowelizowana 27 września 2011 r.; 4) dyrektywa 96/9/WE Parlamentu Europejskiego i Rady z 11 marca 1996 r. o prawnej ochronie baz danych; 5) dyrektywa 2001/84/EC Parlamentu Europejskiego i Rady z 27 września 2001 r., w sprawie prawa autora do wynagrodzenia z tytułu odsprzedaży oryginalnego egzemplarza dzieła sztuki *(droit de suite)*; 6) dyrektywa 2004/48/WE Parlamentu Europejskiego i Rady z 29 kwietnia 2004 r., w sprawie egzekwowania praw własności intelektualnej.

Dla problematyki stosowania prawa autorskiego w kontekście działalności bibliotek szczególne znaczenie mają przede wszystkim dwie dyrektywy:

- dyrektywa 2001/29/WE Parlamentu Europejskiego i Rady z 22 maja 2001 r. o ujednoliceniu niektórych aspektów prawa autorskiego oraz praw pokrewnych w społeczeństwie informacyjnym.

Dla udostępniania przez biblioteki utworów chronionych prawem autorskim najistotniejsze znaczenie mają następujące postanowienia wskazanej powyżej dyrektywy 2001/29/WE: art. 5 ust. 2 lit. „c", art. 5 ust. 3 lit. „n" i art. 5 ust. 5 (jak i pkt 40 preambuły), które statuują dozwolony użytek utworów – podstawę dla korzystania (w pewnym zakresie) z utworów bez konieczności uzyskiwania zgody podmiotu uprawnionego z tytułu praw autorskich. Polski ustawodawca wypełniając zobowiązania traktatowe – w zakresie implementacji tej dyrektywy – dokonał nowelizacji art. 28 pr. aut., szczegółowo kwestia ta będzie omówiona w pkt. IV.9.

Omawiając problematykę bibliotek cyfrowych należy wskazać na wybrane dokumenty unijne, które mają znaczenie dla kształtowania się autorskoprawnych ram funkcjonowania tych instytucji.

Pierwszym z nich, w porządku chronologicznym, są zalecenia Komisji z dnia 24 sierpnia 2006 r. w sprawie digitalizacji i udostępniania w Internecie dorobku kulturowego oraz w sprawie ochrony zasobów cyfrowych (2006/585/WE)[8], w których dostrzeżono potrzebę tworzenia różnego rodzaju kolekcji i ich udostępniania. Zasygnalizowane w tym dokumencie zagadnienia dotyczyły:

- digitalizacji i udostępniania w Internecie materiałów tworzących dorobek kulturalny; w celu uniknięcia powielania działań powinny zostać gromadzone informacje o digitalizowanych obiektach, powinny być wprowadzane wspólne standardy digitalizacji w celu zapewnienia interoperacyjności zasobów; postulowano także zachęcanie instytucji kulturalnych oraz wydawców i podmioty uprawnione do udostępniania ich zasobów w wyszukiwarce europejskiej biblioteki cyfrowej; zalecono stworzenie procedur ułatwiających korzystanie z utworów osieroconych oraz takich, których nakład został wyczerpany (których egzemplarze nie są dostępne w handlu);
- ochrony zasobów cyfrowych; w tym zakresie powinny zostać opracowane strategie na rzecz długoterminowej ochrony i dostępu przy uwzględnieniu praw autorskich, ustawodawstwa krajowe powinny zapewnić instytucjom publicznym, w celu ochrony zasobów cyfrowych wielokrotne powielanie oraz zmianę formatu i/lub nośnika na którym utwory są utrwalone; określenie zasad przekazywania egzemplarza obowiązkowego materiałów powstałych w formie cyfrowej; tzw. Web-harvestingu (gromadzenia materiałów dostępnych w Internecie, na potrzeby ochrony zasobów cyfrowych).

Wykonanie zaleceń zostało zaprezentowane w Komunikacie Komisji do Rady, Parlamentu Europejskiego, Europejskiego Komitetu Ekonomiczno-Społecznego

[8] Polska wersja językowa dostępna na stronie http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2008:0513:FIN:PL:HTML [odczyt: 2.02.2012].

i Komitetu Regionów – Dostęp do dziedzictwa kulturowego Europy poprzez kliknięcie myszką. Postępy w zakresie digitalizacji i udostępnienia w Internecie dorobku kulturowego oraz ochrony zasobów cyfrowych w UE[9].

Drugim dokumentem, który warto wzmiankować jest tzw. *Zielona Księga* (Green Paper) – prawo autorskie w gospodarce opartej na wiedzy[10] z 2008 r. Celem rozpowszechnienia tego typu dokumentu było zachęcenie do publicznej debaty (rozpoczęcie konsultacji społecznych) nad sposobami rozpowszechniania w środowisku online wiedzy związanej z badaniami, nauką i edukacją. W księdze m.in. zasygnalizowano zagadnienia związane z: a) wytwarzaniem cyfrowych kopii materiałów przechowywanych w kolekcjach różnych instytucji oraz b) udostępnianiem kopii utworów użytkownikom drogą elektroniczną. Podkreślono, że digitalizacja pozwala na zachowanie (ochronę) tych utworów dla przyszłych pokoleń (co wynika z faktu sporządzanie dodatkowej, z reguły dobrej jakości i nowej „żywotności" kopii), a równocześnie umożliwia potencjalnie nieograniczone udostępnianie niejako „na bieżąco" (online) zdigitalizowanych (cyfrowych) wersji obecnym użytkownikom. W *Zielonej Księdze* – wymieniono swobodę przepływu wiedzy i innowacji jako tzw. piątą „swobodę" unijną (obok czterech traktatowych swobód przepływu: ludzi, towarów, usług, kapitału).

W *Zielonej Księdze* umieszczono następujące pytania dotyczące problematyki m.in. udostępniania utworów przez biblioteki i archiwa, które miałyby być przedmiotem konsultacji:

- czy dozwolony użytek dotyczący bibliotek i archiwów powinien pozostać niezmieniony, ponieważ wydawcy sami opracują zasady dostępu online do swoich katalogów;
- czy ogólnodostępne biblioteki, instytucje edukacyjne, muzea i archiwa powinny zawierać porozumienia licencyjne z wydawcami, aby zwiększyć dostępność utworów? Czy znane są przykłady funkcjonujących w praktyce porozumień licencyjnych obejmujących dostęp online do kolekcji bibliotecznych;
- czy należy sprecyzować zakres dozwolonego użytku dotyczącego ogólnodostępnych bibliotek, instytucji edukacyjnych, muzeów i archiwów w odniesieniu do: (a) zmiany formy zapisu utworu; (b) liczby kopii, które można wykonać w ramach dozwolonego użytku; (c) skanowania całych kolekcji znajdujących się w posiadaniu bibliotek;
- czy należy sprecyzować w przepisach prawa, czy czynność skanowania utworów przechowywanych w bibliotekach, aby ich treść mogła być przeszukiwana w Internecie, wychodzi poza zakres obecnie obowiązujących wyjątków od prawa autorskiego.

W dokumencie tym poddano także pod dyskusję kwestię utworów osieroconych.

Z treści tych pytań można wywnioskować, jakie kwestie wydają się być problematyczne w odniesieniu do zakresu digitalizacji i udostępniania online zbiorów przez

[9] [SEC(2008)2372], COM/2008/0513 końcowy, dostępny na stronie: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52008DC0513:PL:HTML [odczyt: 2.02.2012].

[10] Dokument z dnia 16 lipca 2008 r., dostępny na stronie: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2008:0466:FIN:EN:PDF [odczyt: 2.02.2012].

biblioteki. Pod koniec 2009 r. ogłoszono Komunikat Komisji[11], który zawierał wyniki konsultacji. Jak można było przypuszczać – potwierdziły one podział występujący w środowisku. Bibliotekarze oraz przedstawiciele innych instytucji naukowych opowiedzieli się za taką treścią dozwolonego użytku, aby w interesie społecznym zapewnić szeroki dostęp do wiedzy; szerszy aniżeli obecnie przewidziany w dyrektywie 2001/29/WE. Natomiast wydawcy, organizacje zbiorowego zarządzania oraz autorzy (podmioty uprawnione z tytułu praw autorskich) opowiedzieli się za utrzymaniem dotychczasowych form dozwolonego użytku i wprowadzeniem praktyki umownego regulowania zasad korzystania z utworów (co daje im większą kontrolę nad sposobami i formami użytkowania utworów). Ponadto Komisja zapowiedziała podjęcie prac nad oceną skutków masowej digitalizacji i kosztów zawierania umów na podstawie których uzyskiwano by na udostępnianie utworów zawartych w kolekcjach. W *Zielonej Księdze* podjęto także problem korzystania utworów osieroconych.

W tym miejscu chciałabym raz jeszcze zasygnalizować dokument, który co prawda nie pochodzi od Komisji (organu unijnego), lecz został w pewnym sensie na jego potrzeby zamówiony i po części wpisuje się w krajobraz europejskich inicjatyw poświęconych bibliotekom cyfrowym. Mianowicie chodzi o raport grupy Sages ze stycznia 2011 r. Zawiera on następujące zalecenia[12]:

- podejmowane powinny być działania służące rozwojowi portalu Europeana, który ma stanowić główny „punkt dostępowy" i „punkt odniesienia" dla europejskiego dziedzictwa kulturowego; za pośrednictwem którego zostanie zapewniony dostęp do najważniejszych utworów znajdujących się w domenie publicznej;
- utwory, o wyczerpanym nakładzie (niedostępne w handlu) powinny być udostępnianie w internecie;
- powinno się dążyć do rozwiązania problemu utworów osieroconych, przy uwzględnieniu ośmiu istotnych warunków[13];
- zwiększenie środków na finansowanie projektów digitalizacyjnych;
- należy wspierać inicjatywy w zakresie partnerstwa publiczno-prywatnego przy inicjatywach digitalizacyjnych oraz ponadgraniczne udostępnianie zdigitalizowanych w tych projektach materiałów;
- uregulować należy system deponowania, depozytu – kopii „zapasowej" (w „archiwum" portalu *Europeana* umieścić drugą kopię cyfrową).

Kolejnym krokiem Komisji było przedstawienie w maju 2011 r. strategii dotyczącej praw własności intelektualnej w celu pobudzenia kreatywności i innowacji[14]. Wśród wskazanych kilku najistotniejszych kwestii – założeń strategii, znalazł się także ten dotyczący bibliotek cyfrowych. Komisja uznała, iż *utworzenie europejskich bi-*

[11] Komunikat z dnia 19 października 2009 r., COM(2009) 532, dostępny na stronie: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2009:0532:FIN:PL:PDF [odczyt: 2.02.2012].

[12] Polska wersja streszczenia jest dostępna na stronie: http://ec.europa.eu/information_society/activities/digital_libraries/doc/refgroup/final_report_cds.pdf [odczyt: 2.02.2012].

[13] O których mowa w pkt. 13.

[14] Dokument z dnia 24 maja 2011 r., dostępny na stronie: http://ec.europa.eu/internal_market/copyright/docs/ipr_strategy/COM_2011_287_en.pdf; http://europa.eu/rapid/pressReleasesAction.do?reference=IP/11/630&format=HTML&aged=0&language=PL [odczyt: 2.02.2012].

bliotek cyfrowych, które chronią i rozpowszechniają bogate dziedzictwo kulturowe i intelektualne Europy, ma kluczowe znaczenie dla rozwoju gospodarki opartej na wiedzy. Dla ułatwienia realizacji tego celu Komisja przygotowała i ogłosiła projekt dyrektywy o utworach osieroconych[15] (czyli opowiedziała się za wprowadzeniem pewnego ułatwienia korzystania z utworów w bibliotekach cyfrowych), ale jednocześnie uwzględniając interesy wydawców „zleciła" zawarcie protokołu ustaleń pomiędzy bibliotekami, wydawcami i organizacjami zbiorowego zarządzania, który dotyczyłby wykorzystania utworów w bibliotekach cyfrowych (w tym utworów niedostępnych w handlu – takich których nakład został wyczerpany). Moim zdaniem taki protokół mógłby mieć w szczególności na celu wyjaśnienie wątpliwości dotyczących interpretacji zakresu dozwolonych działań na podstawie art. 5 ust. 3 lit. „n" dyrektywy 2001/29/WE (i jego krajowych implementacji – byłby zatem pomocny dla interpretacji np. art. 28 pkt 2 i 3 pr. aut.).

W wyniku zapowiedzianych w tym dokumencie działań opublikowano: zieloną księgę dotycząca udostępniania online utworów audiowizualnych (*Green Paper – on the online distribution of audiovisual works in the European Union: opportunities and challenges towards a digital single market*) oraz uchwalono dyrektywę 2011/77/UE zmieniającą dyrektywę o czasie trwania praw artystów wykonawców oraz producentów fonogramów. Tak więc realizacja strategii w pewnym zakresie zaczęła nabierać realnych kształtów.

Zielona Księga w sprawie udostępniania utworów audiowizualnych w Internecie w Unii Europejskiej – możliwości i wyzwania związane z jednolitym rynkiem cyfrowym (KOM (2011) 427) pochodzi z 13 lipca 2011 r.[16]. Mam wrażenie, iż poruszana w niej problematyka ma większe znaczenie dla producentów i podmiotów dystrybuujących treści audiowizualne, telewizyjne i archiwów telewizyjnych (audiowizualnych), aniżeli bibliotek. Nie mniej warto wskazać, iż w księdze zawarto 26 szczegółowych pytań związanych z problematyką dystrybucji utworów audiowizualnych w Internecie, zainicjowano dyskusję nad m.in. takimi kwestiami jak:

- wprowadzenie „Europejskiego kodeksu prawa autorskiego";
- rejestrowanie utworów w celu otrzymania jednego unijnego (opcjonalnego) tytułu prawnego (który byłby ważny w całej UE; w praktyce chodziłoby chyba bardziej o uzyskanie dokumentu (dowodu), który potwierdzałaby prawa autorskie; należałoby się zastanowić na ile koszty (szeroko rozumiane) związane z jego obsługą byłyby adekwatne do korzyści uzyskanych z tytułu posiadania takiego dokumentu, i czy wprowadzenie tego tytułu (i co za tym idzie rejestracji) nie pozostawałaby w sprzeczności z ideą odformalizowania ochrony autorskoprawnej (w przeciwieństwie do wynalazków, znaków towarowych – ochrona autorskoprawna jest niezależna od spełnienia jakichkolwiek formalności, w tym rejestracji, powstaje automatycznie z chwilą stworzenia utworu); można się zastanowić czy ewentualną zaletą wprowadzenia nowego systemu nie było-

[15] Szczegółowo por. pkt. 13.

[16] Polska wersja językowa księgi jest dostępna na stronie: http://ec.europa.eu/internal_market/consultations/docs/2011/audiovisual/green_paper_COM2011_427_pl.pdf [odczyt: 12.01.2012].

by ograniczenie problemu utworów osieroconych wśród utworów audiowizualnych, ułatwienie wymiany informacji o podmiotach uprawnionych z tytułu praw autorskich (rozbudowanie systemu DRM); osobiście na tym wstępnym etapie nie jestem przekonana co do tego rozwiązania;

- zweryfikowanie wyjątków od treści prawa autorskiego (dozwolonego użytku publicznego) w odniesieniu do instytucji dziedzictwa filmowego; tak aby zwiększyć poziom pewności prawnej tych instytucji (obecnie wątpliwości budzi kwestia zasad dokonywania migracji mediów i formatów utworów audiowizualnych oraz przekazywanie/przesyłanie tych utworów w celu zachowania dla przyszłych pokoleń); w tym kontekście postawiono pytanie istotne dla tematu niniejszego opracowania: „Czy konieczne są zmiany legislacyjne w celu pomocy instytucjom dziedzictwa filmowego w pełnieniu zadań leżących w interesie publicznym? Czy wyjątki określone w art. 5 ust. 2 lit. c) (zwielokrotnianie w celu zachowania w bibliotekach) i w art. 5 ust. 3 lit. n) (konsultacje na miejscu dla naukowców) dyrektywy 2001/29/WE należy zmienić w celu zagwarantowania europejskim instytucjom dziedzictwa filmowego pewności prawnej w ich codziennej działalności?"; uważam, że niewątpliwie zróżnicowane sposoby implementacji tych wyjątków w poszczególnych krajach spowodowały, że niestety wbrew oczekiwaniom twórców dyrektywy 2001/29/WE doszło do powstania wątpliwości co do zakresu legalnych działań instytucji ochrony dziedzictwa kulturowego; usunięcie ich może odbyć albo poprzez działania legislacyjne (co jest procesem długotrwałym) albo odpowiednią wykładnię (interpretację) prawa – wydanie wytycznych dla tej interpretacji to drugie rozwiązanie byłoby działaniem pożądanym.

W dniu 27 września 2011 r. została uchwalona dyrektywa 2011/77/UE w sprawie zmiany dyrektywy 2006/116/WE w sprawie czasu ochrony prawa autorskiego i niektórych praw pokrewnych[17]. Skutkuje ona wydłużeniem czasu trwania praw artystów wykonawców i producentów fonogramów z 50 do 70 lat od daty [Kornecka 2011]. Zgodnie z art. 2 dyrektywy państwa członkowskie muszą do dnia 1 listopada 2013 r. dostosować prawo krajowe do nowych rozwiązań – Polska jest zobowiązana do dokonania w tym okresie nowelizacji ustawy o pr. aut., i wydłużenia czasu ochrony artystów wykonawców i producentów fonogramów. W dyrektywie przyjęto tzw. częściową retroaktywność czasu ochrony – obejmuje ona tylko, artystyczne wykonania (i ich nagrania), które w najpóźniejszej dacie implementacji dyrektywy (1.11.2013 r.) będą nadal korzystać z ochrony (czyli nie upłynęło 50 lat od ich nagrania, tj. takie które pochodzą z 1963 r. i lat następnych)[18]. Wydłużenie czasu ochrony skutkuje tym, iż te artystyczne wykonania później o 20 lat wejdą do domeny publicznej.

[17] Opublik. w Dz. U. UE - 11 10 2011, nr L 265; tekst dyrektywy – w języku polskim: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2011:265:0001:01:PL:HTML [odczyt: 12.01.2012]. Krytycznie wobec idei przedłużenia czasu ochrony praw pokrewnych wypowiadał się [Flisak 2009]. Krytyczne stanowisko wyrazili także przedstawiciele środowiska naukowego: [The Proposed Directive 2008].

[18] W komentarzach wskazuje się, że wybór zasad przedłużenia ochrony jest nieprzypadkowy – gdyż ma skutkować objęciem ochroną tzw. evergreeny z lat 60-tych, np. nagrań The Beatles, The Rolling Stones, The Who, The Doors czy Jimiego Hendrixa.

Kolejnym zadaniem, które postawiła sobie do rozwiązania Komisja jest określenie zasad dotyczących korzystania z utworów o wyczerpanym nakładzie [Stanisławska-Kloc 2007] (niedostępnych w handlu). Komisja podjęła się pośrednictwa w zawarciu porozumienia służącego ułatwieniu dostępu do książek o wyczerpanym nakładzie[19]. Jak podano w komunikacie podpisany we wrześniu 2011 r. protokół zawiera pakiet kluczowych zasad, ustalonych pomiędzy bibliotekami, wydawcami, autorami oraz reprezentującymi ich organizacjami zbiorowego zarządzania, które umożliwią europejskim bibliotekom i innym instytucjom kulturalnym digitalizację i udostępnianie w internecie znajdujących się w ich zbiorach książek i czasopism naukowych o wyczerpanym nakładzie. Podstawowa zasada jaką przyjęto jest następująca: prawo do digitalizacji i udostępnienia dzieł o wyczerpanym nakładzie przysługuje w pierwszej kolejności podmiotom praw autorskich, a biblioteki, instytucje kultury na podstawie dobrowolnych umów licencyjnych będą uzyskiwać prawo do digitalizacji i udostępniania w Internecie części ich zasobów (tj. przechowywanych przez nie książek i czasopism o wyczerpanym nakładzie). Jak wskazano w komunikacie przyjęcie protokołu ma duże znaczenie dla dalszego rozwoju europejskich bibliotek cyfrowych oraz portalu *Europeana*.

4. Regulacje międzynarodowe

Najstarsza, pochodząca z 1886 r. umowa międzynarodowa to konwencja dotycząca ochrony dzieł literackich i artystycznych (opubl. w Dz. U. z 1990 r., nr 80, poz. 474), do której należy ponad 160 państw świata[20]. Zakres ochrony na podstawie tej konwencji wyznaczają dwie podstawowe zasady [J. Barta, R. Markiewicz, Prawo autorskie, 2010] międzynarodowego prawa autorskiego:

a) zasada traktowania krajowego (określana także jako zasada asymilacja), z której wynika obowiązek zapewnienia autorom-cudzoziemcom ochrony na takim samym poziomie, jak jest przewidziana dla twórców polskich; zasada asymilacji nakazuje przy ustalaniu poziomu ochrony autorskoprawnej nie czynić różnic ze względu na pochodzenie/obywatelstwo autora, czyli traktować twórców obcych tak jak własnych twórców, tak jak „krajowców" [Barta; Markiewicz 2008, s. 265; Barczewski 2007];
b) zasada minimum ochrony (w konwencji wprowadzono pewne minimalne standardy ochrony, które muszą być zapewnione we wszystkich krajach, które przystąpiły do niej).

Pozostałe konwencje z zakresu prawa autorskiego i praw pokrewnych to: konwencja rzymska z 1961 r. o ochronie wykonawców, producentów fonogramów oraz

[19] Por. Komunikat Komisji z 20 września 2011 r. dotyczący podpisania protokołu ustaleń z http://europa.eu/rapid/pressReleasesAction.do?reference=IP/11/1055&format=HTML&aged=0&language=PL&guiLanguage=en [odczyt: 12.01.2012].

[20] Aktualna lista państw będących stroną konwencji berneńskiej oraz jej tekst (podobnie jak teksty innych ważnych konwencji międzynarodowych z zakresu prawa autorskiego) jest dostępna na stronie internetowej Światowej Organizacji Własności Intelektualnej (WIPO), http://www.wipo.int/treaties/en. Polski tekst konwencji zrewidowanej w 1971 r. w Paryżu (załącznik do Dz. U. z 1990 r., nr 82, poz. 474).

organizacji nadawczych (opubl. w Dz. U. z 1997 r., nr 125, poz. 800), Porozumienie TRIPS (Trade Related Aspects of the Intelectual Property Rights) które stanowi załącznik do porozumienia WTO; (opubl. w Dz. U. z 1996 r., nr 32, poz. 143); konwencje WIPO z 1996 r.: o prawie autorskim (opubl. w Dz. U. z 2005 r., nr 3, poz. 12) oraz o artystycznych wykonaniach i fonogramach (opubl. w Dz. U. z 2004 r., nr 41, poz. 375).

Generalnie w obowiązujących obecnie regulacjach międzynarodowych brak przepisów szczególnych, które odnosiły by się wprost do bibliotek. Zarówno konwencja berneńska, jak i konwencja WIPO, nie zawierają postanowień dotyczących formy dozwolonego użytku, której beneficjentem byłyby biblioteki. Komentatorzy wskazują, iż odpowiednio na podstawie art. 9 ust. 2 i art. 10 ust. 1 przywołanych konwencji dozwolone jest konstruowanie w prawie krajowym instytucji dozwolonego użytku (tzw. wyjątków od treści prawa autorskiego) na rzecz tych instytucji. Z opcji tej skorzystało 128 spośród 184 państw członkowskich WIPO; pierwszym krajem, który wprowadził wyjątek dla bibliotek była prawdopodobnie Wielka Brytania i miało to miejsce w 1956 r. [Crews 2008]. Obecnie przedmiotem prac, konsultacji w ramach WIPO są wyjątki dotyczące wykorzystania utworów przez biblioteki i archiwa oraz wyjątek na rzecz osób niepełnosprawnych[21].

5. Poza autorskie regulacje prawne

Pomimo, że celem tego opracowania jest głównie naświetlenie kwestii autorskoprawnych związanych z funkcjonowaniem bibliotek cyfrowych to jednak przydatne może być wzmiankowanie także innych, wybranych, regulacji prawnych i związanych z nimi zagadnień, które mają znaczenie dla funkcjonowania bibliotek.

a) Ustawa o ochronie danych osobowych

Ustawa z dnia 29 sierpnia 1997 r., o ochronie danych osobowych (tekst pierwotny: Dz. U. 1997 r., nr 133, poz. 883, tekst jednolity: Dz. U. 2002 r., nr 101, poz. 926, ze zm.; dalej „ust. ochr. d. osob.") określa zasady postępowania przy przetwarzaniu danych osobowych oraz prawa osób fizycznych, których dane osobowe są lub mogą być przetwarzane w zbiorach danych [Barta, Fajgielski; Markiewicz 2007; Barta, Litwiński 2009].

Zgodnie z art. 2 „ust. ochr. d. osob." określa zasady postępowania przy przetwarzaniu danych osobowych oraz prawa osób fizycznych, których dane osobowe są lub mogą być przetwarzane w zbiorach danych. Ustawę stosuje się do przetwarzania danych osobowych:

– w kartotekach, skorowidzach, księgach, wykazach i w innych zbiorach ewidencyjnych,

[21] Por. http://www.wipo.int/copyright/en/limitations/index.html [odczyt: 12.01.2012].

– w systemach informatycznych, także w przypadku przetwarzania danych poza zbiorem danych.

Natomiast w odniesieniu do zbiorów danych osobowych sporządzanych doraźnie, wyłącznie ze względów technicznych, szkoleniowych lub w związku z dydaktyką w szkołach wyższych, a po ich wykorzystaniu niezwłocznie usuwanych albo poddanych anonimizacji, mają zastosowanie jedynie przepisy rozdziału 5 (dot. zabezpieczenia danych).

Zakresem podmiotowym ustawa obejmuje:

– organy państwowe, organy samorządu terytorialnego oraz do państwowych i komunalnych jednostek organizacyjnych; podmioty niepubliczne realizujących zadania publiczne,
– osoby fizyczne i osoby prawne oraz jednostki organizacyjne niebędące osobami prawnymi, jeżeli przetwarzają dane osobowe w związku z działalnością zarobkową, zawodową lub dla realizacji celów statutowych,
– które mają siedzibę albo miejsce zamieszkania na terytorium Rzeczypospolitej Polskiej, albo w państwie trzecim, o ile przetwarzają dane osobowe przy wykorzystaniu środków technicznych znajdujących się na terytorium Rzeczypospolitej Polskiej.

Natomiast ustawa nie ma zastosowania do:

– osób fizycznych, które przetwarzają dane wyłącznie w celach osobistych lub domowych,
– podmiotów mających siedzibę lub miejsce zamieszkania w państwie trzecim, wykorzystujących środki techniczne znajdujące się na terytorium Rzeczypospolitej Polskiej wyłącznie do przekazywania danych.

Ustawa zawiera definicję danych osobowych (art. 6 „ust. ochr. d. osob."), którymi są: wszelkie informacje dotyczące osoby fizycznej, zidentyfikowanej lub możliwej do zidentyfikowania; osobą możliwą do zidentyfikowania jest osoba, której tożsamość można określić bezpośrednio lub pośrednio, w szczególności przez powołanie się na numer identyfikacyjny albo jeden lub kilka specyficznych czynników określających jej cechy fizyczne, fizjologiczne, umysłowe, ekonomiczne, kulturowe lub społeczne; informacji nie uważa się za umożliwiającą określenie tożsamości osoby, jeżeli wymagałoby to nadmiernych kosztów, czasu lub działań.

Mimo, że nie wynika to wprost z brzmienia ustawy przyjmuje się, iż dotyczy ona danych osobowych osób żyjących. Do danych osobowych poza imieniem i nazwiskiem, można zaliczyć adres emailowy, wizerunek osoby (fotografię), informacje dotyczące wykształcenia, numer PESEL [Barta; Fajgielski; Markiewicz, 2007 Barta, Litwiński 2009].

W tzw. słowniczku (art. 7„ust. ochr. d. osob.") zawarto definicje najważniejszych pojęć m.in.:

– zbiór danych – każdy posiadający strukturę zestaw danych o charakterze osobowym, dostępnych według określonych kryteriów, niezależnie od tego, czy zestaw ten jest rozproszony lub podzielony funkcjonalnie,

- przetwarzanie danych – jakiekolwiek operacje wykonywane na danych osobowych, takie jak zbieranie, utrwalanie, przechowywanie, opracowywanie, zmienianie, udostępnianie i usuwanie, a zwłaszcza te, które wykonuje się w systemach informatycznych,
- usuwanie danych – zniszczenie danych osobowych lub taką ich modyfikację, która nie pozwoli na ustalenie tożsamości osoby, której dane dotyczą,
- administrator danych – to organ, jednostka organizacyjna, podmiot lub osoba, o których mowa w art. 3, decydujące o celach i środkach przetwarzania danych osobowych,
- zgodzie osoby, której dane dotyczą – rozumie się przez to oświadczenie woli, którego treścią jest zgoda na przetwarzanie danych osobowych tego, kto składa oświadczenie; zgoda nie może być domniemana lub dorozumiana z oświadczenia woli o innej treści,

Podstawą zgodnego z prawem przetwarzania danych jest zgoda podmiotu prawa (osoby, której dane dotyczą), pozostałe 4 podstawy legalności przetwarzania są wymienione w art. 23 „ust. ochr. d. osob." Zasadniczo administrator zbioru danych ma obowiązek jego rejestracji, aczkolwiek liczne wyjątki od tego obowiązku zawiera art. 43 ust. 1„ust. ochr. d. osob." Na przykład zwolnieni z obowiązku rejestracji są administratorzy danych:

- dotyczących osób korzystających z ich usług medycznych, obsługi notarialnej, adwokackiej, radcy prawnego, rzecznika patentowego, doradcy podatkowego lub biegłego rewidenta,
- przetwarzanych wyłącznie w celu wystawienia faktury, rachunku lub prowadzenia sprawozdawczości finansowej,
- powszechnie dostępnych,
- przetwarzanych w celu przygotowania rozprawy wymaganej do uzyskania dyplomu ukończenia szkoły wyższej lub stopnia naukowego,
- przetwarzanych w zakresie drobnych bieżących spraw życia codziennego.

Biblioteki są zobowiązane rejestrować zbiory danych osobowych czytelników (użytkowników)[22], z obowiązku tego zwolnione są uczelniane biblioteki (art. 88 ust. o szkol. wyższ.[23]). W Uniwersytecie Jagiellońskim regulamin określający zasady ochrony danych osobowych znajdujących się w systemie informacji bibliotecznej stanowi załącznik nr 1c do Statutu Uniwersytetu Jagiellońskiego[24].

[22] Rejestracja odbywa się poprzez platformę e-giodo, dostęp: http://egiodo.giodo.gov.pl/index.dhtml [odczyt: 15.01.2012]. Obecnie bibliotek 2335 zarejestrowało zbiory czytelników (użytkowników). Dane na podstawie informacji zawartych w wyszukiwarce na stronie e-giodo.

[23] Ustawa o szkolnictwie wyższym – z dnia 27 lipca 2005 r., (Dz. U. 164, poz. 1365, z późn. zm.) – art. 88 „1. W uczelni działa system biblioteczno-informacyjny, którego podstawę stanowi biblioteka. Organizację i funkcjonowanie systemu biblioteczno-informacyjnego uczelni, w tym zasady korzystania z niego przez osoby niebędące pracownikami, doktorantami lub studentami uczelni, określa statut. (...). 4. Uczelnia w związku z funkcjonowaniem systemu biblioteczno-informacyjnego może przetwarzać określone w jej statucie dane osobowe osób korzystających z tego systemu. 5. Zbiór danych osobowych, o których mowa w ust. 4, jest zwolniony z obowiązku rejestracji zbiorów danych osobowych, o których mowa w art. 40 ustawy z dnia 29 sierpnia 1997 r. o ochronie danych osobowych (Dz. U. z 2002 r. Nr 101, poz. 926 i Nr 153, poz. 1271 oraz z 2004 r. Nr 25, poz. 219 i Nr 33, poz. 285)."

[24] http://www.uj.edu.pl/documents/10172/f9db507c-b869-4d9a-84dc-9cd52268f17d [odczyt: 15.01.2012].

Przetwarzanie danych może odbywać się tylko w celu, który został objęty zgodą podmiotu albo wynika z przepisów prawa, a w innym celu niż ten, dla którego zostały zebrane, jest to dopuszczalne, jeżeli nie narusza praw i wolności osoby, której dane dotyczą, oraz następuje w celach badań naukowych, dydaktycznych, historycznych lub statystycznych.

Ustawa zabezpiecza prawa osób, których dane są przetwarzane i statuuje na ich rzecz konkretne uprawnienia np. prawo kontroli danych, uzyskiwania informacji o źródłach z których dane pochodzą, weryfikacji danych.

Organem, do spraw ochrony danych osobowych jest Generalny Inspektor Ochrony Danych Osobowych (GIODO), na którego stronie internetowej[25] można znaleźć wiele cennych, praktycznych informacji, jak i dokumentów (aktów prawnych, decyzji administracyjnych), wyjaśnień[26] z zakresu ochrony danych osobowych.

Na trzy kwestie istotne z punktu widzenia działalności bibliotecznej należy zwrócić uwagę. Po pierwsze, dane osobowe czytelników podlegają ochronie na podstawie tej ustawy; zatem biblioteka powinna uzyskać zgodę czytelnika (użytkownika) na przetwarzanie jego podstawowych danych np. imienia i nazwiska, adresu zamieszkania.

Po drugie, należy wskazać, iż dane „wytwarzane" w toku świadczenia usług bibliotecznych np. dotyczące wypożyczanych pozycji, czynionych przez użytkownika zamówień, mogą mieć charakter danych osobowych, mogą dotyczyć sfery prywatności użytkownika [Iglezakis 2011].

Po trzecie, ustawa wyraźnie w art. 3 „a" „ust. ochr. d. osob." stanowi, iż nie stosuje się jej przepisów (za wyjątkiem art. 14 – 19 i 36 ust. 1„ust. ochr. d. osob.") do prasowej działalności dziennikarskiej w rozumieniu ustawy z dnia 26 stycznia 1984 r. – Prawo prasowe (Dz. U. Nr 5, poz. 24, z późn. zm.) oraz do działalności literackiej lub artystycznej, chyba że wolność wyrażania swoich poglądów i rozpowszechniania informacji istotnie narusza prawa i wolności osoby, której dane dotyczą.

Publikacja w prasie danych osobowych jest zaliczana do prasowej działalności dziennikarskiej, a ta nie podlega ustawie o ochronie danych. Zgodnie ze stanowiskiem GIODO: *Ustawy o ochronie danych osobowych nie stosuje się więc do zbierania, opracowywania i przygotowywania, a także publikacji materiałów w prasie. Podstawę prawną działalności dziennikarskiej stanowią przepisy prawa prasowego*[27].

Uważam, że dane wykorzystane w utworach literackich – książkach (np. słownikach, opracowaniach biograficznych, pamiętnikach, kronikach, ale także indeksach, wykazach bibliograficznych), które znajdują się w zbiorach (kolekcjach) bibliotecznych nie podlegają ustawie o ochronie danych osobowych (także wtedy, gdy dotyczą autorów książek, zawartych w katalogach)[28]. Przy czym wyraźnie należy odróżnić dwie sytuacje: a) informacje, dane nie znajdują się w zbiorze – np. nazwiska osób

[25] http://www.giodo.gov.pl/168/j/pl [odczyt: 15.01.2012].

[26] Por. np. wyjaśnienia GIODO dotyczące zatrzymywania przez biblioteki dowodów osobistych czytelników; dostęp: http://www.giodo.gov.pl/360/id_art/3375/j/pl/ [odczyt: 15.01.2012].

[27] Wyjaśnienie GIODO, dostęp: http: //www.giodo.gov.pl/398/id_art/3377/j/pl/ [odczyt: 15.01.2012].

[28] Także czynności wyszukiwania według nazwiska autora nie powinny być traktowane jako przetwarzanie danych objęte ustawą.

wymienione w tekście artykułu biograficznego – wtedy w ogóle nie ma zastosowania ustawa jako taka; b) informacje znajdują się w zbiorze – ale równocześnie zbiór ten stanowi przejaw/wynik działalności literackiej lub artystycznej czy też powstał na potrzeby/w ramach działalności literackiej (w tym o charakterze naukowym np. indeksy) lub artystycznej – wtedy znajdzie miejsce wyłączenie z art. 3 „a" „ust. ochr. d. osob.". W przywołanym przepisie nie ma odwołania do ustawy o prawie autorskim i prawach pokrewnych, ale uważam, że każda działalność literacka i artystyczna będzie objęta zakresem tego wyłączenia[29]. Swoje stanowisko w tej sprawie wyraził GIODO: „Publikacje w prasie i książkach, zawierające dane osobowe w formie nieuporządkowanej nie stanowią przetwarzania danych osobowych w zbiorze, a więc nie podlegają przepisom ustawy o ochronie danych osobowych. (...) Podstawowe znaczenie ma tutaj definicja zbioru. Zgodnie z art. 7 pkt 1 „ust. ochr. d. osob." zbiorem danych jest każdy posiadający strukturę zestaw danych o charakterze osobowym, dostępnych według określonych kryteriów, niezależnie od tego, czy zestaw ten jest rozproszony czy podzielony funkcjonalnie. Dostępność dotyczy tu konkretnych danych osobowych, a nie innych informacji, które mogą w sposób pośredni umożliwić dostęp do szukanych danych. Publikacje w prasie i książkach z samej istoty są publikacjami, w których pojawiają się dane osobowe, jednak zasadnicze znaczenie ma tutaj fakt, czy mamy do czynienia z przetwarzaniem w zbiorze. Niewątpliwie zgodnie z definicją określoną w art. 7 pkt 1 „ust. ochr. d. osob." za przetwarzanie w zbiorze uznać należy publikowanie nazwisk w formie list, not biograficznych, czy też opracowań encyklopedycznych lub leksykalnych. Kryterium dostępu jest tu jasno określone i uporządkowane strukturalnie, zazwyczaj w formie alfabetycznej, czasami z dodatkowymi kryteriami (dziedzina, rok ukończenia szkoły itp.). Informacji publikowanych w formie nieuporządkowanej, a więc np. nazwisk, które wymienia się w artykułach, nie można uznać za formę przetwarzania danych osobowych w zbiorze. Tym samym nie podlegają one przepisom ustawy o ochronie danych"[30].

b) ustawa o partnerstwie publiczno–prywatnym

Ustawa o partnerstwie publiczno-prywatnym z 19 grudnia 2008 r. (Dz. U. z 2009 r., nr 19, poz. 100 ze zm.) określa zasady współpracy podmiotu publicznego i partnera prywatnego w ramach partnerstwa publiczno-prywatnego (ppp). Przedmiotem ppp jest wspólna realizacja przedsięwzięcia (służącego realizacji zadania pub-

[29] Uważa, że termin działalność literacka i artystyczna, należy interpretować szeroko obejmując nim wszystkie kategorie utworów ze sfery wyrażanej słowem (w tym naukowe) oraz artystycznej; na uzasadnienie takiego poglądu przywołuję konwencję berneńską z 1886 r. (która w swoim tytule „ogranicza" się tylko do wymienienia dzieł literackich oraz artystycznych, a w treści odnosi się do wszelkich kategorii utworów) oraz dyrektywę Parlamentu Europejskiego i Rady UE z 24 października 1995 r. (95/46/EC, opubl. w Dz. Urz. UE, L 281, z dnia 23 listopada 1995 r.) w sprawie ochrony osób fizycznych w zakresie przetwarzania danych osobowych i swobodnego przepływu tych danych (art. 9 - „the purpose of artistic or literary expression").

[30] Wyjaśnienie GIODO http://www.giodo.gov.pl/320/id_art/978/j/pl/ [odczyt: 15.01.2012].

licznego) oparta na podziale zadań i ryzyk pomiędzy podmiotem publicznym i partnerem prywatnym[31].

W niektórych krajach instytucja ppp była wykorzystywana do realizacji projektów digitalizacyjnych [Stanisławska-Kloc 2009]. Już w Zaleceniach z 2006 r. (w pkt 7 preambuły) wskazywano, iż: *Sponsorowanie digitalizacji przez sektor prywatny oraz partnerstwa publiczno-prywatne mogą prowadzić do włączenia podmiotów prywatnych do starań w zakresie digitalizacji i powinny być nadal wspierane.*

Z kolei w raporcie grupy Sages[32], generalnie pozytywnie oceniono wykorzystanie ppp w dziedzinie digitalizacji. Stwierdzono, iż finansowanie projektów digitalizacyjnych może odbywać się w ramach ppp. Przy czym wskazano, podstawowe warunki – swoiste dobre praktyki w tym zakresie. Po pierwsze, aby chronić interesy instytucji publicznych należy podawać do wiadomości publicznej treść umowy ppp zawartej pomiędzy publiczną instytucją kultury (np. biblioteką) i partnerem prywatnym. Po drugie, zdigitalizowane materiały stanowiące własność publiczną (należące do domeny publicznej) powinny być dostępne bezpłatnie dla użytkowników we wszystkich państwach członkowskich. Po trzecie, partner prywatny powinien dostarczyć instytucji kultury zdigitalizowane pliki tej samej jakości jak te z których sam korzysta. Po czwarte, maksymalny okres obowiązywania preferencyjnych warunków korzystania przez partnera prywatnego z materiałów zdigitalizowanych w ramach ppp nie powinien przekraczać siedmiu lat. Uważam, że zalecenia te powinny być brane pod uwagę przez polskie instytucje kultury decydujące się na realizowanie projektów digitalizacyjnych w ramach instytucji ppp.

Generalnie należałoby promować tę formę współpracy pomiędzy instytucjami kultury i podmiotami prywatnymi (komercyjnymi), która jak wydaje się, w porównaniu do tradycyjnego sponsoringu, mogłaby być odbierana jako bardziej szlachetna w „wydźwięku"[33] społecznym – stanowiąca nowe oblicze instytucji mecenasa kultury.

c) prawo prasowe

Na wzmiankowanie zasługuje regulacja dotycząca prasy / działalności prasowej – prawo prasowe z dnia 26 stycznia 1984 r. (Dz. U. z 1984 r., nr 5, poz.24, ze zm., dalej „pr. pras."). Art. 7 „pr. pras." zawiera definicję prasy: *1) prasa oznacza publikacje periodyczne, które nie tworzą zamkniętej, jednorodnej całości, ukazujące się nie rzadziej niż*

[31] Wiele informacji dotyczących ppp jest dostępnych na stronie „Instytutu partnerstwa publiczno-prywatnego", http://ippp.pl/ [odczyt: 15.01.2012].

[32] Por. streszczenie – s. 7 oraz pkt. – s. 42-46 raportu grupy Sages [2011].

[33] Chociaż jak wskazuje G. Pawelska-Skrzypek [2003, s. 196], wstępnie pozytywnie przyjęte we Francji w latach 80-tych przedsięwzięcia ppp w sferze kultury, z uwagi na ich małą rentowność nie zyskały szerszego zastosowania. O zawarciu ppp z Google informowała Austriacka Biblioteka Narodowa: *Przechowywanie i udostępnianie wiedzy od zawsze było najważniejszym celem bibliotek. Zgodnie z tą fundamentalną zasadą Austriacka Biblioteka Narodowa planuje cyfryzację całości swojego zbioru nieobjętych prawami autorskimi pozycji historycznych i udostępnienie ich w Internecie. W tym celu biblioteka zawarła z firmą Google partnerstwo publiczno-prywatne, które umożliwi udostępnienie internautom 400 000 najważniejszych dzieł z księgozbioru biblioteki. Jesteśmy dumni, że możemy w ten sposób przyczynić się do demokratyzacji wiedzy*, więcej informacji http://www.chip.pl/news/wydarzenia/umowy-i-fuzje/2010/06/austriacka-biblioteka-narodowa-udostepni-w-internecie-caly-swoj-ksiegozbior-historyczny#ixzz1kAT5l3yl [odczyt: 15.01.2012].

raz do roku, opatrzone stałym tytułem albo nazwą, numerem bieżącym i datą, a w szczególności: dzienniki i czasopisma, serwisy agencyjne, stałe przekazy teleksowe, biuletyny, programy radiowe i telewizyjne oraz kroniki filmowe; 2) prasą są także wszelkie istniejące i powstające w wyniku postępu technicznego środki masowego przekazywania, w tym także rozgłośnie oraz tele– i radiowęzły zakładowe, upowszechniające publikacje periodyczne za pomocą druku, wizji, fonii lub innej techniki rozpowszechniania; 3) prasa obejmuje również zespoły ludzi i poszczególne osoby zajmujące się działalnością dziennikarską.

Generalnie należy stwierdzić, iż definicja prasy jest szeroka i pozwala na objęcie nią niektórych „form" przekazywania informacji dokonywanych w internecie. Na tym też tle powstało zagadnienie, które ma znaczenie dla funkcjonowania bibliotek cyfrowych (gdyż posiadają one swoje strony internetowe), a mianowicie dotyczące rejestrowania stron internetowych.

Ustawa wyróżnia wśród form prasowych (art. 7 „pr. pras."):

a) dziennik – ogólnoinformacyjny druk periodyczny lub przekaz za pomocą dźwięku oraz dźwięku i obrazu, ukazujący się częściej niż raz w tygodniu,
b) czasopismo – druk periodyczny ukazujący się nie częściej niż raz w tygodniu, a nie rzadziej niż raz w roku; przepis ten stosuje się odpowiednio do przekazu za pomocą dźwięku oraz dźwięku i obrazu innego niż określony w pkt 2.

Na te dwie w/w formy prasowe nałożono obowiązek (art. 20 „pr. pras.") rejestracji[34] w sądzie okręgowym właściwym miejscowo dla siedziby wydawcy. Wniosek o rejestrację, powinien zawierać: 1) tytuł dziennika lub czasopisma oraz siedzibę i dokładny adres redakcji, 2) dane osobowe redaktora naczelnego, 3) określenie wydawcy, jego siedzibę i dokładny adres, 4) częstotliwość ukazywania się dziennika lub czasopisma. Wydawanie dziennika lub czasopisma można rozpocząć, jeżeli organ rejestracyjny nie rozstrzygnął wniosku o rejestrację w ciągu 30 dni od jego zgłoszenia. Ponowne rozpowszechnianie prasy (np. archiwalnej) w ramach bibliotek cyfrowych nie podlega regułom prawa prasowego (nie stanowi działalności prasowej).

Podstawowy problem dotyczył rozstrzygnięcia czy wszystkie (ewentualnie które) strony internetowe zostałyby zaliczone do dzienników lub czasopism i objęte obowiązkiem rejestracji. Odpowiedź na to pytanie tym bardziej była istotna, że wydawanie dziennika i czasopisma bez rejestracji skutkowało ponoszeniem odpowiedzialności karnej – niezarejestrowanie dziennika lub czasopisma stanowiło do grudnia 2011 r. przestępstwo (art. 45 „pr. pras."[35]).

Omawiane zagadnienie było przedmiotem rozstrzygnięć Sądu Najwyższego[36], dopiero komentarze oraz postanowienie Sądu Najwyższego – Izby Karnej z dnia 15 grudnia 2010 r., sygn. III KK 250/10):

[34] Szczegółowo por. Rozporządzenie Ministra Sprawiedliwości z 9 lipca 1990 r. (Dz.U.90.46.275, ze zm.) w sprawie rejestru dzienników i czasopism. Rejestr (jawny, ogólnodostępny) dzienników i czasopism prowadzą Sądy Okręgowe, wydziały cywilne [Dobosz 2011; Nowińska 2008].

[35] Przepis ten utracił moc, na podstawie wyroku TK z dn. 14.12.2011 r. sygn. akt SK 42/09 (Dz. U. Nr 282, poz. 1660), w zakresie, w jakim wprowadza odpowiedzialność karną za wydawanie czasopisma drukowanego bez rejestracji.

[36] Pierwsze orzeczenie wywołało wiele wątpliwości – postanowienie z dnia 26 lipca 2007 r. (IV KK 174/07, LEX nr 287505, Prok.i Pr.-wkł. 2008/2/4): *1. Osoba rozpowszechniająca bez rejestracji w właściwym sądzie okręgowym dziennik bądź czasopismo za pośrednictwem Internetu zarówno wówczas, gdy prze-*

Prasy ukazującej się w Internecie nie wolno utożsamiać z witryną internetową. Ustawodawca nie nałożył obowiązku rejestracji na strony czy witryny internetowe, a objął tym obowiązkiem jedynie prasę, a ściśle rzecz biorąc – dzienniki i czasopisma.

Przesądziło, iż nie każda strona (witryna) internetowa stanowi prasę (a tym bardziej dziennik lub czasopismo) [Barta; Markiewicz 2008; Barta; Markiewicz 2007; Barta; Markiewicz 1998; Maciąg; Stanisławska 1997; Nowińska 2008]. Dzienniki i czasopisma to publikacje periodyczne, jednakże periodyczności nie można utożsamiać – czy też dopatrywać się w uaktualnianiu zawartości strony czynionym nawet kilka razy dziennie. Zasadniczo należy uznać, iż większość stron internetowych bibliotek cyfrowych (podobnie jak stron wielu instytucji państwowych) nie spełnia wymogu periodyczności wydawania (która powtórzę nie jest tożsama z częstą aktualizacją), nie stanowi dziennika ani czasopisma i stąd nie podlega obowiązkowi rejestracji. Natomiast niektóre strony internetowe mogą być zaliczane do prasy, w ujęciu szerokim (art. 7 ust. 2) – jako powstały w wyniku postępu technicznego nowy środek masowego przekazywania, służący do upowszechniania publikacji periodycznych.

d) egzemplarz obowiązkowy

Kwestie związane z obowiązkiem wydawców w zakresie przekazywania egzemplarzy utworów reguluje ustawa z dnia 7 listopada 1996 r., o obowiązkowych egzemplarzach bibliotecznych (Dz. U. 1996 r., nr 152, poz. 722 ze zm. 2008 r., nr 10, poz. 8, dalej „egz. obow.")[37].

Obecnie 15 bibliotek[38] jest upoważnionych do otrzymywania takich egzemplarzy – rozwiązanie to jest krytykowane, jako zbyt obciążające wydawców [Marciszuk 2010; Byford 2005].

---

kaz taki towarzyszy przekazowi utrwalonemu na papierze, stanowiąc inną elektroniczną jego postać, jak i wówczas, gdy istnieje tylko w formie elektronicznej w Internecie, wyczerpuje znamiona przestępstwa z art. 45 ustawy – Prawo prasowe. 2. Zasada ochrony zaufania obywatela do państwa i prawa wyraża się w takim stanowieniu, jak i stosowaniu prawa, aby nie stawało się ono swoistą „pułapką" dla obywatela i aby mógł on układać swoje sprawy w zaufaniu, iż nie naraża się na prawne skutki, których nie mógł przewidzieć w momencie podejmowania decyzji oraz w przekonaniu, iż jego działania podejmowane zgodnie z obowiązującym prawem i zgodnie z wykładnią stosowaną przez organy tego państwa są zgodne z porządkiem prawnym. Ustawodawca wyraźnie i jednoznacznie stwierdza, że prasą są zarówno dzienniki i czasopisma, jak i „wszelkie istniejące i powstające w wyniku postępu technicznego środki masowego przekazywania (...) upowszechniające publikacje periodyczne za pomocą druku, wizji, fonii lub innej techniki rozpowszechniania" – art. 7 ust. 2 pkt 1 in fine pr.pr. W tej sytuacji jest rzeczą bezsporną, że dzienniki i czasopisma przez to że ukazują się w formie przekazu internetowego, nie tracą znamion tytułu prasowego, i to zarówno wówczas gdy przekaz internetowy towarzyszy przekazowi utrwalonemu na papierze, drukowanemu, stanowiąc inną, elektroniczną jego postać w systemie on line, jak i wówczas, gdy przekaz istnieje tylko w formie elektronicznej w Internecie, ale ukazuje się tylko periodycznie, spełniając wymogi, o których mowa w art. 7 ust. 2 pr.pr.

[37] Rozporządzeniem Ministra Kultury i Sztuki z dnia 6 marca 1997 r. w sprawie wykazu bibliotek uprawnionych do otrzymywania egzemplarzy obowiązkowych oraz zasad ich przekazywania (Dz. U. z 1997 r., nr 29, poz. 161, ze zm.). Por. [Dobosz 2011, s. 111-115] oraz w szczególności w zakresie historycznych oraz obowiązujących regulacji [Czarny-Drożdziejko 2008, s. 121-141].

[38] Wykaz jest dostępny na stronie BN, dostęp: http://www.bn.org.pl/zbiory/egzemplarz-obowiazkowy/wykaz-bibliotek-uprawnionych-do-otrzymywania-bibliotecznych-egzemplarzy-obowiazkowych [odczyt: 12.12.2011].

Zakres przedmiotowy regulacji dotyczącej egzemplarzy obowiązkowych jest szeroki – zgodnie z art. 2 ust. 2 „egz. obow." obejmuje egzemplarze publikacji[39], czyli dzieł[40] zwielokrotnionych dowolną techniką w celu rozpowszechnienia, a w szczególności a) piśmiennicze (np. broszury, gazety, czasopisma, i inne wydawnictwa ciągłe, druki, ulotne, afisze), b) graficzne i graficzno-piśmiennicze, jak mapy, plakaty, plany, wykresy, tabele, rysunki, ilustracje, nuty; c) audiowizualne utrwalające dźwięk, obraz lub obraz i dźwięk, jak płyty, taśmy, kasty, przeźrocza, mikrofilmy, mikroklisze; d) zapisane na informatycznych nośnikach danych; e) oprogramowanie komputerowe. Przekazane egzemplarze biblioteka nabywa na własność. Tryb i zasady przekazywania egzemplarzy reguluje rozporządzenie MKiSz z 6 marca 1997 r. w sprawie wykazu bibliotek uprawnionych do otrzymywania egzemplarzy obowiązkowych poszczególnych rodzajów publikacji oraz zasad i trybu ich przekazywania (Dz. U. z 1997 r., nr 29, poz. 161). Zgodnie z par. 4 ust. 2 rozporządzenia : jeżeli na publikacji nie jest podana wysokość nakładu, wydawca zobowiązany jest podać tę informację przy odpowiedniej pozycji wykazu dołączonego do przesyłki dla Biblioteki Narodowej. Biblioteka Narodowa korzysta z tych danych wyłącznie dla celów urzędowej statystyki wydawnictw i traktuje je jako poufne. Jak wskazuje E. Czarny-Drożdziejko, ustawodawca skonstruował w tym zakresie specyficzną tajemnicę zawodową bibliotekarzy[41].

W aspekcie funkcjonowania bibliotek cyfrowych istotne znaczenie ma objęcie egzemplarzem obowiązkowym „publikacji zapisanych na informatycznych nośnikach danych", czyli publikacji elektronicznych (utrwalonych np. na płytach CD, DVD), które będą mogły być udostępniane czytelnikom, na podstawie art. 28 pkt 3 pr. aut.

Rozważenia wymaga, m.in. w kontekście Web-harvestingu, objęcie obowiązkiem przekazywania (lub nawet „samodzielnego" pozyskiwania przez biblioteki) w ramach „egzemplarza" obowiązkowego publikacji (treści) udostępnianych na statycznych oraz dynamicznych stronach internetowych oraz opracowanie wspólnych standardów technicznych (technologiczno-informacyjnych), tak aby w przyszłości zasoby te mogły być przechowywane (long-term preservation), scalane i udostępniane[42].

e) adresy internetowe (domeny internetowe) bibliotek cyfrowych

Nazwy bibliotek cyfrowych (nazwy/tytuły stron internetowych za pośrednictwem których biblioteki udostępniają swoje zbiory) zasadniczo nie mają charakteru utwo-

[39] Zakres ten rozszerza jeszcze art. 3 ust. 3, a z kolei zawęża art. 4.

[40] Dzieła nie należy utożsamiać tylko i wyłącznie z utworem w rozumieniu prawa autorskiego; tj. dziełem będą zarówno wytwory intelektu które mają charakter utworu, jak i takie które są go pozbawione czy (bez względu na to, czy z powodu wyłączenia na podstawie art. 4, czy dlatego że nie mają charakteru twórczego). Por. na ten temat E. Czarny-Drożdziejko, op. cit. str. 135. Jak wydaje się termin dokumenty (tak postuluje E. Czarny-Drożdziejko) lub materiały byłby bardziej adekwatny.

[41] [Czarny-Drożdziejko 2008, s. 139].

[42] Por. Projekt rekomendacji Komisji (Commission Recommendation of xxx on the digitisation and online accessibility of cultural material and digital preservation; o którym informacja została opatrzona tytułem: The Commission calls on governments to adapt legislation to make more in-copyright and out-of-commerce material available on Europeana), dostęp: http://ec.europa.eu/information_society/activities/digital_libraries/doc/recommendation/new_recommandation28nov11/en_recommendation.pdf

rów, nie podlegają ochronie autorskoprawnej. Natomiast nazwy (tytuły) takie mogą być rejestrowane jako domeny internetowe[43].

f) znaki towarowe

Podstawową polską regulacją, która odnosi się do znaków towarowych jest ustawa prawo własności przemysłowej z dnia 30 czerwca 2000 r. (t.j. Dz. U. z 2003 r., nr 119, poz.1117, ze zm., dalej „pr. wł. przem."), Znaki towarowe służą do odróżniania towarów pochodzących od jednego przedsiębiorstwa od towarów pochodzących od innego przedsiębiorstwa. Zgodnie z art. 120 „pr. wł. przem.": 1. Znakiem towarowym może być każde oznaczenie, które można przedstawić w sposób graficzny, jeżeli oznaczenie takie nadaje się do odróżnienia towarów jednego przedsiębiorstwa od towarów innego przedsiębiorstwa. 2. Znakiem towarowym, w rozumieniu ust. 1, może być w szczególności wyraz, rysunek, ornament, kompozycja kolorystyczna, forma przestrzenna, w tym forma towaru lub opakowania, a także melodia lub inny sygnał dźwiękowy. Najczęściej w praktyce występują znaki słowne lub słowno-graficzne. Zasadniczo można uzyskać ochronę znaków na poziomie krajowym (poprzez rejestrację w krajowym urzędzie patentowym – Urząd Patentowy RP), wspólnotowym – wspólnotowy znak towarowy uzyskuje się poprzez rejestracje w OHIM), międzynarodowym (w wybranych krajach świata – poprzez rejestrację za pośrednictwem WIPO- Światowej Organizacji Ochrony Własności Intelektualnej). Zgodnie z art. 153 „pr. wł. przem.": 1. Przez uzyskanie prawa ochronnego nabywa się prawo wyłącznego używania znaku towarowego w sposób zarobkowy lub zawodowy na całym obszarze Rzeczypospolitej Polskiej. 2. Czas trwania prawa ochronnego na znak towarowy wynosi 10 lat od daty zgłoszenia znaku towarowego w Urzędzie Patentowym. 3. Prawo ochronne na znak towarowy może zostać, na wniosek uprawnionego, przedłużone dla wszystkich lub części towarów, na kolejne okresy dziesięcioletnie.

Rejestracji dokonuje się dla wybranych klas towarów i/lub usług; obecnie obowiązuje tzw. IX edycja klasyfikacji nicejskiej, która zawiera 34 klasy towarów i 11 klas usług. Ponieważ biblioteki zasadniczo nie prowadzą działalności gospodarczej, nie wprowadzają na rynek towarów, nie świadczą usług to też bardzo nieliczne biblioteki zarejestrowały znaki towarowe. Na przykład w Polsce został zarejestrowany znak słowny „BIBLIOTEKA NARODOWA ZAKŁAD NARODOWY IM. OSSOLIŃSKICH – WYDAWNICTWO" (numer ochronny 173325), a wśród wspólnotowych – wspólnotowy znak towarowy British Library (dwa figuratywne-słowno-graficzne, jeden słowny – np. CTM 008960452). Najczęściej rejestracje dokonywane są dla działalności wydawniczej (albumów, książek), usługowej (organizowanie targów, wystaw). Dla zapewnienia podstawowego działania bibliotek cyfrowych, rejestracja nazwy pod którą działa biblioteka w formie znaku towarowego nie jest niezbędna. Przykładowo zarejestrowano 6 wspólnotowych znaków towarowych zawierających słowo

[43] Na temat problematyki adresów internetowych por. Justyna Ożegalska-Trybalska (2003). *Adresy internetowe*. Kraków: Zakamycze, 506 s.

„europeana", w tym m.in. słowno-graficzny (o numerze 010036853) dla Stichting Europeana, fundacji, która prowadzi portal Europeana.

Wybrane zagadnienia szczegółowe dotyczące problematyki autorskoprawnej

6. Kategorie przedmiotów ochrony, które są wykorzystywane w bibliotecznych kolekcjach cyfrowych

Zasadniczo w bibliotekach (w tym cyfrowych) wykorzystywane są niemal wszystkie kategorie utworów – tj. w zbiorach tych instytucji mogą znaleźć się oryginalne egzemplarze, jak i reprodukcje wszystkich utworów, oraz przedmiotów praw pokrewnych. Podstawowe założenia ochrony utworów są identyczne, dla wszystkich kategorii, w pewnym uproszczeniu można stwierdzić, że istotne zróżnicowanie dotyczy tylko zakresu dozwolonego użytku osobistego i publicznego. Tak więc w wielu przypadkach (poza programami komputerowymi i bazami danych) kwalifikacja do konkretnej, wymienionej przykładowo w art. 1 pr. aut.; kategorii utworów, w kontekście działalności bibliotek pozbawiona jest doniosłości prawnej.

Definicję utworu zawiera art. 1 „pr. aut.": *Przedmiotem prawa autorskiego jest każdy przejaw działalności twórczej o indywidualnym charakterze, ustalony w jakiejkolwiek postaci, niezależnie od wartości, przeznaczenia i sposobu wyrażenia (utwór).* Dodatkowo można ją uzupełnić poprzez odwołanie do art. 2, który wskazuje na tzw. utwory zależne (*Opracowanie cudzego utworu, w szczególności tłumaczenie, przeróbka, adaptacja, jest przedmiotem prawa autorskiego bez uszczerbku dla prawa do utworu pierwotnego*) oraz art. 3, który odnosi się do zbiorów, zestawień (*Zbiory, antologie, wybory, bazy danych są przedmiotem prawa autorskiego, nawet jeżeli zawierają nie chronione materiały, o ile przyjęty w nich dobór, układ lub zestawienie ma twórczy charakter, bez uszczerbku dla praw do wykorzystanych utworów*).

Analiza stanowiska doktryny, jak i orzecznictwa wskazuje na dosyć liberalne (niekiedy co prawda krytykowane) kryteria oceny spełnienia przesłanki twórczości, jakie funkcjonują w polskim, jak i europejskim prawie autorskim. Przykładowo w jednym z orzeczeń Trybunału Sprawiedliwości UE przyjęto, iż fragment tekstu składający się z 11 znaków czy logo organizacji może mieć charakter twórczy[44]. W tym kontekście uzasadnione jest stawianie tezy, iż wiele spisów treści[45] utworów literackich (książek),

[44] Odpowiednio dotyczy to orzeczeń: z 16 lipca 2009 r. w sprawie C-5/08 (utworem może już być ciąg 11 słów), z 4 października 2011 r. w sprawach C-403/08 i C-429/08 (utworem może być m.in. logo ligi piłki nożnej).

[45] Poza przypadkami, gdy składają się z kilku banalnych, oczywistych elementów np. 1. Przedmiot ochrony, 2. Podmiot praw, 3. Treść praw, 4. Ograniczenia praw; są oczywiste, zdeterminowane celem pracy. W przypadku spisów treści czasopism, o ile nie ma on charakteru alfabetycznego, to w innych

jak i okładek książek będzie miało charakter twórczy, i będzie podlegać ochronie jako utwory, choć oczywiście ocena spełnienia twórczego charakteru będzie zawsze dokonywana *casu ad casum*. Twórczy charakter mają, a tym samym, stanowią utwory prace: zaliczeniowe, licencjackie, magisterskie[46], dyplomowe.

Jako przykład kategorii wytworów pozbawionych charakteru twórczego można wskazać fotografie reprodukcyjne, legitymacyjne[47]; pozostałe kategorie fotografii (w tym reporterskie, przedstawiające krajobraz, osoby) z reguły będą miały charakter twórczy.

Nie podlegają w ogóle ochronie wytwory wymienione w art. 4 pr. aut.: 1) akty normatywne lub ich urzędowe projekty, 2) urzędowe dokumenty, materiały, znaki i symbole, 3) opublikowane opisy patentowe lub ochronne, 4) proste informacje prasowe), z których ewentualnie największe znaczenie, z uwagi na ich występowanie w zbiorach bibliotek mogą mieć akty normatywne i dokumenty oraz materiały o charakterze urzędowym. Zasadniczo dokumenty partii politycznych, opinie zamawiane na zlecenie: organów państwowych, premiera, prezydenta, dokumenty organizacji międzynarodowych nie mają charakteru dokumentów urzędowych; podobnie jak strony www organów państwowych, zdjęcia osób pełniących funkcje polityczne/urzędowe.

Wymienione przykładowo w art. 5 ust. o bibl. materiały biblioteczne to: *w szczególności dokumenty zawierające utrwalony wyraz myśli ludzkiej, przeznaczone do rozpowszechniania, niezależnie od nośnika fizycznego i sposobu zapisu treści, a zwłaszcza: dokumenty graficzne (piśmiennicze, kartograficzne, ikonograficzne i muzyczne), dźwiękowe, wizualne, audiowizualne i elektroniczne*. Poza utworami wyrażonymi słowami (głównie utworami literackimi, często wzbogaconymi o materiał ilustracyjny), w zbiorach bibliotek znajdują się utwory kartograficzne (mapy), plastyczne (ich reprodukcje), muzyczne (wraz z ich zapisem nutowym , wraz z artystycznymi wykonaniami, utrwalonymi na płytach CD), fotograficzne, filmowe (utrwalone na płytach DVD), bazy danych, programy komputerowe, oraz przedmioty praw pokrewnych (fonogramy, wideogramy, nagrania audycji radiowych i telewizyjnych).

przypadkach stosunkowo często trudno będzie dopatrywać się twórczości w wyborze i układzie artykułów kwalifikowanych do danego numeru. Mam wrażenie, iż często okoliczności organizacyjne, proces recenzyjny decydują o ostatecznej zawartości, stąd można twierdzić, że w układzie i wyborze trudno będzie doszukać się twórczości, aczkolwiek nie przesądzam tej kwestii jednoznacznie.

[46] W mojej ocenie żadną miarą nie jest uzasadniony pogląd, niestety czasem reprezentowany, iż prace studentów stanowią rezultat działalności o charakterze odtwórczym pozbawionym cech twórczości. Z założenia praca, licencjacka, magisterska ma mieć charakter samodzielny, a nawet jeśli jej autor referuje cudze poglądy, dokonuje swoistego zabiegu sprawozdawczego to nie wyłącza to swobody twórczej. Czysto odtwórcza praca licencjacka lub magisterska nie powinna być przyjmowana i stanowić podstawy nadania tytułu zawodowego.

[47] W jednym z ostatnich orzeczeń Trybunału Sprawiedliwości w sprawie o sygn. akt. C-145/10, (opinia rzecznika generalnego Trstenjaka, z dnia 12 kwietnia 2011 r.), przyjęto, iż zdjęcie dziecka (quasi portretowe) może mieć charakter twórczy; dostęp: http://curia.europa.eu/jurisp/cgi-bin/gettext.pl?lang=en&num=79889587C19100145&doc=T&ouvert=T&seance=CONCL; orzeczenie z dnia 1 grudnia 2011 r., dostęp: http://curia.europa.eu/juris/document/document.jsf?text=&docid=115785&pageIndex=0&doclang=pl&mode=doc&dir=&occ=first&part=1&cid=99146 [odczyt: 12.12.2011].

Ponieważ zarówno definicja materiałów bibliotecznych, jak i utworu jest szeroka, niejako otwarta to nie ma przeszkód aby każdy istniejący, jak i taki który powstanie w przyszłości rodzaj utworu, a który znajdzie się w zbiorach bibliotecznych zakwalifikować do materiałów bibliotecznych. Z kolei każdy materiał biblioteczny, który spełnia cechu utworu, będzie podlegać ochronie autorskoprawnej.

Przypomnieć należy, iż ochrona autorskoprawna powstaje bez wymogu spełnienia jakichkolwiek formalności; tj. nie wymaga dokonania zgłoszenia, rejestracji, zresztą w Polsce nie funkcjonuje żaden urząd, instytucja, która przyjmowałaby utwory i wydawała poświadczenia ich twórczego charakteru (art. 1 ust. 4: „Ochrona przysługuje twórcy niezależnie od spełnienia jakichkolwiek formalności"). Dla przyznania ochrony nie mają znaczenia następujące oznaczenia/oświadczenia autora lub wydawcy umieszczane na utworach:

a) noty copyrightowi (np. „Copyright by A. Kowalski, Kraków 2011"),
b) zastrzeżenia praw autorskich („All rights reserved", „Some rights reserved"),
c) zakazy kopiowania całości lub części utworu („Wszystkie prawa autorskie i prawa producenta do nagranych utworów zastrzeżone. Kopiowanie, wynajmowanie, wypożyczanie, wykorzystywanie tego nośnika do wykonań publicznych i nadań RTV bez zezwolenia zabronione").

Nota copyrightowi służy głównie celom informacyjnym [Barta; Markiewicz 2011; Stanisławska-Kloc; Ptak 2007] – wskazuje, kto jest podmiotem autorskich praw majątkowych (uprawnionym z tytułu tych praw), z kim kontaktować się celem zawarcia umowy o wykorzystanie utworu. Zwyczajowo składa się ona z : litery „c" umieszczonej w kółeczku, nazwiska (nazwy) podmiotu, któremu przysługują autorskie prawa majątkowe oraz roku pierwszej publikacji (© 2007 by XY; © by XY; 2007, (c), (C)), powszechnie używa się także zwrotu „Copyright by XY". Pochodzi ona z konwencji powszechnej o prawie autorskim z 1952 r., natomiast w polskiej ustawie o prawie autorskim nie została szczegółowo uregulowana, to też nie ma wiążących reguł dotyczących tego kto powinien być wskazany w jej treści (autor, wydawca, licencjobiorca), co więcej zwyczaje w tym zakresie są dosyć zróżnicowane.

Natomiast ostatnia z przywołanych form oświadczeń nie wyłącza/nieogranicza[48] możliwości kopiowania a nawet rozpowszechniania utworów na podstawie dozwolonego użytku czy to osobistego (art. 23 pr. aut.) czy publicznego (np. art. 27, 28 pr. aut.). Tak więc biblioteka może egzemplarz opatrzony takim oświadczeniem np. skopiować na podstawie art. 28 ust. 2 pr. aut.[49].

[48] Wyjątek dotyczy art. 25 ust. 1 pkt 1 lit/. „b" pr. aut.; oraz może być dyskutowany na gruncie wykładni art. 5 ust. 3 lit. „n" dyrektywy 2001/29/WE.

[49] Z informacji prasowych wynika, iż już w 2006 r. Prezes UOKiK wszczął postępowania przeciwko podmiotom dokonującym dystrybucji filmów na płytach DVD – Warner Bros i Vision Film. UOKiK zakwestionował umieszczanie na okładkach płyt informacji sugerujących całkowity zakaz pożyczania oraz kopiowania utworu bez zgody właściciela praw autorskich. Napisy te zostały zmodyfikowane. Por. informacje zawarte na stronie UOKIK, http://www.uokik.gov.pl/aktualnosci.php?news_id=263 [odczyt: 12.12.2011].

Na koniec tej części rozważań podkreślenia wymaga, iż strona internetowa (portal, jak i podstrony) biblioteki cyfrowej może stanowić utwór[50]. Twórczy charakter portalu może wynikać m.in. z wyboru, układu, zestawienia poszczególnych jej elementów (np. słownych, graficznych, fotograficznych).

7. Domena publiczna – „wolne" treści wykorzystywane w kolekcjach cyfrowych

Termin domena publiczna nie został zdefiniowany w przepisach prawa autorskiego, nie występuje wprost w ustawach prawnoautorskich[51]. Powszechnie, potocznie przyjmuje się, że do domeny publicznej należą te rezultaty pracy intelektualnej:

- które nie są chronione prawami własności intelektualnej[52], w tym prawem autorskim
- pomysły, idee (art. 1 ust. 2^1 pr. aut. – które nie spełniają cechy twórczości),
- materiały (wyłączone spod ochrony na podstawie art. 4 pr. aut.),
- w stosunku do których ochrona wygasła – np. takie utwory w stosunku do których ochrona autorskich praw majątkowych wygasła (najczęściej, gdy upłynęło 70 lat od śmierci autora – art. 36 pr. aut. – lub utwór został stworzony w czasach, gdy ochrona prawnoautorska nie obowiązywała).

Te utwory, materiały, utwory które znajdują się w domenie publicznej mogą być swobodnie wykorzystywane, w bibliotekach cyfrowych; tj. bez ograniczeń co do sposobu, czasu, miejsca – stąd mówimy, iż są to treści „wolne", nieograniczone monopolem twórcy. Natomiast należy pamiętać, że pomimo wygaśnięcia praw majątkowych, cały czas trwają autorskie prawa osobiste i w związku z tym niezbędne jest oznaczanie autorstwa wykorzystywanych utworów, nienaruszanie ich treści i/lub formy.

Domena publiczna przeciwstawiana jest domenie twórcy – treści chronionej prawem autorskim na rzecz konkretnego twórcy. Projektem UE, poświęconym domenie publicznej jest Communia[53]. Natomiast za pośrednictwem strony Koalicji Otwartej Edukacji można zapoznać się z polska wersją tzw. kalkulatora domeny publicznej[54], który stanowi narzędzie ułatwiające ustalenie czy utwór może znajdować się w domenie publicznej. W wielu przypadkach wynik testu może być wiarygodny, ale trzeba mieć na uwadze, że nie zawsze na jego podstawie w sposób pewny (bezbłędny) zo-

[50] W jednym z orzeczeń przyjęto, iż „stworzenie układu i formy graficznej portalu, jak również jej ulepszanie i zmiany w okresie eksploatacji portalu internetowego mieści się w definicji utworu, o jakim mowa w art. 1 ust. 1 ustawy z dnia 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych (wyrok SA w Warszawie, z dnia 11 maja 2007 r., sygn. akt I ACa 1145/2006). Por. Barta; Markiewicz 1998, s. 120; Stanisławska-Kloc 1999, s. 32-35; Flisak 2008; Michalak 2010.

[51] Na temat domeny publicznej zob. Targosz 2007; Stec 2008. Por. też raport *WIPO – Committee on Development and Intellectual Property* 2011; Barta; Markiewicz 2010, s. 427.

[52] Termin ten funkcjonuje także na gruncie np. prawa patentowego.

[53] Dostęp: http://communia-project.eu/ [odczyt: 12.12.2011].

[54] Został on przygotowany na podstawie schematu rozpowszechnianego w haśle „Domena publiczna" w *Wikipedii* http://pl.wikipedia.org/wiki/Domena_publiczna; z tym, że kalkulator jest bardziej szczegółowy, http://domena.nowoczesnapolska.org.pl [odczyt: 12.12.2011].

stanie określony status utworu jako należącego do domeny publicznej (gdyż wiele okoliczności dotyczących twórcy, jak i utworu, szczególnie w przypadku utworów anonimowych, publikowanych pod pseudonimem, pracowniczych wpływa na zakwalifikowanie utworu do domeny publicznej). Niedawno został zgłoszony interesujący pomysł powołania rzecznika (ombudsmana) domeny publicznej[55], jak na razie nie został on zrealizowany.

Utwór, który wszedł do domeny publicznej nie może „z powrotem" uzyskać ochrony autorsko-prawnej; „wrócić" do sfery objętej prawem autorskim. Digitalizacja utworu twórcy, który żył w XIX w. nie skutkuje powstaniem czy to na rzecz osoby dokonującej digitalizacji czy spadkobierców autora jakichkolwiek praw autorskich. Tylko w sytuacji, gdy dojdzie do twórczego opracowania takiego utworu – np. tłumaczenia, to wtedy będziemy mieć do czynienia z nowy utworem, który będzie podlegać ochronie przez okres życia tłumacza i 70 lat po jego śmierci.

W wytycznych dotyczących digitalizacji zbiorów bibliotecznych wskazuje się, aby nie umieszczać na cyfrowych wersjach utworów należących do domeny publicznej tzw. znaków wodnych z oznaczeniami praw autorskich na rzecz podmiotu, który dokonał digitalizacji. W opisach katalogowych takich utworów, przy oznaczaniu ich statusu powinna zostać umieszczona informacja o przynależności do domeny publicznej. Rozpowszechnianie utworu, który znajduje się w domenie publicznej wraz z chronioną prawem autorskim przedmową i/lub wstępem nie ma wpływu na odżycie ochrony tego utworu „zasadniczego".

Utwory należące do domeny publicznej niekiedy oznaczane są poniżej zaprezentowanym symbolem (nieoficjalnym, nie pochodzącym z żadnej regulacji prawnej):



Moim zdaniem utworów w domenie publicznej nie można mylić z utworami rozpowszechnianymi na tzw. licencjach wolnościowych typu CC. Zasadnicza różnica pomiędzy tymi kategoriami utworów wynika z tego, iż utwory udostępniane na licencjach CC nadal podlegają ochronie (nie wygasły w stosunku do nich autorskie prawa majątkowe). Jednakże twórcy zdecydowali się na szerokie udostępnianie utworów użytkownikom, na zasadach określonych w tych licencjach, które to zasady przede wszystkim nie nakładają obowiązku indywidualnego negocjowania, kontaktowania się pomiędzy twórcą a użytkownikiem. Sama umowa zostaje zawarta w momencie, gdy użytkownik zdecyduje się na skorzystanie z utworu na warunkach licencji CC (np. poprzez ściągnięcie utworu z sieci), użytkownik nie musi informować autora o rozpoczęciu korzystania, czy składać mu jakichkolwiek oświadczeń pisemnych. Dopóki użytkownik przestrzega zasad licencji (tekst lub link do strony www, na której

[55] Por. Peukert, Alexander (2011). *A European Public Domain Supervisor.* "IIC" nr 2 vol 42, s. 125 -129.

znajduje się tekst licencji jest dołączany do utworu, podawany/rozpowszechniany wraz z utworem) dopóty legalnie korzysta z utworu, na podstawie licencji udzielonej mu przez autora. Najdalej idącą licencją CC, która ma wywoływać „skutek"[56] taki, jak gdyby utwór był w domenie publicznej (swoiste „przeniesienie" utworu do domeny publicznej), a co za tym idzie, autor zrzekłby się niemal wszelkich możliwych praw jest licencja CC0 („Creative Commons 0")[57].

Na marginesie trzeba wskazać, iż do utworów rozpowszechnianych za pośrednictwem bibliotek cyfrowych nie będzie mieć zastosowania art. 40 pr. aut. – który statuuje obowiązek uiszczenia opłat (określanych niekiedy nazwą: „domain public payant"), w przypadku korzystania, z niepodlegających już ochronie w zakresie autorskich praw majątkowych, utworów (literackich, muzycznych, plastycznych, fotograficznych i kartograficznych). Ponieważ przesłanką uiszczania tych opłat jest uzyskanie wpływów ze sprzedaży egzemplarzy wymienionych powyżej kategorii utworów opublikowanych na terytorium RP – nie mamy do czynienia z jej spełnieniem w przypadku umieszczania/udostępniania utworów w bibliotekach cyfrowych, stąd też brak podstaw do nakładania tego obowiązku na biblioteki cyfrowe.

Dzień[58] domeny publicznej jest obchodzony każdego roku w dniu 1 stycznia, gdyż wtedy trafiają do domeny publicznej utwory, których twórcy zmarli 70 lat temu (np. utwory wszystkich twórców, którzy zmarli od dnia 1 stycznia do dnia 31 grudnia 1941 r., w dniu 1 stycznia 2012 r. trafiły do domeny – por. art. 39 pr. aut.).

8. Kategorie podmiotów, którym mogą przysługiwać prawa do treści wykorzystywanych w kolekcjach cyfrowych

Skrótowe omówienie kategorii podmiotów, którym mogą przysługiwać prawa autorskie ma na celu ułatwienie zidentyfikowania z jakimi kategoriami tych podmiotów mogą spotkać się w praktyce osoby, które będą tworzyć biblioteki cyfrowe, od kogo będą nabywać prawa autorskie do utworów wykorzystywanych w tych bibliotekach.

Każdorazowo przystępując do korzystania z utworu należy ustalić czy trwają jeszcze autorskie prawa majątkowe i komu przysługują:

- jednemu twórcy lub kilku współtwórcom (i w jakim zakresie),
- pracodawcy,
- osobie (osobom), która nabyła na podstawie umowy (przenoszącej lub licencyjnej),
- spadkobiercom.

Zasadniczo ustalenie czasu trwania autorskich praw majątkowych wymaga posiadania informacji o pierwotny autorze (twórcy – jego imieniu i nazwisku, które

[56] Przy założeniu, iż dany system prawny zezwala na zrzeczenie się praw autorskich; na gruncie prawa polskiego jest to nieskuteczne (por. art. 16 i 18 pr. aut.).

[57] Na temat licencji CC por. pkt 12.

[58] Funkcjonuje strona dnia domeny publicznej – dostęp: http://domenapubliczna.org/

pozwalają na jego zidentyfikowanie i ustalenie czy żyje czy też zmarł i upłynęło już 70 lat od jego śmierci[59]).

Pomimo, że kwestia noty copyrightowej nie jest szczegółowo uregulowana w przepisach prawa to zwyczajowo przyjmuje się, że w nocie wskazywana jest osoba, której przysługują autorskie prawa majątkowe i w czasie gdy one jeszcze trwają. Ustawa o pr. aut. posługuje się terminem twórca (którego synonimem w języku potocznym jest „autor") lub uprawniony; z tym, że w wielu przypadkach czynione przez ustawodawcę rozróżnienie terminologiczne nie ma charakteru jednoznacznego, uporządkowanego – stąd uzasadnione jest przyjęcie uogólnienia, iż terminy: twórca i uprawniony są używane[60] w ustawie zamiennie.

W odniesieniu do utworów współautorskich zawsze należy pamiętać o ustaleniu udziału poszczególnych twórców w prawie autorskim, umowa powinna być zawarta ze wszystkim współtwórcami (stroną umowy na podstawie, której biblioteka ma uzyskać prawo do korzystania z utworu jest każdy współtwórca, natomiast nie musi on osobiście podpisywać umowy może upoważnić do tego np. innego współtwórcę).

Jeżeli chodzi o utwory pracownicze uregulowane w art. 12 pr. aut. (inne aniżeli programy komputerowe, do których stosuje się przepis szczególny z art. 74 ust. 4), to należy mieć na uwadze, następujące istotne kwestie:

- pracodawca nabywa tylko prawa majątkowe, a prawa osobiste (np. do oznaczenia autorstwa) pozostają przy twórcy;
- pracodawca nabywa prawa tylko w stosunku do utworów stworzonych w ramach wykonywania obowiązków wynikających z umowy o pracę (zakres ten określa umowa o pracę, aneksy do umowy, regulamin pracy, a czasem może być doprecyzowany przez ustne lub pisemne polecenie pracodawcy);
- wykonanie przez pracownika utworu na podstawie zawartej z pracodawcą umowy o dzieło lub zlecenia nie powoduje nabycia przez pracodawcę praw autorskich na zasadach określonych w art. 12.

Szczególny status autorskoprawny posiadają tzw. utwory zbiorowe (art. 11), które często znajdują się w zbiorach bibliotecznych (np. encyklopedie, leksykony, czasopisma), wynika on z tego, że można w stosunku do tych utworów wyróżnić dwie płaszczyzny praw:

- do całości (w znaczeniu, układu, zestawienia), które przysługują wydawcy,
- do części składowych np. poszczególnych haseł encyklopedycznych, artykułów prasowych, które przysługują ich twórcom, poprzez wydawcę).

Na podstawie umów zawieranych pomiędzy autorem w wydawcą, ten drugi nabywa prawa autorskie majątkowe do utworu (wkładu) inkorporowanego do utworu zbiorowego. Art. 11 nie stanowi samoistnej podstawy nabycia praw autorskich do wszystkich części składowych utworu zbiorowego.

Po śmierci twórcy autorskie prawa majątkowe wchodzą w skład spadku i podlegają dziedziczeniu. Spadkobiercami są osoby, które nabyły autorskie prawa majątkowe

[59] Zasada ta obejmuje także utwory pracownicze.

[60] Aczkolwiek kwestia ta jest dyskusyjna. Na liczne problemy związane z brakiem jednolitego, uporządkowanego, przemyślanego nazewnictwa zwracają uwagę: Barta; Markiewicz 2010, s. 70-71.

na podstawie orzeczenia sądowego (postanowienia o stwierdzeniu nabycia spadku) lub notarialnego potwierdzenia dziedziczenia. Sam testament (o ile twórca go pozostawił) nie stanowi dokumentu potwierdzającego dziedziczenie, dopiero na jego podstawie sąd lub notariusz mogą wydać dokumenty, które w obrocie cywilnoprawnym stanowią dowód nabycia praw po zmarłym twórcy. Jeżeli twórca nie pozostawił ważnego testamentu to krąg osób powołanych do dziedziczenia oraz ich udział ustala się według zasad określonych w przepisach księgi IV k.c., zatytułowanej „Spadki" (art. 922 i nast. kc.).

W celu ułatwienia wykazaniu faktu przysługiwania praw autorskich należy:

- archiwizować dowody stworzenia utworu, projekty (wstępne wersje) utworu, przygotować i prowadzić własny rejestr (wykaz) utworów – bazę własnych utworów, umieszczać na egzemplarzach utworów noty copyrightowi i wyraźne oznaczać autorstwo (współautorstwo);
- archiwizować dowody nabycia praw autorskich od osób trzecich – archiwizować umowy, korespondencję prowadzoną z pracownikami, z osobami które wykonywały utwór na podstawie np. umowy o dzieło, odbierać oświadczenia z których np. wynika, że osobie przekazującej nam utwór przysługują prawa autorskie.

W kontekście omawiana podmiotu praw autorskich zasygnalizować należy problem utworów osieroconych[61], czyli takie, których twórców (tj. ich tożsamość) trudno ustalić lub trudno z nimi nawiązać kontakt. Wskazane okoliczności nie powodują modyfikacji zasad ochrony, nie wyłączają ochrony w stosunku do tych utworów, utwór osierocony nie jest utworem „niczyim".

9. Działalność organizacji zbiorowego zarządzania

Zgodnie z art. 104 pr. aut. organizacjami zbiorowego zarządzania prawami autorskimi lub prawami pokrewnymi (w skrócie „ozz") są stowarzyszenia zrzeszające twórców, artystów wykonawców, producentów lub organizacje radiowe i telewizyjne, których statutowym zadaniem jest zbiorowe zarządzanie i ochrona powierzonych im praw autorskich lub praw pokrewnych oraz wykonywanie uprawnień wynikających z ustawy.

Idea działania „ozz" związana jest także z tym, że korzystanie z utworów odbywa się na coraz większą skalę, zarówno użytkownicy, jak i sami uprawnieni (twórcy, ich spadkobiercy) nie byliby w stanie (z powodów organizacyjnych, finansowych) kontrolować i zawierać indywidualnie umów. Na przykład właściciel dyskoteki czy restauracji, w której odtwarzane są setki utworów musiałby w związku z odtwarzaniem każdej piosenki zawrzeć umowę z autorem: słów, aranżacji, warstwy muzycznej, z artystami wykonawcami, producentem fonogramu, indywidualnie negocjować umowy z setkami osób – obecnie w wielu przypadkach wystarczy, że zostanie zawarta umowa za pośrednictwem odpowiednich ozz. Z pośrednictwa ozz korzystają stacje radiowe

[61] Zagadnienie utworów osieroconych jest omawiane w pkt. 13.

i telewizyjne, które nadają utwory. Ponadto niektóre formy dozwolonego użytku (np. przewidziane w art. 30 ust. 2, 20, 20[1] pr. aut.) wymagają zapłaty wynagrodzenia, które odbywa się za pośrednictwem ozz[62]. „Ozz" pobierają opłaty z tytułu sprzedaży urządzeń, które służą do zwielokrotniania utworów (magnetowidy, kserokopiarki, skanery) oraz tzw. czystych nośników (służących do utrwalania utworów). Uzyskane przez ozz kwoty (po odliczeniu kosztów działalności ozz) są wypłacane uprawnionym (tj. twórcom, artystom wykonawcom, wydawcom), potocznie określane one są terminem „tantiemy autorskie".

W Polsce działają następujące ozz: Stowarzyszenie Artystów Wykonawców Utworów Muzycznych i Słowno-Muzycznych (SAWP), Związek Polskich Artystów Plastyków (ZPAP), Związek Artystów Scen Polskich (ZASP), Stowarzyszenie Autorów ZAiKS, Stowarzyszenie Filmowców Polskich, Stowarzyszenie Twórców Ludowych, Związek Producentów Audio-Video (ZPAV), Związek Polskich Artystów Fotografików, Związek Stowarzyszeń Artystów Wykonawców (STOART), Stowarzyszenie Zbiorowego Zarządzania Prawami Autorskimi Twórców Dzieł Naukowych i Technicznych (KOPIPOL), Stowarzyszenie Architektów Polskich (SARP), Stowarzyszenie Autorów i Wydawców „Polska Książka".

Członkiem ozz może być: twórca, artysta wykonawca, producent, organizacje radiowe i telewizyjne. Na podstawie umów o wzajemnej reprezentacji, polskie „ozz" reprezentują prawa twórców zrzeszonych w zagranicznych „ozz". Do celów statutowych ozz należy: zbiorowe zarządzanie prawami autorskimi i pokrewnymi; ochrona powierzonych praw; wykonywanie innych uprawnień wynikających z ustawy. Podjęcie przez ozz działalności wymaga zezwolenia Ministra Kultury i Dziedzictwa Narodowego, natomiast tabele wynagrodzeń za korzystanie z utworów i przedmiotów praw pokrewnych (w zakresie objętym zarządem ozz) zatwierdza powołana przez MKiDZN Komisja Prawa Autorskiego. Często tabele te są przywoływane także w sporach z tytułu naruszenia autorskich praw majątkowych; służą jako podstawa dla obliczania stosownego wynagrodzenia, o którym mowa w art. 79 ust. 1 pkt 3 „b" pr. aut.[63].

W odniesieniu do działalności bibliotek (zarówno tradycyjnych, jak i cyfrowych), obecnie funkcjonowanie „ozz" nie odgrywa on istotnej roli (tzn. biblioteki nie korzystają z utworów w taki sposób, który wiązałby się z pośrednictwem ozz), sytuacja może się zmienić o ile:

- zostaną wprowadzone opłaty z tytułu wypożyczeń utworów (tzw. public lending rights), wtedy opłaty te byłyby pobierane/dystrybuowane przez upoważnione ozz (takie które uzyskałyby zgodę MKiDZN na zarząd w tym zakresie);
- zostanie wprowadzone pośrednictwo ozz (jak np. ma to miejsce w krajach nordyckich) w zakresie korzystania z utworów osieroconych.

[62] Zob. w tej kwestii orzeczenie Trybunału Sprawiedliwości (trzecia izba) z dnia 21 października 2010 r. (wniosek o wydanie orzeczenia w trybie prejudycjalnym złożony przez Audiencia Provincial de Barcelona – Hiszpania) w sprawie Padawan SL przeciwko Sociedad General de Autores y Editores (SGAE), sygn. sprawy C-467/08, sentencja orzeczenia opublikowana w Dz. U. UE z 2010, seria C, 2010 nr 346, s. 8-9.

[63] Przykładowo tabele Związku Polskich Artystów Fotografików dostępne są na stronie : http://www.zpaf.waw.pl/tabele-minimalnych-wynagrodze.html [odczyt: 12.12.2011].

10. Licencje ustawowe jako podstawa wykorzystywania utworów chronionych prawem autorskim (i przedmiotów praw pokrewnych) w bibliotekach cyfrowych

Terminem licencje ustawowe określane są przepisy art. 23-33[5] pr. aut., które statuują na rzecz użytkowników możliwość korzystania z utworu (w przeciwieństwie do licencji umownej tutaj mamy do czynienia z licencją wynikającą z przepisu ustawy), gdy trwają autorskie prawa majątkowe, bez konieczności uzyskiwania zgody twórcy. Podkreślić należy, iż instytucja dozwolonego użytku ma znaczenie dla zapewnienia „usprawiedliwionego" korzystania z utworów, gdy trwają jeszcze autorskie prawa majątkowe. Zakres tego usprawiedliwionego korzystania wynika m.in. z konieczności zapewnienia dostępu do informacji, postępu i rozwoju w zakresie nauki, sztuki. Gdy prawa te wygasną utwór może być zwielokrotniany, wykorzystywany publicznie bez ograniczeń, poza tymi wynikającymi z obowiązku przestrzegania autorskich praw osobistych (prawa do autorstwa, prawa do integralności). Zasadniczo (poza wyjątkiem z art. 25 pr. aut.), licencja ustawowa nie może być wyłączona, ograniczona na podstawie jednostronnego oświadczenia twórcy, wydawcy. Na przykład oświadczenia zakazujące kopiowania utworu (umieszczane na stronach redakcyjnych książek) nie wyłączają możliwości skserowania książki w punkcie ksero w ramach dozwolonego użytku osobistego. Przepisy o dozwolonym użytku utworów stosuje się do przedmiotów praw pokrewnych (por. art. 100 pr. aut.)

Dozwolony użytek dzieli się na dwie formy:

- użytek prywatny (osobisty), uregulowany w art. 23 pr. aut.
- użytek publiczny, uregulowany w art. 24-33[5] pr. aut.

Do obydwu form użytku mają zastosowanie reguły wyrażone w art. 34 (obowiązek podania nazwiska i źródła) oraz art. 35 („dozwolony użytek nie może naruszać normalnego korzystania z utworu lub godzić w słuszne interesy twórcy"). Przy czym klauzula generalna z art. 35 rodzi trudności interpretacyjne[64], jest krytykowana jako niejasna; wywodząca się z regulacji międzynarodowej może być przydatna jako wskazówka dla ustawodawcy, aby nie tworzył ram prawnych dla zbyt szerokiego dozwolonego użytku. Natomiast klauzula nie powinna być stosowana jako instrument indywidualnego korygowania (w kierunku zawężenia) przez uprawnionych zakresu dozwolonego użytku (np. jako uzasadnienie ograniczenia kopiowania tylko 10% lub jednego rozdziału w ramach dozwolonego użytku osobistego). Aprobata dla szerokiego stosowania tej klauzuli stanowiłaby istotne zagrożenie dla pewności prawa w zakresie dozwolonego użytku.

[64] Por. Barta; Markiewicz 2010, s. 136 i nast.

Art. 23. 1. Bez zezwolenia twórcy wolno nieodpłatnie korzystać z już rozpowszechnionego utworu w zakresie własnego użytku osobistego. Przepis ten nie upoważnia do budowania według cudzego utworu architektonicznego i architektoniczno-urbanistycznego oraz do korzystania z elektronicznych baz danych spełniających cechy utworu, chyba że dotyczy to własnego użytku naukowego niezwiązanego z celem zarobkowym.

2. Zakres własnego użytku osobistego obejmuje korzystanie z pojedynczych egzemplarzy utworów przez krąg osób pozostających w związku osobistym, w szczególności pokrewieństwa, powinowactwa lub stosunku towarzyskiego.

W odniesieniu do dozwolonego użytku osobistego, na który mogą powoływać się indywidualni użytkownicy (osoby fizyczne), należy zwrócić uwagę na następujące kwestie:

- użytek ten nie obejmuje 3 kategorii utworów: architektonicznych, elektronicznych baz danych oraz co najistotniejsze programów komputerowych (por. art. 77); obejmuje natomiast pozostałe kategorie utworów i przedmiotów praw pokrewnych zarówno w wersji cyfrowej, jak i analogowej;
- użytek obejmuje tylko korzystanie przez jedną osobę (indywidualnie) lub przez grupę osób (ale pod warunkiem, że pozostają one w związku osobistym); weryfikacja spełnienia tego warunku obecnie przy komunikacji w środowisku cyfrowym, jak i ustalenie wytycznych „ilościowo-jakościowych" dla interpretacji terminu „związek osobisty" rodzi wiele problemów, pytań na które odpowiedź nie niewykluczone, że będzie miała fundamentalne znaczenie dla przyszłości instytucji dozwolonego użytku osobistego;
- oczywistym jest, że przepis o dozwolonym użytku nie może być podstawą dla funkcjonowania bibliotek cyfrowych, ale użytkownicy w ramach dozwolonego użytku osobistego mogą wykorzystywać utwory zawarte w zbiorach bibliotek cyfrowych;
- niektóre biblioteki wprowadziły ograniczenia (ilościowe) w kserowaniu przez użytkowników utworów na podstawie instytucji dozwolonego użytku – zasadność takiego działania była przedmiotem oceny sądowej – na początku grudnia 2011 r. Sąd Ochrony Konkurencji i Konsumentów, w sprawie XVII Amc 113/11, orzekł, iż zawarta w regulaminie jednej z bibliotek klauzula ograniczająca kopiowanie do jednego arkusza wydawniczego ma charakter niedozwolony (jest tzw. klauzulą abuzywną [Wagłowski 2011a; 2011b]) i tym samym nie wiąże stron[65].

Pozostałe formy dozwolonego użytku (odnoszą sie tylko do użytku publicznego) utworów i przedmiotów praw pokrewnych obejmują m.in.:

- korzystanie przez instytucje naukowe i oświatowe (art. 27),
- cytat (art. 29),
- korzystanie przez ośrodki informacji i dokumentacji (art. 30).

[65] Instytucja klauz niedozwolonych jest szczegółowo uregulowana w art. 385^1 k.c., zgodnie z którym postanowienia umów zawieranych z konsumentem nie uzgodnione indywidualnie nie wiążą go, jeżeli kształtują jego prawa i obowiązki w sposób sprzeczny z dobrymi obyczajami, rażąco naruszają jego interesy. Pisemne uzasadnienie przykładowego orzeczenia, którego jeszcze nie udostępniono, będzie zawierało szczegółowe wyjaśnienie stanowiska sądu i argumenty które przeważyły za uznaniem tego ograniczenia za niedozwolone.

Zawsze korzystając z utworu na podstawie konkretnego przepisu o dozwolonym użytku należy sprawdzić czy są spełnione wszystkie przesłanki – warunki legalności działania; nie można tworzyć nowych „hybrydowych" form dozwolonego użytku, które zawierałyby część wymogów z jednej, a część z drugiej formy dozwolonego użytku (łączyć elementów – przesłanek warunkujących legalność korzystania z licencji z art. 27 i 28).

Najważniejsze znaczenie wśród przepisów o dozwolonym użytku, dla funkcjonowania bibliotek ma art. 28 pr. aut. Poniżej zostanie zaprezentowana pierwotna[66] (1994 r.) oraz obecnie obowiązująca (po nowelizacji z 2004 r.) wersja tego przepisu. Już wstępne porównanie treści przepisów pozwala na zajęcie stanowiska, iż w istotny sposób został poszerzony dla bibliotek zakres działań dozwolonych. Jest to jedyny przepis polskiego pr. aut., w którym, wprost uczyniono odwołanie do bibliotek.

Prawo autorskie od 1994 – 2004	Prawo autorskie od 2004 – obecnie
Art. 28: Biblioteki, archiwa i szkoły mogą:	Art. 28: Biblioteki, archiwa i szkoły mogą:
1) udostępniać nieodpłatnie, w zakresie swoich zadań statutowych, egzemplarze utworów opublikowanych,	1) udostępniać nieodpłatnie, w zakresie swoich zadań statutowych, egzemplarze utworów rozpowszechnionych,
2) sporządzać lub zlecać sporządzenie pojedynczych egzemplarzy utworów opublikowanych, niedostępnych w handlu – w celu uzupełniania, ochrony swoich zbiorów i nieodpłatnego ich udostępniania.	2) sporządzać lub zlecać sporządzanie egzemplarzy rozpowszechnionych utworów w celu uzupełnienia, zachowania lub ochrony własnych zbiorów,
	3) udostępniać zbiory dla celów badawczych lub poznawczych za pośrednictwem końcówek systemu informatycznego (terminali) znajdujących się na terenie tych jednostek.a

Przepis ten ma zastosowanie do tych podmiotów, które spełniają cechy biblioteki przewidziane w ustawie o bibliotekach, nie każda inicjatywa, grupa osób posługująca się nazwą „biblioteka" będzie uprawniona do wykorzystywania utworów z powołaniem się na ten przepis.

Art. 28 ma zastosowanie wszystkich kategorii utworów poza: bazami danych (art. 30^1 pr. aut.) i programami komputerowymi (art. 77 pr. aut.).

Na podstawie art. 28 pkt 1 pr. aut. biblioteki mogą udostępniać (użyczać) egzemplarze utworów rozpowszechnionych (a nie tylko opublikowanych jak miało to miejsce do 2004 r.), w porównaniu do pierwotnej wersji, zrezygnowano także z wymogu niedostępności w handlu oraz pojedynczych egzemplarzy[67]. W praktyce biblioteki głównie wypożyczają utwory opublikowane, których egzemplarze były wprowadzone do obrotu, nie mniej jednak coraz częściej zostają przekazywane do biblioteki pojedyncze egzemplarze utworów (które nie były w ogóle publikowane – niewydane ręko-

[66] Por. uwagi J. Marcinkowskiej, A. Matlaka w książce: Barta; Markiewicz 1999, s. 141-143.

[67] Odnośnie kwestii odpłatności za użyczanie szczegółowo por. Sokołowska, D., *Wynagrodzenie twórców z tytułu udostępniania utworów przez biblioteki*, „Przegląd Prawa Handlowego", luty 2007, nr s. 39; Jewuła, B., Stanisławska-Kloc, S., 2008; Kurowska, A., *Problem implementacji dyrektywy 2006/115/WE w zakresie public lending right*. „Zeszyty Naukowe Uniwersytetu Jagiellońskiego. Prace z Prawa Własności Intelektualnej" z. 3 (105), s. 60).

pisy, prace doktorskie [Barta; Markiewicz 2010, s. 175]; lub też zostały rozpowszechnione w innej formie – na przykład za pośrednictwem Internetu). Uprawnienie to bezpośrednio nie odgrywa istotnej roli w działalności bibliotek cyfrowych.

Uprawnienia przyznane bibliotekom na podstawie art. 28 pkt 2 i pkt 3 zostały bądź to zmodyfikowane (pkt 2) bądź wprowadzone (pkt 3) w związku z implementacją do prawa polskiego dyrektywy 2001/29/WE. Poniżej zostały przywołane[68] adekwatne do zakresu omawianej problematyki fragmenty tej dyrektywy:

Recital 40 Member States may provide for an exception or limitation for the benefit of certain non–profit making establishments, such as publicly accessible libraries and equivalent institutions, as well as archives. However, this should be limited to certain special cases covered by the reproduction right. Such an exception or limitation should not cover uses made in the context of on–line delivery of protected works or other subject–matter. This Directive should be without prejudice to the Member States' option to derogate from the exclusive public lending right in accordance with Article 5 of Directive 92/100/EEC. Therefore, specific contracts or licences should be promoted which, without creating imbalances, favour such establishments and the disseminative purposes they serve.	Preambuła pkt 40 Państwa Członkowskie mogą uwzględnić wyjątek lub ograniczenie na rzecz niektórych instytucji non-profit, takich jak ogólnodostępne biblioteki lub inne instytucje o podobnym charakterze, jak również archiwa. Jednakże powinno to być ograniczone do niektórych szczególnych przypadków objętych prawem do zwielokrotniania. Taki wyjątek lub ograniczenie nie powinien pokrywać korzystania w ramach dostarczania przez Internet utworów lub innych przedmiotów objętych ochroną. Niniejsza dyrektywa nie powinna naruszać możliwości Państw Członkowskich do odstąpienia od wyłącznego prawa do publicznego użyczania, zgodnie z art. 5 dyrektywy 92/100/EWG. Stosowne jest promowanie specjalnych umów lub licencji, bez wprowadzania nierówności, wspierających takie instytucje i rozpowszechnianie celów, które realizują.
Art. 5 /2/c Member States may provide for exceptions or limitations to the reproduction right provided for in Article 2 in the following cases: (...) – in respect of specific acts of reproduction made by publicly accessible libraries, educational establishments or museums, or by archives, which are not for direct or indirect economic or commercial advantage;	Art. 5 ust. 2 lit. „c" Państwa Członkowskie mogą przewidzieć wyjątki lub ograniczenia w odniesieniu do prawa do zwielokrotniania określonego w art. 2, w następujących przypadkach: (...) – w odniesieniu do szczególnych czynności zwielokrotniania dokonywanym przez ogólnodostępne biblioteki, instytucji edukacyjnej lub muzea, lub przez archiwa, które nie są skierowane na osiągnięcie bezpośredniej lub pośredniej korzyści gospodarczej lub handlowej;

[68] Polska wersja językowa (co do której można mieć wątpliwości zarówno redakcyjne, jak i merytoryczne) pochodzi z oficjalnego tłumaczenia, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:2001L0029:20010622:PL:PDF [odczyt: 12.12.2011].

Art 5/3/n 3. Member States may provide for exceptions or limitations to the rights provided for in Articles 2 and 3 in the following cases: (...) –use by communication or making available, for the purpose of research or private study, to individual members of the public by dedicated terminals on the premises of establishments referred to in paragraph 2(c) of works and other subject–matter not subject to purchase or licensing terms which are contained in their collections; (...)	Art. 5 ust. 3 lit. „n" 3. Państwa Członkowskie mogą przewidzieć wyjątki lub ograniczenia w odniesieniu do praw określonych w art. 2 i 3 w następujących przypadkach: (...) – korzystania poprzez udostępnienia lub podanie do wiadomości, w celu badań i prywatnych studiów, indywidualnym postronnym osobom z dzieł lub innych przedmiotów objętych ochroną, które nie podlegają zasadom zakupu lub licencji, znajdujących się w zbiorach instytucji określonych w ust. 2 lit. c), przy użyciu wyposażonych terminali w pomieszczeniach wymienionych instytucji; (...)
4. Where the Member States may provide for an exception or limitation to the right of reproduction pursuant to paragraphs 2 and 3, they may provide similarly for an exception or limitation to the right of distribution as referred to in Article 4 to the extent justified by the purpose of the authorised act of reproduction.	4. Jeżeli Państwa Członkowskie mogą przewidzieć wyjątek lub ograniczenie w odniesieniu do prawa do zwielokrotniania zgodnie z ust. 2 lub 3, mogą one również dopuścić wyjątek lub ograniczenie w odniesieniu do prawa rozpowszechniania określonego w art. 4, o ile ten wyjątek jest uzasadniony celem zezwolenia do zwielokrotniania.
5. The exceptions and limitations provided for in paragraphs1, 2, 3 and 4 shall only be applied in certain special cases which do not conflict with a normal exploitation of the work or other subject–matter and do not unreasonably prejudice the legitimate interests of the rightholder.	5. Wyjątki i ograniczenia przewidziane w ust. 1, 2, 3 i 4 powinny być stosowane tylko w niektórych szczególnych przypadkach, które nie naruszają normalnego wykorzystania dzieła lub innego przedmiotu objętego ochroną ani nie powodują nieuzasadnionej szkody dla uzasadnionych interesów podmiotów praw autorskich.

W odniesieniu do zwielokrotniania utworów polski ustawodawca przewidział legalność dokonywania tej czynności w ramach realizacji 3 celów: uzupełnienia, zachowania, lub ochrony własnych zbiorów. Na podstawie art. 28 pkt 2 dopuszczalna jest cyfryzacja zbiorów bibliotecznych[69].

Na podstawie art. 28 pkt 2 może być dokonywana każda[70] forma zwielokrotnienia, np. przy zastosowaniu techniki drukarskiej, reprograficznej, zapisu magnetycznego (por. art. 50 pr. aut.); dozwolone jest sporządzenie fotokopii, mikrofilmu, kopii cyfrowej np. na dyskietce, płycie CD.

[69] J. Barta i R. Markiewicz opowiadają się w tym zakresie za wykładnią tego przepisu pozwalającą w szerokim stopniu na digitalizacje zasobów [Barta; Markiewicz 2010, s. 178]; szczegółowe omówienie interpretacji przesłanek zwielokrotnienia dla celów uzupełnienia, zachowania, ochrony zbiorów zostało dokonane w opracowaniu: Jewuła; Stanisławska-Kloc 2008. Por. też np. dyskusję, która toczy się obecnie w Wielkiej Brytanii: dokument Consulation on Copyright, listopad 2011, s. 70 [dok. elektr.] http://www.ipo.gov.uk/consult-2011-copyright.pdf [odczyt: 2.01.2012].

[70] Por. też Poźniak-Niedzielska, Maria; Szczotka, Jerzy; Mozgawa, Marek (2006). *Prawo autorskie i prawa pokrewne. Zarys wykładu.* Red. Maria Poźniak-Niedzielska. Bydgoszcz; Warszawa; Lublin: Wydaw. Branta, s.101; Gienias, Krzysztof (2011). *Treść prawa autorskiego.* [W:] Ferenc-Szydełko, Ewa red. Tomczyk, Sławomir; Szyjewska-Bagińska, Joanna; Drzewiecki, Andrzej; Krzysztof Gienas aut. (2011). *Ustawa o prawie autorskim i prawach pokrewnych: komentarz.* Warszawa: C.H. Beck, s. 223-225.

Z kolei na podstawie art. 28 pkt 3 pr. aut. biblioteki mogą udostępniać swoje zbiory (czyli udostępniać znajdujące się w nich utwory) dla celów badawczych lub poznawczych za pośrednictwem terminali znajdujących się na terenie tych instytucji (w intranecie).

Termin „udostępnianie" nie został zdefiniowany w ust. o pr. aut. W kontekście art. 6 pkt 3 pr. aut., który stanowi, iż „utworem rozpowszechnionym jest utwór, który za zezwoleniem twórcy został w jakikolwiek sposób udostępniony publicznie" oraz mając na względzie treść art. 50 pkt 3 pr. aut. można przyjąć, iż biblioteki są upoważnione do udostępniania czytelnikom utworów w każdy sposób, który odbywa się poprzez (za pośrednictwem) komputery usytuowane na terenie tych jednostek.

Omawiany przepis nie stanowi podstawy legalizacji udostępniania utworów (do których nie wygasły autorskie prawa majątkowe) w Internecie (w ogólnodostępnych sieciach komputerowych). Zatem na gruncie prawa polskiego jedynie umowa z podmiotem uprawnionym (z twórcą) jest źródłem legalnego udostępniania w Internecie utworów (których czas ochrony autorskich praw majątkowych jeszcze nie upłynął) przez biblioteki cyfrowe.

Literalne brzmienie art. 28 pkt 3 pr. aut. jest wzorowane na treści art. 5 ust. 3 lit. „n" dyrektywy 2001/29/WE, stąd też sformułowanie odwołujące się do terminu „teren jednostek". Powinno ono być interpretowane celowościowo w ten sposób, aby [Jewuła; Stanisławska-Kloc 2008]:

- objąć pomieszczenia należące do biblioteki tj. będące własnością biblioteki lub wykorzystywane na podstawie umowy najmu; dotyczy to tych bibliotek, które mają podmiotowość prawną;
- wydzielone, przeznaczone na działalność biblioteczną pomieszczenia[71]; dotyczy to tych bibliotek, które nie mają podmiotowości prawnej, stanowią jednostki organizacyjne (ale nie są samodzielnymi jednostkami – por. art. 11 pr. o bibl.) innych podmiotów prawa nie będących beneficjentami wyjątku z art. 28 pr. aut., np. biblioteki zakładowe;
- pomieszczenia przeznaczone na działalność biblioteczną oraz pomieszczenia należące do innego, aniżeli biblioteka podmiotu prawa będącego beneficjentem wyjątku z art. 28 pr. aut. np. szkoły podstawowej, szkoły ponadpodstawowej oraz szkoły wyższej, w których działają biblioteki (np. biblioteki będące jednostkami organizacyjnymi szkół).

Polski ustawodawca nie wprowadził ograniczenia tej licencji tylko do utworów „które nie podlegają zasadom zakupu lub licencji"– art. 5 ust. 3 „n"[72].

11. Umowy dotyczące autorskich praw majątkowych

Umowy z punktu widzenia prawa cywilnego to tzw. dwustronne czynności prawne. Dla skutecznego zawarcia umowy wymagane jest złożenie oświadczenia woli prze

[71] Stan posiadania takich pomieszczeń może ulec zmianie w związku ze zmianami organizacyjnymi podmiotu, w ramach którego biblioteka funkcjonuje.

[72] Wątpliwość co do zgodności prawa polskiego z unijnym zgłaszali w tym zakresie Barta, Markiewicz 2010, s. 395.

dwie strony (autora/uprawnionego z tytułu praw autorskich/licencjodawcę oraz nabywcę/wydawcę/licencjobiorcę). Bardzo często autor, który dokonuje samo archiwizacji swojego utworu w ramach repozytorium cyfrowego udziela bibliotece prowadzącej to repozytorium licencji na korzystanie z jego utworu. Także student, który „przekazuje swoja pracę magisterską" do uczelnianej bazy prac dyplomowych udziela uczelni licencji na jej wykorzystywanie[73].

Do umów z zakresu praw autorskich mają zastosowanie przepisy kodeksu cywilnego, a w szczególności księgi pierwszej „Część ogólna" oraz księgi trzeciej „Zobowiązania".

Do zawarcia umów dot. praw autorskich co do zasady wymagana jest pełna zdolność do czynności prawnych, którą posiada każda osoba fizyczna mająca ukończone 18 lat (i nie będąca ubezwłasnowolnioną). Twórca może zawrzeć umowę samodzielnie lub ustanowić pełnomocnika, który w jego imieniu zawrze umowę. Szczegółowe uregulowanie dotyczące umów z zakresu autorskich praw majątkowych zawierają przepisy art. 41-68 pr. aut.

Wyróżnia się zasadniczo dwa rodzaje umów dotyczących autorskich praw majątkowych: **umowy przenoszące** prawa oraz **umowy na korzystanie tzw. umowy licencyjne**. Te pierwsze wywołują skutek taki jak umowy sprzedaży (dochodzi do zmiany podmiotu prawa, nabywca niejako „wstępuje" w miejsce autora). Natomiast na podstawie umów drugiego typu osoba trzecia (nazywana licencjobiorcą) uzyskuje od twórcy zgodę (swoiste upoważnienie) do korzystania z utworu (przez okres czasu oraz w sposób – tj. na polach eksploatacji określonych w umowie). Wyróżnia się dwa rodzaje umów licencyjnych:

- wyłączne (w których licencjodawca udziela zgody na korzystanie tylko jednej osobie; zapewnia licencjobiorcy swoistą wyłączność na korzystanie z utworu)
- niewyłączne (takich umów licencjodawca może zawrzeć nieograniczoną liczbę, wtedy wielu licencjobiorców korzysta „równolegle" z utworu).

W razie rozbieżności pomiędzy treścią umowy (np. udzielona licencja) a tytułem (umowa przenosząca prawo) o zakresie przekazanych praw rozstrzyga treść umowy a nie jej tytuł. Zwyczajowo strony, głównie w celach ułatwienia obrotu i archiwizowania umów nadają im nazwy oraz numery porządkowe.

Dla umów przenoszących autorskie prawa majątkowe przewidziany jest wymóg formy pisemnej, pod rygorem nieważności. Oznacza to, że umowa przenosząca prawo zawarta w innej aniżeli pisemna formie będzie od początku nieważna. Forma pisemna jest zachowana w przypadku złożenia przez strony własnoręcznego podpisu pod treścią umowy (na dokumencie zawierającym spisaną treść umowy – umowa może być przygotowana w komputerze i wydrukowana, nie musi być spisana odręcznie). Także posłużenie się formą bezpiecznego podpisu elektronicznego jest dopuszczalne i skuteczne.

Umowy licencyjne inaczej zwane umowami na korzystanie z utworu, mogą być zawarte w każdej formie np. poprzez wymianę emaili, ustnie. Jednak dla celów dowodowych i archiwalnych (aby wykazać fakt nabycia praw – ciężar w tym zakresie spoczywa

[73] Por. pkt 2-4 oświadczenia stanowiące załącznik do zarządzenia Rektora UJ nr 39 z dnia 11 kwietnia 2011 r. w sprawie elektronicznego archiwizowania prac dyplomowych (licencjackich i magisterskich), http://www.bip.uj.edu.pl/dokumenty/zarzadzenia-rektora [odczyt: 12.12.2011].

na korzystającym) zalecane jest zawieranie umów w wersji pisemnej, archiwizowanie korespondencji, w której autor wyraził zgodą na wykorzystanie jego utworu. Forma pisemna umowy ogranicza ryzyko wystąpienia wątpliwości interpretacyjnych (które mogą wystąpić w przypadku, gdy strony umowy zawartej ustnie, po pewnym czasie nie będą pamiętać szczegółów tej umowy w sposób jednolity.) Także dokonywanie zmian, uzupełnień czy aneksów do umowy powinno odbywać się w formie pisemnej.

Umowa dotycząca praw autorskich może być zawarta w każdym języku. Prawo polskie nie wprowadza obowiązku zawarcia umowy tylko i wyłącznie w języku polskim. Jeżeli umowa będzie wykonywana w Polsce zalecane jest :

- sporządzenie jej w wersji bilateralnej (tj. dwujęzycznej np. polsko-angielskiej, ze wskazaniem która z wersji w razie wątpliwości interpretacyjnych będzie miała decydujące znaczenie),
- dokonanie tłumaczenia wersji obcojęzycznej na język polski.

Dla umów z zakresu prawa autorskiego oraz praw pokrewnych, prawa producenta bazy danych nie istnieje w Polsce oficjalny ich rejestr, stąd nie ma możliwości sprawdzenia czy i kiedy, które strony zawarły umowę. Należy podkreślić, że samo zawarcie umowy o dzieło czy umowy zlecenia nie skutkuje nabyciem autorskich praw majątkowych. Z faktu zawarcia jednej z wymienionych umów nie można domniemywać przeniesienia praw autorskich. Niekiedy zasadnie próbuje się, stosując instytucję wykładni oświadczeń woli (badając cel umowy i zgodny zamiar stron, charakter umowy) wyinterpretowywać, zawarcie dorozumianej licencji niewyłącznej na korzystanie z utworu. Niestety rezultat takiego „zabiegu", dokonywanego najczęściej już na etapie procesu sądowego (np. wytoczonego przez autora, który uważa, że nie udzielił zgody na korzystanie z jego utworu), jest niepewny; a udowodnienie uzyskania takiej licencji dosyć czasochłonne i trudne. Dlatego należy zawsze pamiętać o wprowadzeniu do umowy o dzieło (lub innej na podstawie której dochodzi do stworzenia utworu) postanowień dotyczących praw autorskich.

Na potrzeby wykorzystania utworu w bibliotece cyfrowej, wystarczające będzie zawarcie umowy licencyjnej, niewyłącznej, w zakresie takich pól eksploatacji jak: utrwalenie, zwielokrotnienie, publiczne udostępnienie – rozpowszechnienie. Zasadniczo, wydaje się, że przy typowym wykorzystywaniu utworów w bibliotece cyfrowej nie dochodzi do ich opracowywania, tłumaczenia, lub dokonywania innych jeszcze modyfikacji, które wymagałyby nabycia od twórcy praw do opracowania (art. 2 pr. aut.).

Poniżej zostaną wymienione podstawowe elementy umowy (licencyjnej lub przenoszącej prawa autorskie), na podstawie której utwór będzie inkorporowany i następnie udostępniany w ramach biblioteki cyfrowej:

- data i miejsce zawarcia umowy,
- oznaczenie stron umowy, które pozwoli na ich identyfikację (uprawniony – nabywca, licencjodawca – licencjobiorca i szczegółowe dane identyfikujące oraz kontaktowe),
- oznaczenie, opisanie utworu (-ów), których dotyczy umowa (tytuł, objętość, dodatkowo może zostać wskazane miejsce pierwotnej publikacji; jedna umowa może dotyczyć kilku utworów),

- oświadczenie o przysługiwaniu praw autorskich (z tytułu stworzenia utworu, z tytułu dziedziczenia praw autorskich zmarłego twórcy, w tym drugim przypadku celowe jest podanie rodzaju i oznaczenia dokumentu, który poświadczy fakt dziedziczenia),
- oświadczenie o udzieleniu licencji/wyrażeniu zgody na korzystanie z utworu lub o przeniesieniu praw; jeżeli wprost nie wynika, z oświadczenia, iż doszło do przeniesienia praw to domniemywa się, że została udzielona licencja;
- postanowienie o zakresie udzielonych praw, zawierające wymienienie pól eksploatacji, przy czym najlepiej, aby zostały one opisane przy użyciu terminologii z art. 50 pr. aut.,
- oświadczenie o nieodpłatnym udzieleniu praw, a jeżeli udzielenie praw odbywa się za wynagrodzeniem wskazanie tego wynagrodzenia lub zasad jego obliczania (ryczałtowe, procentowe) i terminów wypłaty,
- oświadczenie o czasie trwania umowy licencyjnej i terytorium (np. licencja udzielona na cały okres ochrony utworu, na terytorium Polski oraz za granicą, we wszystkich krajach świata).

Przykładowe postanowienie umowne dotyczące pól eksploatacji utworu, wykorzystywanego w bibliotece cyfrowej:

Licencja wyłączna/niewyłączna, o której mowa w pkt. X powyżej obejmuje wyłączne prawo Licencjobiorcy do korzystania z Utworu w całości lub we fragmentach, przez cały okres ochrony autorskich praw majątkowych do Utworu w kraju i za granicą, na następujących polach eksploatacji:

- utrwalania Utworu w wersji drukowanej oraz pamięci komputerów, w tym spełniających funkcje serwerów,
- zwielokrotniania Utworu bez żadnych ograniczeń ilościowych, techniką: drukarską, zapisu magnetycznego, reprograficzną oraz cyfrową, w pamięci komputera jak i w sieciach multimedialnych, w internecie, intranecie, w szczególności online, a także poprzez wydruk komputerowy, na każdym znanym w dacie podpisania niniejszej umowy nośniku,
- udostępniania Utworu, w tym także przesyłanie za pośrednictwem sieci multimedialnych, w szczególności internetu i intranetu, online, w ramach komunikacji na życzenie (on-demand), w tym również publiczne udostępnianie w taki sposób, aby każdy mógł mieć do utworu (lub jego fragmentu) dostęp w miejscu i w czasie przez siebie wybranym.

12. Biblioteki cyfrowe – etapy tworzenia i udostępniania zbiorów z punktu widzenia prawa autorskiego

W pewnym uproszczeniu, proces tworzenia biblioteki cyfrowej składa się z dwóch etapów:

1) zbierania, przygotowania zasobów, który wiąże się z utrwalaniem, zwielokrotnianiem utworów składających się na zbiory:
- jeżeli utwór (np. książka) jest w wersji analogowej, tradycyjnej to niezbędna będzie jego digitalizacja, cyfryzacja, aby mógł zaistnieć w cyfrowym świecie – biblioteki cyfrowej, w jej zbiorach;
- jeżeli utwór istniał od początku w wersji cyfrowej (tzw. digitaly born) to jego cyfrowe wersje powinny zostać utrwalone (zapisane) na serwerze, z którego będzie dochodziło do rozpowszechniania – udostępniania ograniczonego lub nieograniczonego terytorialnie i podmiotowo; takie utrwalenie stanowi odrębne od zwielokrotnienia pole eksploatacji utworu,

2) udostępniania publicznego zasobów (zbiorów), czyli rozpowszechniania poszczególnych utworów składających się na ten zbiór.

Wymienione w pkt 1 i 2 czynności, dopóki trwają autorskie prawa majątkowe objęte są monopolem twórcy.

W przypadku pierwszego etapu istotną rolę odgrywa digitalizacja, a potem utrwalenie (przechowywanie) cyfrowej wersji na serwerze. Digitalizacja stanowi:
- formę zwielokrotnienia utworu objętą monopolem autorskoprawnym,
- dokonanie zapisu w formie cyfrowej, jest tylko i wyłącznie czynnością techniczną.

Digitalizacja nie stanowi:
- opracowania w rozumieniu prawa autorskiego, nie skutkuje powstaniem po stronie osoby dokonującej digitalizacji tzw. praw zależnych (art. 2 pr. aut.).

Digitalizacja nie wpływa na przedłużenie czasu trwania autorskich praw majątkowych do utworu (wersja tradycyjna, jak i cyfrowa tego samego utworu są chronione przez ten sam okres, czyli zasadniczo przez okres życia twórcy i 70 lat po jego śmierci). Nie można jednak wykluczyć, iż po zdigitalizowaniu nowa (z technicznego punktu widzenia) kopia utworu zostanie np. przetłumaczona (na inny język), wzbogacona o dodatkowe materiały (inne utwory lub ich fragmenty – np. ilustracje) i wtedy możemy mieć do czynienia z powstaniem opracowania (z tzw. dziełem zależnym – art. 2 pr. aut.)[74].

W omawianym kontekście pojawiają się dwa istotne zagadnienia dotyczące tego czy, można robić kopie cyfrowe całości zbiorów analogowych oraz tego jaki jest status (losy) wersji tradycyjnej po stworzeniu wersji cyfrowej. Uważam, że brak w prawie polskim podstaw dla legalizacji działań polegających na całościowej jednorazowej retrodigitalizacji zbiorów analogowych. Natomiast stopniowo, realizując cel jakim jest zachowanie oraz ochrona zbiorów biblioteki mogą na podstawie art. 28 pkt 2 dokonywać digitalizacji zasobów i następnie udostępniać albo na zasadach indywidualnie ustalonych (w umowie) z autorem lub wydawcą albo na zasadach dozwolonych w ramach art. 28 pkt 3 pr. aut. Eydaje się, że wersja tradycyjna może podlegać wykorzystaniu na dotychczasowych warunkach (np. użyczenie).

[74] Chyba tylko wyjątkowo można by było jeszcze rozważać zaistnienie utworów połączonych do rozpowszechniania (art. 10 pr. aut.).

Drugi aspekt funkcjonowania biblioteki cyfrowej związany jest z publicznym udostępnianiem jej zbiorów – czyli z rozpowszechnianiem tych zbiorów w rozumieniu pr. aut. Rozpowszechnianie stanowi:

- publiczne udostępnienie utworu, które może odbywać się na wiele sposobów, za pośrednictwem odrębnych pól eksploatacji w rozumieniu art. 50 pr. aut.,
- udostępnianie publiczne utworu (udostępnianie utworu w czasie i miejscu wybranym przez użytkownika),
- udostępnianie cyfrowej wersji utworu w internecie (wirtualizacja) to forma udostępniania publicznego, rozpowszechniania utworu (por. art. 6 pkt 3 pr. aut.),
- udostępnianie utworu w formie prezentacji multimedialnej na stronie www.

13. Znaczenie wolnych licencji dla funkcjonowania bibliotek cyfrowych

Zagadnieniem, które wymaga zasygnalizowania jest wykorzystywanie tzw. wolnych licencji (np. CC) przy tworzeniu i funkcjonowaniu bibliotecznych kolekcji cyfrowych[75]. Precyzyjnie mówiąc, chodzi o wykorzystanie utworów rozpowszechnianych (udostępnionych) na licencjach CC w bibliotekach cyfrowych. Na licencjach CC mogą być rozpowszechniane te utwory, które korzystają z ochrony (nie wygasły w stosunku do nich autorskie prawa majątkowe).

Licencje Creative Commons[76] (CC) są jednymi z najpopularniejszych wolnych licencji; używają ich zwolennicy szerokiego udostępniania utworów (dzielenia się własną twórczością z innymi), zapewnienia sobie tylko „pewnych" praw autorskich. Autor (lub jego spadkobiercy) udostępniając utwór na licencji CC (czyli „opatrując" utwór w taką licencję) i zastrzegając dla siebie tylko „pewne prawa", godzi się na to, że każdy przy zachowaniu wymogów określonych w tej licencji będzie mógł wykorzystywać utwór (kopiować, poprawiać, udoskonalać). Podstawowe formy licencji CC i ich oznaczanie to:

a) uznanie autorstwa (BY:) (*attribution*), na podstawie której można korzystać z utworu pod warunkiem, że zostanie przywołane nazwisko autora;

[75] Zob. Kapellakou, Galateia; Markellou, Marina; Vagena, Evangelia (2011). *Open content in Libraries: Contracutal Issues.* [W:] Iglezakis, Ioannis; Synodinou, Tatiana-Eleni; Kapidakis, Sarantos red. (2011). *E-Publishing and Digital Libraries : Legal and Organizational Issues.* New York: Hershey, s. 342-361 oraz Iglezakis, Ioannis (2011). *Personal Data Protection in Digital Libraries.* [W:] Iglezakis, Ioannis; Synodinou, Tatiana-Eleni; Kapidakis, Sarantos red. (2011). *E-Publishing and Digital Libraries : Legal and Organizational Issues.* New York: Hershey, s. 413-429; Szpringer, Włodzimierz (2011). *Koncepcja open access w świetle ekonomicznej analizy praw własności intelektualnej.* Warszawa: Wolters Kluwer Polska Sp. z o.o., 293 s.

[76] Por. Barta; Markiewicz 2010, s. 170 i nast.; Machała, Wojciech (2009). *Wybrane cywilnoprawne aspekty licencji creative commons.* „Monitor Prawniczy" nr 8, s. 422 i nast.; Jastrzębska, Anna (2010). *Organizacje zbiorowego zarządzania prawami autorskimi a Creative Commons.* „Zeszyty Naukowe Uniwersytetu Jagiellońskiego. Prace z Prawa Własności Intelektualnej" nr 2 (108), s. 20-33; Wasilewski, Piotr (2008). *Open kontent: zagadnienia prawne.* Warszawa: Wolters Kluwer, 244 s. Szczegółowo por. http://creativecommons.pl/

b) użycie niekomercyjne  (*noncommercial*), na podstawie której można korzystać z utworu jedynie dla celów niekomercyjnych;
c) użycie bez utworów zależnych  (*no derivative works*), na podstawie której można korzystać z utworu jedynie w jego oryginalnej postaci – tworzenie utworów zależnych nie jest dozwolone;
d) użycie na tych samych warunkach  (*share alike*), na podstawie której można korzystać z utworu zależnego jedynie na licencji identycznej do tej, na której udostępniono utwór w wersji pierwotnej (oryginalny).

Przy czym zawsze musi być spełniony warunek oznaczenia autorstwa, przy każdej wersji licencji CC.

Najdalej idącą, nową licencją CC, która ma wywoływać „skutek"[77] taki, jak gdyby utwór był w domenie publicznej, a co za tym idzie, tak jakby autor zrzekł się niemal wszelkich możliwych praw jest licencja CC0 („Creative Commons 0", „CC zero"). Przy czym, specjalne nawiązanie w jej nazwie do zera, tłumaczy się tym, że utwory udostępnianie na tej licencji są usytuowanie na „osi uprawnień"[78] w punkcie zerowym (autor nie ma żadnych uprawnień „no copyright" albo „no copyrights reserved"). Potem w dalszej kolejności na tej osi (może w środku) znalazłyby się utwory udostępnianie na pozostałych licencjach CC (które to licencje pozostawiają po stronie twórcy pewne prawa – stąd formuła „some rights reserved"). Najdalej jako przeciwwaga punktu zero – można umownie określić go punktem „10" (niejako na końcu) byłyby utwory w pełni chronione prawem autorskim (których twórcy posługują się formułą „all rights reserved"), udostępniane na podstawie „tradycyjnych" umów negocjowanych pomiędzy uprawnionym z tytułu praw autorskich a użytkownikiem.

Biblioteki cyfrowe mogą zawierać (umieszczać w swoich zbiorach) utwory udostępniane na licencjach CC. Należy mieć na uwadze, iż włączenie utworu do kolekcji biblioteki cyfrowej nie może skutkować pozbawieniem go statusu dzieła udostępnianego na warunkach licencji CC (a także nie może odbywać się z pominięciem oznaczeń CC). Jak wydaje się nie ma potrzeby wydzielania specjalnego zbioru utworów udostępnianych na licencji CC, utwory te mogą być i katalogowane i udostępniane wspólnie zarówno z utworami należącymi do domeny publicznej, jak i podlegającymi „pełnym" rygorom ochrony autorskoprawnej. Niektóre biblioteki promują rozpowszechnianie utworów na licencji CC0. Na gruncie prawa polskiego mając na względzie wyraźny zakaz zrzekania się autorskich praw osobistych (które to prawa, gdy trwają i tak nie ograniczają zasadniczo możliwości wykorzystywania utworu w bibliotece cyfrowej); wątpliwości co do przyjęcia możliwości zrzeczenia[79] się wszelkich praw

[77] Przy założeniu, iż dany system prawny zezwala na zrzeczenie się praw autorskich; na gruncie prawa polskiego jest to nieskuteczne (por. art. 16 i 18 pr. aut.).

[78] Por. Siewicz, Krzysztof (2011). *Analiza prawna „Creative Commons 0"* [dok. elektr.] http://koed.org.pl/2011/10/creative-commons-zero-w-polskim-prawie/ [odczyt: 12.01.2012]. Por też: Informacja z dnia 22 stycznia 2008 r. *Nowa licencja Cre-ative Commons?* [dok. elektr.] http://pl.wikinews.org/wiki/2008-01-22:_Nowa_licencja_Creative_Commons%3F; http://creativecommons.pl/2008/01/cc-zero/ [odczyt: 12.01.2012].

[79] Praw tych można nie wykonywać, lub zobowiązać się do ich niewykonywania w stosunku do określonych osób.

majątkowych ze skutkiem w postaci wygaśnięcia tych praw, i „zaliczenia" utworu do domeny publicznej – „wystarczające" jest posługiwanie się licencją „CC Uznanie autorstwa", natomiast nie wydaje się niezbędne odwoływanie do licencji CC0.

14. Problem utworów osieroconych

Dla funkcjonowania bibliotek cyfrowych, a w szczególności udostępniania ich kolekcji w wersji cyfrowej istotne znaczenie ma rozwiązanie problemu utworów[80] „osieroconych"[81]. Do tej kategorii zaliczane są utwory, z których twórcami (podmiotami uprawnionymi) nie można się skontaktować (twórcy ci są nieznani – nie można ustalić ich tożsamości lub znając tożsamość nie można nawiązać z nimi kontaktu). Problem z utworami osieroconymi polega na tym, że nie mogąc skontaktować się z ich autorem :

- nie można ustalić czy żyje on czy też zmarł (oraz czy upłynęło już 70 lat od śmierci, które skutkowałoby wygaśnięciem ochrony autorskich praw majątkowych);
- nie można uzyskać zgody od autora na korzystanie z utworu, jeśli jest on nadal chroniony prawem autorskim.

Nie bez powodu wskazuje się, iż to bibliotekarze ujawnili na szeroką skalę problem utworów osieroconych[82]. Doszło do tego w związku z digitalizacją zbiorów – które przy tej okazji – zostały poddane próbie ustalenia statusu autorskoprawnego. W praktyce chodzi o ustalenie, które utwory znajdują się już w domenie publicznej – i mogą podlegać swobodnemu wykorzystaniu przez kolejne pokolenia. Zakwalifikowanie utworu do domeny publicznej (w związku z upływem czasu trwania autorskich praw majątkowych) ma podstawowe znaczenie dla funkcjonowania bibliotek cyfrowych oraz reguł postępowania w zakresie wyznaczenia działań zmierzających do realizowania zasady zachowania dziedzictwa kulturowego dla przyszłych pokoleń.

Rozstrzygając problem korzystania z utworów osieroconych, z jednej strony widzę potrzebę zapewnienia uprawnień do korzystania z utworów, które znajdują się w zbiorach instytucji naukowych, kulturalnych (biblioteki, muzea, archiwa, jednostki naukowe). Instytucji niewątpliwie działających na rzecz rozwoju naukowego i artystycz-

[80] Zauważmy, że w literaturze przedmiotu głównie autorzy posługują się terminem utwory „osierocone" to status „osieroconych" można nadać także przedmiotom praw pokrewnych. – dotyczy to fonogramów, wideogramów, nietwórczych zdjęć (chronionych w niektórych systemach w ramach praw pokrewnych).

[81] Por. Barta; Markiewicz 2010, s. 82-84; Vetulani, Agnieszka (2009). *Dzieła osierocone – w poszukiwaniu europejskiego rozwiązania.* „Zeszyty Naukowe Uniwersytetu Jagiellońskiego. Prace Instytutu Prawa Własności Intelektualnej" z. 103, s. 26-69; Stanisławska-Kloc, Sybilla (2007). *Utwory osierocone.* Prawo własności intelektualnej – wczoraj, dziś i jutro. „Zeszyty Naukowe Uniwersytetu Jagiellońskiego. Prace z Prawa Własności Intelektualnej" z. 100, s. 453-476.

[82] Por. Szczepańska, Barbara (2010). *Biblioteki cyfrowe w europejskich dyskusjach o prawie autorskim.* [W:] *Seminarium prawne dla bibliotekarzy cyfrowych.* Warszawa, s. 75 i nast.; Papadopoulou, Maria-Daphne (2011). *The issue of Orphan 'Works' in Digital Libraries.* [W:] Iglezakis, Ioannis; Synodinou, Tatiana–Eleni; Kapidakis, Sarandos red. (2011). *E-Publishing and Digital Libraries: Legal and Organizational Issues.* Hershey; New York: Information Science Reference, s. 198-231.

nego kolejnych pokoleń. Z drugiej strony zdaję sobie sprawę, iż problem korzystania z utworów osieroconych dotyczy także innych środowisk; wydawcy także często stają przed nie lada wyzwaniem/ryzykiem – w związku ze wznowieniami tradycyjnych, jak i cyfrowych wydań tych utworów, z których część ma status utworów osieroconych. Niekiedy bez tych ryzykownych decyzji nie bylibyśmy w stanie obcować z wieloma utworami o znacznej wartości artystycznej, naukowej, wielce wzbogacającymi nasz dorobek[83].

Jak najbardziej dostrzegam konieczność ochrony interesów – trzeciej grupy podmiotów autorów (oraz ich spadkobierców), których nie można „wywłaszczyć" z ich praw tylko dlatego, że ich nie wykonują. Natomiast całkowity brak regulacji w zakresie utworów osieroconych może przyczynić się do zniknięcia (i to bynajmniej nie zawsze tylko czasowego, „chwilowego") części utworów (często zbyt duże ryzyko korzystania z utworów o nieustalonym statusie prawnym może skutkować ich pomijaniem, a w konsekwencji całkowitym zapomnieniem i zniszczeniem). A przecież prawo autorskie ma na celu chronić nie tylko majątkowe – mające bezpośrednie finansowe przełożenie interesy autora, ale i wieczne, „duchowe" interesy (realizowane poprzez system praw autorskich osobistych), do których nie dojdzie, gdy utwór zostanie zaliczony do kategorii „nietykalnych" utworów osieroconych.

Funkcjonujące w obecnym kształcie w Polsce (art. 33 pkt 3 pr. aut.) rozwiązanie dotyczące korzystania z utworów, z których twórcą nie można się skontaktować jest niewystarczające, zarówno z punktu widzenia przedmiotowego, jak i podmiotowego. Natomiast reguły odpowiedzialności cywilnej i karnej z tytułu korzystania z cudzego utworu bez zezwolenia są stosunkowo surowe.

Sygnalizowałam, iż problem utworów osieroconych był przedmiotem analizy twórców raportu Sages. Wydaje się celowe, przed omówieniem założeń projektu dyrektywy, przedstawienie głównych postulatów wysuwanych przez twórców raportu Sages, pod kierunkiem przyszłych regulacji prawnych dotyczących zasad korzystania z utworów osieroconych. Twórcy raportu zaproponowali ośmiostopniowy test, któremu według nich powinna zostać poddana propozycja prawna (akt prawny) zawierający rozwiązanie problemu utworów osieroconych – spełnienie tych warunków zawartych w teście zapewni skuteczność proponowanego rozwiązania[84]. Są to następujące warunki:

1) rozwiązanie powinno zostać wdrożone we wszystkich krajach UE (zatem powinno być przyjęte w formie dyrektywy lub rozporządzenia, obecnie w UE wybrano pierwszą opcję),

[83] Gdyby nie „odważna" decyzja jednego z polskich wydawców do polskich czytelników nie trafiłaby np. polska edycja biografii Honore Balzaka autorstwa Stefana Zweiga, z posłowiem i opracowaniem Richarda Friedenthala, która to w zakresie tłumaczenia autorstwa Wandy Kragen, jest utworem osieroconym. Przy czym wydaje się, że mimo możliwości dokonania tłumaczenia przez inną osobę (koszty mogą być porównywalne) to względy artystyczne, uznanie dla (dorobku) tłumaczki mogły spowodować, że zdecydowano się na taki krok. Co ciekawe wydawca ten nie ukrywał faktu, wykorzystania utworu w pewnym zakresie (tj. tłumaczeniu) osieroconego.

[84] W raporcie poruszono także kwestię korzystania z utworów o wyczerpanym nakładzie, rekomendując następujące rozwiązania: pierwszeństwo wykorzystania przez podmioty uprawnione (które decydują czy są zainteresowane rozpowszechnianiem utworów o wyczerpanym nakładzie), dopiero w drugiej kolejności należy rozważyć zapewnienie instytucjom kultury możliwości korzystania z tych utworów.

2) rozwiązanie powinno obejmować wszystkie kategorie utworów (a nie tylko wybrane jak postąpiono w projekcie dyrektywy),
3) należy zapewnić transgraniczność osieroconego charakteru utworu (uznanie utworu za osierocony w jednym państwie, skutkuje nadaniem mu takiego statusu we wszystkich krajach UE, przy czym zasadnie zwrócono uwagę, aby nie odwoływać się do historycznych granic, a uwzględniać współczesne granice państwowe),
4) zapewnić transgraniczną skuteczność takiego uznania, rozumianą w ten sposób, iż utwór osierocony udostępniony w Internecie w jednym państwie powinien być udostępniony w Internecie we wszystkich państwach (a nawet na całym świecie); mam wrażenie, iż ten warunek ma charakter „techniczny" i organizacyjny (należałoby koordynować bazy danych utworów osieroconych, może zastanowić się nad wprowadzeniem specjalnego oznaczenia – symbolu dla tej kategorii utworów),
5) rozwiązanie powinno być zgodne z zasadami wdrażania idei partnerstwa publiczno-prywatnego (ppp) w dziedzinie digitalizacji (z uwagi na zróżnicowane zasady ppp, raczej odbieram ten warunek jako promowanie idei ppp w projektach digitalizacji i przeciwdziałanie nowemu zawłaszczaniu cyfrowej wersji „starych" utworów),
6) rozwiązanie powinno przewidywać w przypadku komercyjnego korzystania (nawet przez instytucje kultury, nauki) z utworów osieroconych, wynagrodzenie dla podmiotów prawa autorskiego (na wypadek gdy się odnajdą); zasugerowano, iż takie wynagrodzenie może być przechowywanie na rachunku depozytowym,
7) rozwiązanie powinno zabezpieczać uzasadnione koszty „transakcyjne" z tytułu obsługi utworów osieroconych; rozumiane w ten sposób, iż poszukiwania podmiotów uprawnionych do utworów starszych mogą być mniej intensywne aniżeli nowszych; w praktyce oznacza to zgodę na zróżnicowanie, ale wydaje się, że dopuszczalne (uzasadnione), standardów oceny spełnienia warunku starannych poszukiwań,
8) rozwiązanie prawne powinno być wspierane poprzez powołanie bazy danych zawierającej informacje o prawach (podmiotach prawa), zasugerowano, aby baza taka była połączona z Europeaną.

Obecnie trwają prace w UE nad wprowadzeniem reguł korzystania z utworów osieroconych, a ich konkretnym efektem jest projekt dyrektywy Parlamentu Europejskiego i Rady w sprawie niektórych dozwolonych możliwości wykorzystywania utworów osieroconych, ogłoszony w maju 2011 r.[85]. Dokument, który zawiera projekt

[85] Projekt dyrektywy jest dostępny: http://ec.europa.eu/internal_market/copyright/docs/orphan-works/proposal_en.pdf; wersja polskojęzyczna: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0289:FIN:PL:PDF; wraz z projektem dyrektywy został opublikowany stosunkowo obszerny dokument dotyczący konsekwencji transgranicznego udostępniania on-line utworów osieroconych: SEC(2011) 615 final, Commision Staff Working Paper Impact Assessment on the Cross-border Online Access to Orphan Works, Accompanying the document Proposal for a Directive of the European Parliament and the Council on certain permitted uses of orphan works [dok. elektr.] http://ec.europa.eu/internal_market/copyright/docs/orphan-works/impact-assessment_en.pdf [dostęp: 12.05.2012].

składa się ze wstępnego uzasadnienia (5 stron, które zawiera informacje o czynionych w UE konsultacjach z zainteresowanymi stronami), ocenę 6 najczęściej rozważanych sposobów (wariantów) postępowania, do których należą[86]:

1) niepodejmowanie żadnych działań,
2) wprowadzenie licencji ustawowej,
3) wprowadzenie systemu rozszerzonych zbiorowych licencji (rozszerzony zbiorowy zarząd),
4) wprowadzenie specjalnych licencji udzielanych przez organizacje zbiorowego zarządzania,
5) wprowadzenie specjalnych licencji udzielanych przez podmioty publiczne,
6) wzajemne uznawanie krajowych rozwiązań w odniesieniu do utworów osieroconych.

Natomiast tekst projektu dyrektywy składa się z: 23 punktów preambuły, 13 artykułów oraz załącznika. W preambule, która stanowi typowy element każdej dyrektywy zawarto motywy podjęcia działań legislacyjnych oraz opis sytuacji związanej z powstaniem kategorii utworów osieroconych, wyjaśnienie znaczenia niektórych proponowanych rozwiązań (ujawnienie intencji ustawodawcy). W skrócie można zreferować preambułę następująco: wspieranie przepływu wiedzy (i informacji) oraz innowacji stanowi istotny element europejskiej agendy cyfrowej; jednym z działań przewidzianych w agendzie jest stworzenie ram prawnych dla digitalizowania utworów, których autorzy są nieznani lub nie można z nimi nawiązać kontaktu; co powoduje, iż nie można uzyskać ich zgody (a ta jest niezbędna) na zwielokrotnienie i publiczne udostępnienie. Wyraźnie wskazano, iż uznanie utworu za osierocony powinno zostać poprzedzone odpowiednimi, starannymi poszukiwaniami podmiotów uprawnionych, natomiast w celu ochrony interesów uprawnionych należy im zapewnić możliwość unieważnienia statusu utworu osieroconego. Następnie wyjaśniono, iż pewne instytucje kultury powinny mieć możliwość udostępniania utworów osieroconych (na podstawie licencji ustawowej) w ramach realizacji zadań publicznych. Warto podkreślić, iż równocześnie wskazano na utrzymanie w mocy rozwiązań w zakresie zbiorowego rozszerzonego zarządu oraz prawa do wynagrodzenia za korzystanie z utworów osieroconych poza zakresem realizacji zadań publicznych.

Art. 1 projektu dyrektywy zatytułowany „Przedmiot i zakres stosowania" obejmuje:

- zakres podmiotowy – (1) biblioteki publiczne, (2) placówki oświatowe, (3) muzea, (4) archiwa, (5) instytucje odpowiedzialne za działania w sferze zachowania dziedzictwa filmowego oraz (6) nadawców publicznych; oczywistym jest, iż nie są to wszyscy potencjalni użytkownicy utworów osieroconych np. wyłączeni spod niego są użytkownicy „komercyjni", podmioty prywatne czy ozz;
- zakres przedmiotowy – zasadniczo 2 kategorie utworów opublikowanych lub nadanych w państwach UE: utwory opublikowane w postaci książek, czasopism, gazet, magazynów, tekstów (other writings), które znajdują się w zbio-

[86] Interesująca analizę tych wariantów, jeszcze na etapie wstępnych prac nad projektem przedstawił Tilman Lüder (2010). *The „orphan works" challenge.* "GRUR Int.", s. 677 i nast.; zob. też Corbett, Susan (2010). *Regulation for Cultural Heritage Orphans : Time Does Matter.* "WIPO Journal" nr 2, s. 180-196.

rach instytucji wymienionych powyżej w pkt 1-4, oraz filmowe (kinematograficzne) i audiowizualne w zbiorach instytucji wymienionych powyżej w pkt 5-6 (w przypadku tej ostatniej kategorii także utwory audialne, ale tylko te utwory, które zostały wyprodukowane przez nadawców publicznych przed dniem 31 grudnia 2002 r., i znajdują się w ich archiwach, czyli status osierocony mogą uzyskać tylko utwory co najmniej 10-letnie, ale już nie późniejsze – w tym zakresie można upatrywać ograniczonego znaczenia dyrektywy dla przyszłych osieroconych utworów filmowych i audiowizualnych, pomimo krytyki tego rozwiązania[87] wydaje się, że nie stanowi ono błędu, gdyż założyć zasadnie można, że bardzo mało utworów powstałych po 2002 r. będzie miało charakter osierocony, m.in. z uwagi na stosowane obecnie zasady oznaczania praw autorskich), jeżeli chodzi o zakres przedmiotowy to projekt jest krytykowany m.in. z tego powodu, iż nie odnosi się do utworów nieopublikowanych, wśród których jak wskazuje się jest najwięcej utworów osieroconych.

Art. 2 wprowadza krótką definicję utworów osieroconych: „utwór uznaje się za utwór osierocony jeżeli podmiot praw autorskich nie jest znany lub, nawet jeżeli jest znany, nie został odnaleziony w wyniku starannego poszukiwania podmiotu praw autorskich, przeprowadzonego i zarejestrowanego zgodnie z art. 3".

Kolejny artykuł zawiera uregulowanie dotyczące starannych poszukiwań, które polegają na sprawdzeniu informacji we wskazanych źródłach dla poszczególnych kategorii utworów (ustalanych w poszczególnych krajach, przy uwzględnieniu źródeł wskazanych w załączniku). Zasadnie przyjęto, iż prowadzenie poszukiwań powinno odbywać się w państwie, w którym utwór opublikowano lub nadano po raz pierwszy. W celu uniknięcia „dublowania" poszukiwań państwa członkowskie mają rejestrować ich wyniki w stworzonych w tym celu bazach danych. Konieczność zapewnienia państwom członkowskim swobody w zakresie ustalenia tych źródeł wydaje się uzasadniona biorąc pod uwagę historyczne, kulturowe uwarunkowania, a jedynie można postulować wprowadzenie (uszczegółowienie) nieformalnej procedury weryfikowania w ramach procesu monitorowania, o których mowa w art. 11 projektu.

Projekt realizuje zasadę wzajemnego (w UE) uznawania statusu utworu osieroconego (art. 5) oraz nakłada na państwa członkowskie wprowadzenie instytucji unieważnienia statusu utworu osieroconego; unieważnienie to może odbyć się zawsze, pod warunkiem odnalezienia się twórcy (podmiotu uprawnionego), bez spełnienia jakichkolwiek innych warunków. Zasadniczo to unieważnienie ma sens praktyczny, o ile dokona się przed upływem czasu trwania autorskich praw majątkowych do utworu (lub do 5 lat po upływie tej ochrony – w związku z art. 7 pkt 5 projektu); słusznie przyjęto, iż unieważnienie może odbyć się zawsze (nie ma i nie może być możliwości „przedawnienia" unieważnienia; nie ma bowiem żadnych racji prawnych, które przemawiałyby za odmową odebrania tego statusu utworowi np. po 20 latach).

[87] Zob. krytyczne stanowisko EBIB i KOED w sprawie projektu, pkt 3, dostęp: http://www.nowyebib.info/wiadomosci/471-stanowisko-ebib-w-sprawie-dzie-osieroconych. lub http://koed.org.pl/wp-content/uploads/2012/01/konsultacje-dziela-osierocone.pdf; oraz Communi, dostęp: http://www.communia-association.org/wp-content/uploads/communia_orphan_works_policy_paper.pdf.

Kolejne przepisy (art. 6 i 7) zawierają uregulowanie w zakresie, nazwijmy to w pewnym uproszczeniu odpowiednio „obligatoryjnej" i „fakultatywnej" formy korzystania z utworów osieroconych. Pierwsza forma obejmuje dozwolony użytek w zakresie:

- udostępniania utworów osieroconych, w rozumieniu art. 3 dyrektywy 2001/29/WE[88] – czyli publicznego udostępnienia;
- dokonywania czynności zwielokrotniania utworu osieroconego, w rozumieniu art. 2 dyrektywy 2001/29/WE, do celów digitalizacji, udostępniania, indeksowania, katalogowania, ochrony i odnawiania.

Forma ta jest stosunkowo szeroka pozwala na realizację podstawowych zadań bibliotek cyfrowych; obejmuje dokonywane dla w stosunkowo szerokich celów zwielokrotnianie, które także pozwoli na realizacje zadań bibliotek cyfrowych; krytykowane ograniczenie dozwolonego wykorzystania dla celów „realizacji zadań leżących w interesie publicznym, obejmujące zapewnienie dostępu w celach kulturalnych i edukacyjnych" traktuję jako, element wyważenia pomiędzy interesami beneficjentów tej nowej formy dozwolonego użytku a interesami podmiotów uprawnionych z tytułu praw autorskich.

Na instytucjach korzystających z utworów osieroconych ciąży obowiązek prowadzeniu rejestru działań podejmowanych w ramach starannych poszukiwań oraz ogólnodostępnego rejestru przypadków wykorzystania danego utworu osieroconego (art. 6 pkt 4). Spełnienie tego drugiego obowiązku, jeżeli chodzi o jego aspekty praktyczne („monitorowanie i rejestrowanie" tego użytku przez uprawnione instytucje oraz jego ujawnianie) rodzi moje wątpliwości, aczkolwiek na pewno ułatwi „śledzenie losów" utworów osieroconych. Wydaje się, że celowym byłoby nałożenie na korzystających z utworu obowiązku podawania informacji o jego osieroconym statucie.

Druga forma, którą państwa członkowskie mogą wprowadzić dotyczy wykorzystywania utworów o statusie osieroconym dla innych celów, aniżeli związanych z realizacją zadań w interesie publicznym i jest obwarowane m.in. uzyskaniem wynagrodzenia dla podmiotów, uprawnionych z tytułu utworów osieroconych „unieważnionych" (roszczenie o to wynagrodzenie nie mogłoby się przedawnić wcześniej aniżeli po upływie 5 lat (por. art. 7 projektu).

Ta druga forma tym różni się od pierwszej, że nie musi przyjąć charakteru dozwolonego użytku (licencji ustawowej), jak wydaje się, może odbywać się, na podstawie zezwoleń udzielanych np. przez ozz; rozwiązanie to wychodzi to naprzeciw praktykom istniejącym w niektórych państwach członkowskich (w zakresie pośrednictwa ozz).

Wśród przepisów proceduralno-formalnych zasygnalizować należy art. 11 projektu dyrektywy, który wprowadza tzw. klauzulę przeglądową, pozwalającą na weryfiko-

[88] Por. art. 3: „Prawo do publicznego udostępniania utworów i prawo podawania do publicznej wiadomości innych przedmiotów objętych ochroną. 1. Państwa Członkowskie powinny zapewnić autorom wyłączne prawo do zezwalania lub zabraniania na jakiekolwiek publiczne udostępnianie ich utworów, drogą przewodową lub bezprzewodową, włączając podawanie do publicznej wiadomości ich utworów w taki sposób, że osoby postronne mają do nich dostęp w wybranym przez siebie miejscu i czasie. 2. Państwa Członkowskie powinny przewidzieć wyłączne prawo do zezwalania lub zabraniania na jakiekolwiek podawanie do publicznej wiadomości utworów, drogą przewodową lub bezprzewodową, w taki sposób, że osoby postronne mają do nich dostęp w wybranym przez siebie miejscu i czasie (...)".

wanie i zmienianie dyrektywy w przyszłości w odniesieniu do źródeł informacji (wymaganych dla wypełnienia starannych poszukiwań), jak i jej zakresu przedmiotowego (rozszerzenie np. na fonogramy).

Załącznik do projektu zawiera podstawowy wykaz źródeł, w których należy dokonać starannych poszukiwań, są nimi w odniesieniu do:

- książek: egzemplarze obowiązkowe, istniejące bazy i rejestry (ARROW, WATCH, ISBN), bazy ozz,
- czasopism, periodyków : egzemplarze obowiązkowe, ISSN, katalogi biblioteczne, bazy ozz, biblioteczne indeksy i katalogi,
- gazet i czasopism: egzemplarz obowiązkowy, stowarzyszenia wydawców i autorów, bazy ozz,
- utworów wizualnych (np. fotografie, ilustracje) zawarte w książkach, czasopismach, gazetach: wszystkie powyżej wymienione źródła, oraz agencje fotograficzne (picture agencies),
- utworów audiowizualnych: egzemplarz obowiązkowy, bazy instytucji działających w zakresie ochrony dziedzictwa filmowego, bibliotek narodowych, bazy ISAN, bazy ozz.

Podsumowując projekt dyrektywy stanowi pierwsze unijne, stosunkowo kompleksowe rozwiązanie problemu korzystania z utworów osieroconych. Moim zdaniem rozwiązanie to wychodzi na przeciw oczekiwaniom bibliotek cyfrowych, jeśli chodzi o zakres nowej formy dozwolonego użytku utworów osieroconych (jest to stosunkowo szeroki zakres, aczkolwiek nie do końca zadowalający same biblioteki). Niestety problem praktyczny z wykorzystaniem tego rozwiązania przez biblioteki cyfrowe może pojawić się, w związku z kosztownym i długotrwałym procesem ustalania statusu osieroconego, każdego indywidualnego utworu, który wymaga poczynienia starannych poszukiwań. Dlatego, przynajmniej na pierwszym etapie, stosowania nie będzie to rozwiązanie pomocne przy procesach masowych digitalizacji dużych kolekcji, zbiorów bibliotecznych.

Na koniec rozważań dotyczących utworów osieroconych chciałabym doprecyzować dwie kwestie. Po pierwsze, za zasadne uważam powołanie bazy danych utworów osieroconych. Pomimo wiążących się z tym pewnych problemów natury prawnej i organizacyjnej to wydaje mi się celowe rozważenie kwestii dotyczących:

- podmiotu (podmiotów) odpowiedzialnego i administrującego taką bazą,
- zasad/podstaw wpisywania informacji o niemożności ustalenia autora, lub braku możliwości nawiązania kontaktu z autorem (podmiotem uprawnionym),
- kosztów jej utrzymania, źródeł finansowania bazy oraz opłat pobieranych z tytułu udzielania informacji.

Może dodatkowo rozważenia wymagałaby celowość i zasady przechowywania w takiej bazie kopii utworu uznanego za osierocony (tak, aby np. móc porównywać z innymi wersjami tego samego lub podobnych utworów, których status byłby ustalany).

Po drugie, obecnie skupiamy się na istniejących utworach osieroconych, a wydaje się, że nie powinniśmy zapominać o działaniach prewencyjnych. Chodzi o to, żeby uniknąć/lub ograniczyć w praktyce w przyszłości powiększanie się skali tego zjawi-

ska w stosunku do utworów, których autorów (wydawców) obecnie jeszcze znamy i mamy z nimi kontakt.

Bibliografia

1. Barczewski, Maciej (2007). *Traktatowa ochrona praw autorskich i praw pokrewnych.* Warszawa: Wolters Kluwer Polska, 207, [1] s.
2. Barta, Janusz; Fajgielski, Paweł; Markiewicz, Ryszard (2007). *Ochrona danych osobowych. Komentarz.* Warszawa: Wolters Kluwer Polska sp. z o.o., 924 s.
3. Barta, Janusz; Markiewicz, Ryszard (1998). *Internet a prawo.* Kraków: Towarzystwo Autorów i Wydawców Prac Naukowych „Universitas", 357 s.
4. Barta, Janusz; Markiewicz, Ryszard (2007). *Rejestracja e-dzienników i czasopism* [dok. elektr.] *Rzeczpospolita* 09.10.2007 http://www.rp.pl/artykul/60888.html [odczyt: 12.01.2012].
5. Barta, Janusz; Markiewicz, Ryszard (2008). *Prawo autorskie.* Warszawa: Wolters Kluwer Polska, 395, [1] s.
6. Barta, Janusz; Markiewicz, Ryszard (2008). *Postęp techniczny w mediach.* [W:] Barta, Janusz; Markiewicz, Ryszard; Matlak, Andrzej red. (2008). *Prawo mediów.* Warszawa: Wydawnictwo Prawnicze LexisNexis, 776 s.
7. Barta, Janusz; Markiewicz, Ryszard (2010). *Prawo autorskie.* Wyd. 2. Warszawa: Wolters Kluwer Polska, 512 s.
8. Barta, Janusz; Markiewicz, Ryszard red.; Czajkowska-Dąbrowska Monika; Ćwiąkalski Zbigniew; Traple Elżbieta aut. (2011). *Prawo autorskie i prawa pokrewne: komentarz.* Warszawa: Wolters Kluwer Polska Sp. z o.o., 821 s.
9. Barta, Janusz; Markiewicz, Ryszard red. *Biała Księga. Polska – Unia Europejska. Ochrona danych osobowych. Opracowania i Analizy.* Seria: Prawo, Z. 32, Urząd Rady Ministrów, Biuro Pełnomocnika Rządu ds. Integracji Europejskiej oraz Pomocy Zagranicznej. Warszawa, 47 s.
10. Barta, Paweł; Litwiński, Paweł (2009). *Ustawa o ochronie danych osobowych. Komentarz.* Warszawa: Wydawnictwo C.H. Beck, 574 s.
11. Byford, John (2005). *Współpraca wydawców i bibliotek upoważnionych do otrzymywania egzemplarza obowiązkowego w Wielkiej Brytanii od 1610 roku: owocna, czy nie?* [dok. elektr.] *EBIB* nr 3, http://ebib.oss.wroc.pl/2005/64/byford.php [odczyt: 12.01.2012].
12. Crews, Kenneth (2008). *Study on copyright limitations and exceptions for libraries and archives* [Opracowanie przygotowane na Standing Committee on Copyright and related rights, Seventeenth session, Geneva, listopad 2008] [dok. elektr.] http://www.wipo.int/meetings/en/doc_details.jsp?doc_id=109192 [odczyt: 2.02.2012].
13. Czarny-Drożdziejko, Elżbieta (2008). *Egzemplarze obowiązkowe dla bibliotek* [W:] Barta, Janusz; Markiewicz, Ryszard; Matlak, Andrzej red. (2008). *Prawo mediów.* Warszawa: Wydawnictwo Prawnicze LexisNexis, s. 121-141.
14. Dobosz, Izabela (2011). *Prawo prasowe.* Wyd. 2. Warszawa: Wolters Kluwer Polska, 281, [1] s.
15. Dzienniki Ustaw [dok. elektr.] http://dziennikustaw.gov.pl/DziennikiUrzedowe.aspx [odczyt: 2.02.2012].
16. Flisak, Damian (2008). *Utwór multimedialny w prawie autorskim.* Warszawa : Wolters Kluwer Polska, 203, [1] s.

17. Flisak, Damian (2009). *Wydłużenie okresu ochrony praw pokrewnych, czyli jak z poważnego tematu uczynić groteskę.* „Zeszyty Naukowe Uniwersytetu Jagiellońskiego. Prace z Prawa Własności Intelektualnej" nr 105, s. 76-88.
18. Gienias, Krzysztof (2011). *Treść prawa autorskiego.* [W:] Ferenc-Szydełko, Ewa red. Tomczyk, Sławomir; Szyjewska-Bagińska, Joanna; Drzewiecki, Andrzej; Krzysztof Gienas aut. (2011). *Ustawa o prawie autorskim i prawach pokrewnych: komentarz.* Warszawa: C.H. Beck, s. 223-225.
19. Iglezakis, Ioannis (2011). *Personal Data Protection in Digital Libraries.* [W:] Iglezakis, Ioannis; Synodinou, Tatiana-Eleni; Kapidakis, Sarandos red. (2011). *E-Publishing and Digital Libraries: Legal and Organizational Issues.* Hershey; New York: information Science Reference, s. 413-439.
20. Iglezakis, Ioannis; Synodinou, Tatiana-Eleni; Kapidakis, Sarandos red. (2011). *E-Publishing and Digital Libraries: Legal and Organizational Issues.* Hershey; New York: information Science Reference, 528 s.
21. Jewuła, Beata; Stanisławska-Kloc, Sybilla (2008). *Prawo autorskie a działalność bibliotek (licencja dla bibliotek z art. 28 ust. pr. aut. i pr. pokr.).* „Zeszyty Naukowe Uniwersytetu Jagiellońskiego. Prace z Prawa Własności Intelektualnej" nr 102, s. 132-136.
22. Kocójowa, Maria red. (2004). *Przestrzeń informacji i komunikacji społecznej.* Kraków: Wydaw. Uniw. Jagiel., 416 s.
23. Kornecka, Anna (2011). *Utwory muzyczne będą dłużej chronione* [dok. elektr.] „Rzeczpospolita" 14.października.2011 r. http://www.rp.pl/artykul/732566.html?print=tak&p=0 [odczyt: 2.02.2012].
24. Maciąg, Joanna; Stanisławska, Sybilla (1997). *Internet – prasa – prawo.* „Zeszyty Naukowe Uniwersytetu Jagiellońskiego. Prace z Prawa Własności Intelektualnej" z. 68, s. 87-96.
25. Marciszuk, Piotr (2010). *Egzemplarz obowiązkowy dzieli środowisko wydawców* [dok. elektr.] Internetowy Dwutygodnik Wydawniczy http://idw.org.pl/index.php?option=com_content&view=article&id=182%3Aegzemplarz–obowizkowy–dzili–rodowisko&catid=49%3Agosc–idw&Itemid=71&lang=pl [odczyt: 12.01.2012].
26. Michalak, Arkadiusz (2010). *Przegląd cywilnoprawnych instrumentów ochrony portali internetowych.* „Przegląd Prawa Handlowego" nr 4, s. 19-24.
27. Nowińska, Ewa (2008). Prasa drukowana. [W:] Barta, Janusz; Markiewicz, Ryszard; Matlak, Andrzej red. (2008). Prawo mediów. Warszawa: Wydawnictwo Prawnicze LexisNexis, s.189 i nast.
28. Papadopoulou, Anthoula (2011). *The Digitalization of contents in Digital Libraries: Moral rights and Limits.* [W:] Iglezakis, Ioannis; Synodinou, Tatiana-Eleni; Kapidakis, Sarandos red. (2011). *E-Publishing and Digital Libraries: Legal and Organizational Issues.* Hershey; New York: Information Science Reference, s. 180-197.
29. Poźniak-Niedzielska, Maria; Szczotka, Jerzy; Mozgawa, Marek (2006). *Prawo autorskie i prawa pokrewne. Zarys wykładu.* Red. M. Poźniak-Niedzielska, Bydgoszcz; Warszawa; Lublin 2006, s.101.
30. Prawelska-Skrzypek, Grażyna (2003). *Polityka kulturalna polskich samorządów.* Kraków: Wydaw. Uniw. Jagiel., 258 s.
31. *Raport WIPO – Committee on Development and Intellectual Property* (CDIP), Seventh Session, Geneva, May 2 to 6 (2011). Scoping Study on Copyright and related rights and the Public Domain [przygotowany przez S. Dusollier] [dok. elektr.] http://www.wipo.int/meetings/en/details.jsp?meeting_id=22102 [odczyt: 2.12.2011].
32. *Report of the 'Comite des Sages'* (2011). [Raport „Grupy Sages" ze stycznia 2011 r., w skład którego wchodzili Elisabeth Niggemann, Jacques De Decker, Maurice Lévy]

[dok. elektr.] http://ec.europa.eu/information_society/activities/digital_libraries/doc/refgroup/final_report_cds.pdf [odczyt: 2.02.2012].

33. Stanisławska-Kloc, Sybilla (1999). *Przedmiot prawa autorskiego*. [W:] Barta, Janusz; Markiewicz, Ryszard red. (1999). *Prawo autorskie a postęp techniczny*, Kraków: Universitas, s. 32-35.
34. Stanisławska-Kloc, Sybilla (2007). *Utwory osierocone*. Prawo własności intelektualnej – wczoraj, dziś i jutro. „Zeszyty Naukowe Uniwersytetu Jagiellońskiego. Prace z Prawa Własności Intelektualnej" z. 100, s. 453-476.
35. Stanisławska-Kloc, Sybilla (2009). *Prawne aspekty ochrony własności intelektualnej*. [W:] Wisła, Rafał red. (2009). *Rozwój współpracy sektorów publicznego i prywatnego z wykorzystaniem praw własności intelektualnej*. Warszawa: Ministerstwo Gospodarki, s. 81-84.
36. Stanisławska-Kloc, Sybilla (2009). *Własność intelektualna w projektach partnerstwa publiczno-prywatnego*. [W:] Wisła, Rafał red. (2009). *Rozwój współpracy sektorów publicznego i prywatnego z wykorzystaniem praw własności intelektualnej*. Warszawa: Ministerstwo Gospodarki, s. 97-101.
37. Stanisławska-Kloc, Sybilla (2010). *Udostępnianie kolekcji cyfrowych utworów fotograficznych i filmowych*. Wybrane zagadnienia z zakresu prawa autorskiego. [W:] *Zbiory fotograficzne i filmowe w Internecie. Aspekty prawne i technologiczne* (2010). Gdańsk: Biblioteka Uniwersytecka, s. 58.
38. Stanisławska-Kloc, Sybilla; Ptak, Łukasz (2007). *Wybrane skróty, oznaczenia, symbole i zwroty używane w zakresie własności intelektualnej; Nazwy i skróty organizacji oraz konwencji z zakresu własności intelektualnej. Prawo własności intelektualnej – wczoraj, dziś i jutro*. Zeszyty Naukowe Uniwersytetu Jagiellońskiego. Prace z Prawa Własności Intelektualnej z. 100, s. 653- 671.
39. Stec, Piotr (2008). *Pojęcie i zakres domeny publicznej w prawie polskim*. „Przegląd Sądowy" R. 18, [nr] 4, s. 45-64.
40. Szczepańska, Barbara (2007). *Prawo autorskie – ochrona dziel elektronicznych*. [W:] Woźniak-Kasperek, Jadwiga; Franke, Jerzy red. (2007). *Biblioteki cyfrowe: projekty, realizacje, technologie: praca zbiorowa* (2007). Warszawa: Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich, s. 51-66.
41. Targosz, Tomasz (2007). *Domena publiczna w prawie autorskim*. Prawo własności intelektualnej – wczoraj, dziś i jutro. „Zeszyty Naukowe Uniwersytetu Jagiellońskiego. Prace z Prawa Własności Intelektualnej" z. 100, s. 531-554.
42. *The Proposed Directive for a Copyright Term Extension – A backward–looking packane* (2008) [dok. elektr.] Joint Academic Statement on Term Extension (CIPPM/CIPIL/IViR/MPI–IP), 27 October 2008, http://www.cippm.org.uk/downloads/Term%20Statement%2027_10_08.pdf [odczyt: 12.01.2012].
43. Timmer, John (2011). *Authors' Guild sues universities over book digitization project* [dok. elektr.] http://arstechnica.com/tech–policy/news/2011/09/authors–guild–sues–universities–over–book–digitization–project.ars [odczyt: 2.02.2012].
44. Waglowski, Piotr (2011a). *Biblioteczne ograniczanie konsumentów w reprodukcji „do 1 arkusza wydawniczego" klauzulą abuzywną* [dok. elektr.] http://prawo.vagla.pl/node/9588 [odczyt: 12.12.2011].
45. Waglowski, Piotr (2011b). *Reprografia w bibliotekach* [dok. elektr.] http://prawo.vagla.pl/node/9176 [odczyt: 12.12.2011].

Koszty i finansowanie digitalizacji

Małgorzata Kowalska

W tekście przedstawiono zarówno koszty, jak i możliwości finansowania digitalizacji. Utrwalanie bogatego dziedzictwa kulturowego w postaci cyfrowej jest bowiem wyzwaniem ambitnym i kosztownym. Niezależnie od tego, czy planowana cyfryzacja ma być realizowana samodzielnie czy we współpracy z partnerem zewnętrznym, jako proces wieloletni czy doraźny, wymaga odpowiedniego zaplecza finansowego.

Uruchomienie nawet najmniejszego projektu digitalizacyjnego poprzedzać musi wnikliwa analiza kosztów poszczególnych jego elementów składowych. Dopiero po ich oszacowaniu można rozpocząć etap poszukiwania odpowiednich funduszy, a następnie realizacji inwestycji. W przypadku procesu digitalizacji problem finansowania nie znika jednak w momencie uruchomienia projektu, ale dotyczy także utrzymywania i rozwijania zasobu w czasie. Jak podkreślają liczne organizacje wspierające ochronę dziedzictwa kulturowego z wykorzystaniem techniki cyfrowej, w tym Komisja Europejska i Amerykańska Organizacja Normalizacyjna, dla zasobów zawierających materiały o długotrwałym znaczeniu należy tworzyć długofalowe plany ich dalszego rozwoju uwzględniające zarówno perspektywy organizacyjne, techniczne, jak i finansowe [Framework 2007].

Outsourcing vs. własna pracownia digitalizacji

Wysokość całkowitych kosztów procesu digitalizacji determinują w pierwszej kolejności sposób realizacji procesu, wielkość inwestycji oraz jej przeznaczenie. Innych nakładów finansowych wymaga przecież digitalizacja prowadzona z udziałem partnera zewnętrznego, innych – uruchomienie własnej pracowni.

Powierzanie cyfryzacji (w części lub całości) podmiotom zewnętrznym (ang. *outsourcing*) wydaje się szczególnie celowe w przypadku tych instytucji kultury, których ze względów finansowych, lokalowych czy personalnych po prostu nie stać na uruchomienie specjalistycznej pracowni. Realizując jednorazowy bądź niewielki projekt digitalizacyjny bezzasadne wydaje się inwestowanie w zakup drogiego sprzętu i opro-

gramowania oraz szkolenie i zatrudnianie wyspecjalizowanych pracowników. Parafrazując słowa Henry'ego Forda, w sytuacji, gdy danej pracy nie można wykonać wydajniej, taniej i lepiej niż konkurenci, lepiej zlecić ją komuś innemu [Outsourcing 2007].

Za korzystaniem z usług zewnętrznych przemawiać może jednak nie tylko nierentowność uruchomienia pracowni digitalizacji we własnych pomieszczeniach, ale także pewne cechy charakterystyczne materiału poddawanego konwersji, tj. jego szczególna wartość, daleko posunięty proces zniszczenia, wrażliwość bądź niestandardowy format. W związku z tym, iż cyfryzacja tego typu dokumentów wymaga zastosowania wysoko specjalistycznego i bardzo kosztownego sprzętu, a także odpowiednich umiejętności i kwalifikacji, lepiej powierzyć ją profesjonalistom.

Doświadczenia bibliotek na świecie udowadniają, że outsourcing w digitalizacji wdrażany jest na różnych etapach realizacji tego procesu, od tworzenia cyfrowych wersji dokumentów i opisu ich kolekcji za pomocą metadanych poczynając, na archiwizacji i konserwacji zasobu czy jego promocji kończąc [Paradowski 2010, s. 153-154]. W Polsce do zadań najczęściej zlecanych firmom komercyjnym należą skanowanie i mikrofilmowanie materiałów, cyfrowa obróbka oraz przetwarzanie i rejestracja plików na nośnikach elektronicznych. Nieco rzadziej zakres tych usług obejmuje tworzenie opisu kolekcji cyfrowych i konsulting prawny.

Bez względu na obszar zastosowania, outsourcing w digitalizacji niesie ze sobą kilka niewątpliwych zalet. Z punktu widzenia kwestii finansowych są to:

- eliminacja kosztów związanych z koniecznością przygotowywania, wyposażania i utrzymywania pomieszczeń,
- likwidacja kosztów związanych z zakupem sprzętu, jego utrzymaniem, konserwacją i modernizacją,
- redukcja kosztów związanych z koniecznością szkolenia, rekrutacji i wynagradzania pracowników [Paradowski 2010, s. 153].

Badania stanu digitalizacji bibliotek polskich przeprowadzone w latach 2003-2005 [zob. Potrzebnicka 2005; 2006; Kowalska 2007] dowodzą, że cyfryzacja zbiorów własnych z udziałem partnera zewnętrznego stała się już codziennością. Wśród realizatorów tego procesu najczęściej wymieniane są firmy Digital-Center z Poznania, Mikrofilm Service z Raszyna, Telecomp Service oraz DDP Sp. z o.o. z Warszawy. Szczególnym uznaniem cieszy się współpraca z firmą poznańską. Profesjonalizm jej informatyków, umiejętność obchodzenia się z dokumentami zabytkowymi, używanie wysoko specjalistycznych skanerów, krótki czas skanowania, otwartość pracowników firmy na wszelkie sugestie natury merytorycznej i technicznej oraz rzetelność i terminowość wykonywania usług, powodują, że wiele instytucji kultury udziela jej rekomendacji.

W związku z tym, iż rynek usług digitalizacyjnych w Polsce jest bardzo rozchwiany, niezwykle trudno wycenić rzeczywiste koszty digitalizacji z udziałem partnera zewnętrznego. Jak zauważają przedstawiciele firm komercyjnych, pod terminem *digitalizacja* każdy rozumie coś innego (konserwacja i skanowanie?, mikrofilmowanie ze skanowaniem i obróbką cyfrową?, archiwizacja i zarządzanie materiałem cyfrowym?),

a to utrudnia porównywanie istniejących ofert. Bez wątpienia każdy projekt digitalizacyjny przygotowywany jest indywidualnie, a na jego ostateczne koszty wpływają wielkość i przeznaczenie projektu, stan zachowania obiektów, stosowana technologia oraz rodzaj „produktu końcowego" (plik, kopia użytkowa, prezentacja multimedialna, biblioteka cyfrowa). Samo oszacowanie kosztów procesu skanowania niewiele daje, bowiem przy stale malejących cenach[1] zdarza się, że wykonanie skanu karty dokumentu w jednym przypadku kosztuje zaledwie 10 gr, w innym – 50 zł brutto. Niestety, najwięcej wydatków (niekiedy nawet ponad 50%) generują pozostałe czynności, tj. przygotowanie materiału, transport, konwersja czy indeksacja [Buczek 2011]. O ostatecznej wysokości nakładów decydują wreszcie także wymagania stawiane samym digitalizatom i czas potrzebny na „przerobienie" formy analogowej na cyfrową. Im lepsza jakość produktu końcowego, tym bardziej precyzyjny, dłuższy, ale i droższy proces [Komarnicki 2011].

Z całą pewnością największych inwestycji finansowych wymaga uruchomienie własnej linii technologicznej digitalizacji i mikrofilmowania zbiorów. Wiąże się ono bowiem z koniecznością poniesienia wydatków na:

- zakup sprzętu i wdrażanie nowych technologii, na których zasoby cyfrowe mogą być przechowywane, skutecznie zarządzane i udostępniane,
- szkolenie i zatrudnienie pracowników odpowiedzialnych za selekcję, opracowanie i digitalizację obiektów,
- przygotowanie analiz i ekspertyz prawnych,
- pozyskiwanie praw autorskich,
- archiwizację i konserwację zasobów,
- budowanie sieci współpracy krajowej i międzynarodowej,
- promocję zasobów cyfrowych w Polsce i za granicą [Bednarek-Michalska 2010].

Mimo że wyżej wymienione aspekty procesu digitalizacji nie wyczerpują katalogu przesłanek, które należy rozważyć podejmując decyzję o samodzielnym prowadzeniu konwersji cyfrowej, to bez wątpienia stanowią warunek *sine non qua* każdego szacunkowego kosztorysu. Przykładowe szczegółowe koszty, które należy w nim uwzględnić prezentuje tabela 1, stanowiąca aneks do tej pracy. Warto podkreślić, że zamieszczone w tabeli koszty poszczególnych składowych digitalizacji są wyłącznie wartościami szacunkowymi. Ich zestawienie unaocznia jednak, że największych nakładów finansowych wymaga organizacja nowoczesnej pracowni i zakup specjalistycznego sprzętu (minimum 300 tys. zł, a w przypadku pracowni materiałów audiowizualnych nawet 700 tys. Euro [Program digitalizacji 2009]). Kolejne wydatki generują etaty pracow-

[1] Dane pochodzące z prywatnych pracowni archiwizacji potwierdzają spadek cen usług reprograficznych. O ile w roku 2005 opłata za skanowanie obiektów z oryginału nie przekraczała 3 zł za skan formatu A4, skanowanie mikrofilmu lub mikrofiszy – 2 zł za skan, a nagranie płyty CD lub DVD 15 zł (wraz z kosztami nośników), to w roku 2011 ceny te wyniosły odpowiednio: od 1 do 3 zł za skan czarno-biały i od 5 do 10 zł za skan kolorowy, a dla mikrofilmów nawet od 0,45 gr do 1 zł za skan. Potaniał także zapis skanów na płyty CD lub DVD i wahał się w granicach od 3 do 5 zł (wraz z kosztami nośników).

Por. przykładowy cennik usług reprograficznych firmy komercyjnej: *Usługi reprograficzne* [dok. elektr.] Biblioteka Gdańska PAN http://www.bgpan.gda.pl/ [odczyt: 27.06.2011].

nicze, które w średniej wielkości pracowni digitalizacji (zatrudniającej 2–3 osoby) wynieść mogą około 70 tys. zł, oraz przeprowadzenie samego procesu konwersji, na który przewidzieć trzeba fundusze rzędu 10–30 tys. zł rocznie [Kowalska 2007, s. 256–257]. Niestety, najtrudniej wycenić wysokość wynagrodzenia za świadczone usługi informatyczne i konsultingowe, bowiem ich ceny – zależne od szczegółowego zakresu zlecenia, harmonogramu i warunków realizacji – bywają każdorazowo przedmiotem indywidualnych ustaleń.

Mecenat państwa

Ze względu na to, iż koszty digitalizacji zazwyczaj przekraczają możliwości finansowe pojedynczych instytucji kultury, podmioty te zmuszone są do poszukiwania zewnętrznych źródeł finansowania.

W Polsce prace digitalizacyjne mogą być finansowane z pieniędzy publicznych, funduszy strukturalnych, dotacji celowych, projektów międzynarodowych oraz środków pochodzących od prywatnych darczyńców.

Poszukiwanie informacji na temat potencjalnych źródeł finansowania warto rozpoczynać od serwisów informacyjnych poszczególnych ministerstw (szczególnie Kultury i Dziedzictwa Narodowego, Nauki i Szkolnictwa Wyższego, Edukacji Narodowej, Rozwoju Regionalnego) i fundacji (Bankowej im. Leopolda Kronenberga, Batorego, Roberta Boscha, Współpracy Polsko-Niemieckiej), które jako instytucje od lat zajmujące się finansowaniem kultury, udostępniają informacje najbardziej kompletne, a do tego rzetelnie opracowane. Dodatkową pomocą mogą służyć witryny organizacji pozarządowych, strony WWW dedykowane poszczególnym programom strukturalnym i funduszom europejskim oraz portale bibliotekarskie, w których poza ogólnymi informacjami o istniejących subwencjach niejednokrotnie udostępniane są wykazy instytucji finansujących kulturę, poradniki, generatory wniosków czy przykłady dobrych praktyk[2].

Rozważając wybór stosownej formy finansowania digitalizacji ze środków publicznych warto pamiętać, że choć instytucja mecenatu państwa pozwala na ubieganie się o dotacje wszystkim zainteresowanym, to liczba aplikujących systematycznie rośnie. Przy dużej konkurencji, uzyskanie dofinansowania nie jest zatem zabiegiem prostym i wymaga umiejętności przebicia się z pomysłem, wytrwałości i cierpliwości w wypełnianiu różnorodnych wniosków oraz samozaparcia w pokonywaniu przeszkód administracyjnych. Niestety, również znalezienie instytucji, firmy czy osoby prywatnej, która bezinteresownie wyasygnuje określone fundusze na dofinansowanie pewnego obszaru działalności digitalizacyjnej (rzadkością jest bowiem finansowanie całości danego projektu) nie należy do najprostszych. Wielu potencjalnych darczyńców cechuje postawa zamknięta, wyrastająca z przekonania, że finansowanie kultury to proces mało opłacalny i przynoszący mniej wymierne efekty niż dotowanie przedsięwzięć

[2] Przykłady źródeł zawierających informacje nt. możliwości finansowania projektów digitalizacyjnych zob. Bednarek-Michalska 2010, s. 240-242 oraz Burska 2004, s. 166-168.

biznesowych czy inwestycyjnych. By przekonać potencjalnego darczyńcę do danego pomysłu, warto zatem skupić się na właściwym przygotowaniu wniosku projektowego. Powinien on uwzględniać przynajmniej:

- charakterystykę instytucji, która stara się o fundusze wraz z rzeczową prezentacją projektu i wskazaniem jego ważkości,
- uzasadnienie istnienia deficytu i potrzeby wsparcia finansowego z określeniem wysokości udziału środków własnych oraz zakresów odpowiedzialności poszczególnych podmiotów uczestniczących w projekcie,
- szczegółowy opis projektu wraz z metodologią, harmonogramem, terminarzem i kosztorysem prac oraz wskazaniem jego oczekiwanych efektów oraz szans jego powodzenia i niepowodzenia.

Oczywiście o powodzeniu danej inicjatywy w trybie konkursowym decyduje nie tylko dobrze przygotowana oferta, ale także społeczne zapotrzebowanie na nią, a w przypadku sponsoringu dodatkowo „przystawalność" do założeń potencjalnego darczyńcy. Przystępując do pozyskiwania funduszy należy więc przeanalizować, jakie możliwości oferuje dana instytucja finansująca i jakich ekwiwalentów oczekuje w zamian.

Zgodnie z narodowym programem digitalizacji organem odpowiedzialnym za tworzenie polskich zasobów cyfrowych jest Ministerstwo Kultury i Dziedzictwa Narodowego, a więc to na nim w dużej mierze spoczywa obowiązek dotowania projektów digitalizacyjnych. Wypełnianiu tego zadania służą corocznie uruchamiane programy operacyjne, w ramach których organizacje pozarządowe, instytucje upowszechniania kultury, muzea, archiwa, biblioteki, media, instytucje i szkoły filmowe, jak i przedsiębiorstwa działające dla sektora kultury mogą aplikować o określone środki finansowe. Wśród dotychczas zainicjowanych programów, digitalizacja jako priorytet znalazła się m.in. w programach *Dziedzictwo kulturowe* (2005-2009) oraz *Zasoby cyfrowe* (2010 r.). Możliwość ubiegania się o fundusze na budowę i wyposażanie specjalistycznych pracowni oraz upowszechnianie dóbr kultury w środowisku cyfrowym stwarzały także programy *Edukacja kulturalna* oraz *Promocja czytelnictwa* (obydwa uruchamiane cyklicznie od 2005 r.)[3].

Od roku 2011 realizację wszystkich zadań związanych z finansowaniem procesów digitalizacyjnych przeniesiono do nowo uruchomionego *Wieloletniego Programu KULTURA+*. Jego celem strategicznym jest poprawa dostępu do kultury oraz uczestnictwa w życiu kulturalnym na terenach wiejskich i wiejsko-miejskich, poprzez modernizację i budowę infrastruktury bibliotecznej oraz digitalizację zasobów polskich muzeów, bibliotek i archiwów. Program składa się z dwóch priorytetów:

- „Biblioteka+. Infrastruktura bibliotek", w ramach którego wspierany jest remont, przebudowa, rozbudowa i budowa budynków bibliotek w małych gminach,

[3] Regulaminy poszczególnych programów, ich priorytety i uprawnieni beneficjenci w układzie chronologicznym od roku 2008 – dostępne na stronie: Ministerstwo Kultury i Dziedzictwa Narodowego (2011a); Zob. też Bednarek-Michalska 2010, s. 238-262.

- „Digitalizacja", który zmierza do poszerzenia i ułatwienia dostępu do cyfrowych zasobów polskiego dziedzictwa kulturowego, poprzez cyfryzację zbiorów i rozbudowę pracowni digitalizacyjnych [Ministerstwo Kultury i Dziedzictwa Narodowego 2011b].

Obsługa drugiego priorytetu została powierzona Narodowemu Instytutowi Audiowizualnemu, a wśród zadań objętych finansowaniem wyróżniono:

- organizowanie sieci pracowni digitalizacyjnych, która znacząco przyśpieszy digitalizację zasobów kulturowych,
- digitalizację zasobów dziedzictwa kulturowego w tym: materiałów bibliotecznych, muzealiów, archiwaliów oraz materiału audiowizualnego,
- organizowanie sieci profesjonalnych repozytoriów cyfrowych umożliwiających właściwe przechowywanie zdigitalizowanych zbiorów,
- udostępnienie zbiorów za pośrednictwem Internetu w postaci cyfrowych muzeów, bibliotek, archiwów i wortali audiowizualnych,
- zwiększanie dostępności zbiorów polskich muzeów, bibliotek, archiwów oraz zbiorów audiowizualnych [Ministerstwo Kultury i Dziedzictwa Narodowego 2011b].

Na realizację wyżej wymienionych celów, w okresie programowania obejmującym lata 2011-2015, minister kultury przeznaczył łącznie 85 mln zł (rocznie po 18 mln, w ostatnim roku finansowania – 13 mln). Z kwoty tej przewiduje się dofinansowywanie projektów, których okres realizacji nie przekracza 12 miesięcy. Inwestycje państwowych instytucji kultury i archiwów państwowych mogą być dotowane w 100%, zaś projekty samorządowych instytucji kultury i instytucji filmowych w 80%. Minimalna kwota wsparcia wynosi przy tym 100 tys. zł, maksymalna – 2 mln zł. Program nie przewiduje przekazywania środków na finansowanie digitalizacji obiektów nie zachowujących standardów technicznych i metadanych [Ministerstwo Kultury i Dziedzictwa Narodowego 2011b].

Z uwagi na fakt, że w myśl narodowego programu digitalizacji instytucją wspomagającą tworzenie polskich treści cyfrowych jest Ministerstwo Nauki i Szkolnictwa Wyższego, także z jego inicjatywy uruchamiane są programy, granty i dotacje, z których pośrednio możliwe jest dotowanie przedsięwzięć digitalizacyjnych. Szansę taką stwarza *Program rozwoju infrastruktury informatycznej nauki na lata 2007-2013*. W jego ramach ośrodki akademickie mogą ubiegać się o finansowe wsparcie na utrzymanie centrów usług i obliczeń dużej mocy (przechowujących m.in. biblioteki cyfrowe i bazy danych), budowę wirtualnych bibliotek nauki oraz tworzenie zasobów cyfrowych i usług dla e-społeczeństwa [Program rozwoju 2007].

Niestety, realizacja wszystkich polskich inicjatyw digitalizacyjnych z pieniędzy publicznych nie jest możliwa. Dlatego ważnym elementem polityki ich finansowania są środki pochodzące z funduszy strukturalnych i programów wspólnotowych Unii Europejskiej.

Fundusze europejskie

Najbardziej powszechnym unijnym źródłem finansowania inicjatyw związanych z ochroną dorobku kulturowego w postaci cyfrowej są Krajowe Programy Operacyjne. Z ogółem sześciu uruchomionych, trzy (*Program Operacyjny Infrastruktura i Środowisko, Program Operacyjny Kapitał Ludzki, Program Operacyjny Innowacyjna Gospodarka*) umożliwiają dotowanie przedsięwzięć digitalizacyjnych.

Największą alokację środków na kulturę (576 mln euro) przewiduje *Program Operacyjny Infrastruktura i Środowisko*. Istniejący w jego ramach priorytet XI pt. „Kultura i dziedzictwo kulturowe" zakłada wykorzystanie potencjału kultury i dziedzictwa kulturowego o znaczeniu światowym i europejskim dla zwiększenia atrakcyjności Polski. O wsparcie finansowe z programu mogą ubiegać się realizatorzy projektów z zakresu konserwacji i zabezpieczenia ruchomych obiektów dziedzictwa kulturowego, takich jak zabytkowe muzealia, starodruki, księgozbiory, oraz projektów z zakresu rozwoju cyfrowych zasobów bibliotecznych, archiwalnych i muzealnych. Dotacją mogą zostać objęte także działania związane z tworzeniem wirtualnych instytucji kultury[4].

Programami pośrednio pozwalającymi na uzyskanie subwencji na digitalizację są także dwa pozostałe z wymienionych. Mimo że nie uruchomiono w nich odrębnych osi priorytetowych dedykowanych ochronie dziedzictwa kulturowego w postaci cyfrowej, to stworzono tu pewne mechanizmy finansowania systemu oświaty i związanej z nią infrastruktury informatycznej. I tak, w ramach *Programu Operacyjnego Kapitał Ludzki* (w obrębie priorytetu III *Wysoka jakość systemu oświaty*) mogą być dotowane projekty mające na celu rozwój e-learningu jako formy kształcenia oraz służące upowszechnianiu innowacyjnych programów i metod nauczania[5]. Do działań takich z pewnością zaliczyć można tworzenie wirtualnych zasobów edukacyjnych, portali dydaktycznych czy tutoriali online. Drugi z wymienionych – *Program Operacyjny Innowacyjna Gospodarka* – w ramach priorytetów „Badania i rozwój nowoczesnych technologii" oraz „Infrastruktura sfery B+R" przewiduje wsparcie finansowe dla projektów z zakresu rozwoju zasobów informacyjnych nauki w postaci cyfrowej, zaawansowanych aplikacji i usług teleinformatycznych dla środowiska naukowego, jak i projektów rozwojowych, tj. posiadających szczególny wymiar społeczny[6].

Doskonałym uzupełnieniem funduszy strukturalnych są unijne Regionalne Programy Operacyjne na lata 2007-2013 (łącznie 16, po jednym dla każdego z województw). Instytucjami zarządzającymi są w tym przypadku urzędy marszałkowskie,

[4] Łączna suma środków finansowych zaangażowanych w realizację *Programu Operacyjnego Infrastruktura i Środowisko* wynosi 37,6 mld euro, z czego wkład unijny to 27,9 mld euro, krajowy – 9,7 mld euro. Por. *Program Operacyjny Infrastruktura i Środowisko* (2007).

[5] Całość kwoty, jaką przewidziano na realizację *Programu Operacyjnego Kapitał Ludzki*, wynosi prawie 11,5 mld euro. Wkład finansowy Europejskiego Funduszu Społecznego to ponad 9,7 mld euro, pozostałą część stanowią środki krajowe. Por. *Program Operacyjny Kapitał Ludzki (2007)*.

[6] Wsparcie przewidziane w ramach *Programu Operacyjnego Innowacyjna Gospodarka* przekroczy kwotę 9,71 mld euro, z czego 8,25 mld stanowić będą środki pochodzące z Europejskiego Funduszu Rozwoju Regionalnego, a 1,46 mld euro z budżetu krajowego. Por. *Program Operacyjny Innowacyjna Gospodarka, 2007-2013. Narodowe Strategiczne Ramy Odniesienia 2007-2013 (2007)*.

su finansowego, należy w znaczącym stopniu zwiększyć ilość środków przeznaczanych na ten cel. Ze względu na uwarunkowania lokalne (konteksty społeczne, polityczne i kulturowe) finansowanie cyfryzacji powinno odbywać się jednak przede wszystkim na szczeblu krajowym bądź regionalnym, nie zaś europejskim. Ten typ udzielania wsparcia zarezerwować należy bowiem dla projektów o charakterze transgranicznym oraz dla portalu Europeana [Niggermann 2010].

Mając na uwadze wysokie koszty digitalizacji oraz jej nadrzędny cel, jakim jest zwiększenie dostępu do europejskiego dziedzictwa kulturowego, rekomendacji wymaga także zawieranie partnerstw publiczno-prywatnych. Udział środków finansowych przeznaczanych na ten cel z sektora komercyjnego nie powinien jednak zastępować subwencji publicznych. By współpraca instytucji kultury ze środowiskiem gospodarczym przyczyniała się do powstawania nowych możliwości i rozwoju innowacyjnych usług w sektorach edukacji, badań naukowych i turystyki, należy tworzyć korzystne warunki do angażowania się różnych podmiotów w digitalizację. W tym kontekście szczególnego rozważenia wymaga wprowadzenie systemu zachęt podatkowych do inwestowania środków prywatnych oraz powoływanie klastrów przedsiębiorstw we współpracy z instytucjami kultury[10].

Sponsoring i fundraising

Jak słusznie konstatują analitycy Komisji Europejskiej, partnerstwo z sektorem publicznym może sprzyjać przyśpieszeniu digitalizacji utworów i zbiorów stanowiących własność publiczną, zwłaszcza w czasach, gdy środki publiczne są ograniczone.

W Stanach Zjednoczonych czy Japonii instytucje kultury od wielu lat aktywnie korzystają ze sponsoringu korporacyjnego, fundacji oraz indywidualnych donacji odpisywanych ze zobowiązań podatkowych. Szacuje się, że te trzy źródła stanowią ponad 40% rocznego przychodu amerykańskiego sektora kulturalnego [Gordon 2011], a niektóre biblioteki wypracowują w ten sposób nawet 50-60% rocznego budżetu [Gębołyś 2009]. W Europie – ze względu na uwarunkowania polityczne i podatkowe – zwyczaj „prywatnych" datków jest znacznie mniej popularny niż w USA. Nie oznacza to jednak, że Europejczycy nie wykorzystują wypracowanych przez Amerykanów metod i sposobów działania. Obszarem ich zastosowania jest coraz częściej właśnie digitalizacja.

W jej przypadku szeroko rozumiany sponsoring przyjmuje zazwyczaj formę finansowego wsparcia zakupu odpowiedniego sprzętu, pokrycia kosztów usług skanowania, mikrofilmowania lub archiwizacji materiałów, czy wreszcie budowy infrastruktury informatycznej dla zasobu cyfrowego. Ponieważ ten rodzaj finansowania zawsze zorientowany jest na wzajemność, instytucje korzystające z funduszy prywatnych pamiętać muszą o zapewnieniu określonych korzyści sponsorowi. W przypadku pro-

[10] Dokument omawia istniejące modele partnerstwa publiczno-prywatnego i precyzuje, w jaki sposób nawiązywać je w zakresie digitalizacji, jak pobudzać przepływ funduszy prywatnych oraz jak chronić interesy publicznych instytucji kultury, w sytuacji, gdy digitalizacja ich zbiorów finansowana jest ze środków prywatnych [Niggermann 2010].

jektów digitalizacyjnych korzyściami takimi mogą być stworzenie szansy zbudowania pozytywnego wizerunku lub prestiżu firmy w środowisku lokalnym poprzez umożliwienie dotarcia do wybranej części rynku i określonej grupy potencjalnych klientów, zamieszczenie reklamy/logo sponsora bądź odnośnika do jego witryny na stronie biblioteki cyfrowej, udział w gremiach oraz bezpłatny dostęp do zasobów licencjonowanych danej instytucji.

Pewną odmianą sponsoringu jest fundraising. Pod tym terminem w kontekście działalności informacyjnej rozumiane jest udostępnienie bibliotece określonych zasobów (funduszy, czasu, usług, materiałów, know-how) na konkretny cel, mające jednak jednokierunkową postać, tzn. bez wymogu wzajemności [Gębołyś 2009]. Doświadczenia wielu krajów europejskich, w tym np. Niemiec, Szwecji, Wielkiej Brytanii i Czech, dowodzą, że fundraising na stale zagościł już w realiach bibliotecznych i przybiera różnorodną postać: kwesty, promocji projektów bibliotecznych w Internecie (bannery), organizacji eventów i spotkań okolicznościowych, subwencji, dotacji, fundacji lub koła przyjaciół biblioteki [Vollmer 2001]. Ta ostatnia forma wspierania działalności bibliotek coraz częściej wykorzystywana jest także jako sposób zbierania funduszy na ochronę, zabezpieczanie i digitalizację dziedzictwa kulturowego.

Godną polecenia formą fundraisingu w digitalizacji, szczególnie dobrze rozwiniętą w krajach niemieckiego obszaru językowego, jest tzw. „adopcja książki" (niem. *Buchpatenschaft*), polegająca na sponsorowaniu kosztów zabiegów renowacji lub cyfryzacji wybranego przez fundatora dzieła. W zamian za wyasygnowanie odpowiedniej pomocy finansowej, donator liczyć może na szereg uprawnień. Poza otrzymaniem kopii cyfrowej „sponsorowanego" dokumentu są to specjalne certyfikaty i nagrody (od 500 euro), bezpłatne karty wstępu na koncerty, wykłady i wystawy przygotowywane przez i odbywające się w bibliotece (do 500 euro), zaproszenia do uczestnictwa w szkoleniach i wycieczkach kursowych organizowanych zarówno w bibliotece, jak i poza nią (powyżej 1 tys. euro) [*Aktion Buchpatenschaft* 2007-2011].

Mimo że dotowanie procesu digitalizacji przez sektor prywatny może wnosić do niego znaczący wkład, ten rodzaj partnerstwa nie jest jeszcze w Polsce wystarczająco dobrze rozwinięty. Umieszczane w witrynach bibliotecznych [*Sponsorzy* 2009; *Lista sponsorów* 2008-2011] i na stronach poszczególnych bibliotek cyfrowych dLibry[11] wykazy instytucji wspierających cyfryzację polskiego dziedzictwa kulturowego potwierdzają jednak tezę, że fundraising korporacyjny i w polskich realiach istnieje. Na listę stałych sponsorów polskiej digitalizacji wpisały się już Fundacja Bankowa im. Leopolda Kronenberga[12], Fundacja Batorego, Fundacja Roberta Boscha, Fundacja

[11] Przykłady pozycji podanych digitalizacji z udziałem sponsora: 1. firmy prywatnej – *Odra (Wrocław 1961) R.1 Nr 1 marzec 1961* [dok. elektr.] *Biblioteka Cyfrowa Uniwersytetu Wrocławskiego* http://www.bibliotekacyfrowa.pl/dlibra/docmetadata?id=28031&from =publication [odczyt: 27.06.2011]; 2. towarzystwa naukowego – *Ziemia Dobrzyńska: Zeszyty Historyczne Dobrzyńskiego Oddziału WTN, V: Dekanat Rypiński: Z archiwaliów diecezjalnych płockich XIX wieku* [dok. elektr.] *Kujawsko-Pomorska Biblioteka Cyfrowa* http://kpbc.umk.pl/dlibra/docmetadata?id=41364&from=publication& [odczyt: 27.06.2011]; 3. biblioteki szkoły wyższej – *Przemysł i Handel* R. 1 (1920) [dok. elektr.] *e-biblioteka Uniwersytetu Warszawskiego* http://ebuw.uw.edu.pl/dlibra/publication?id=4320&tab=3 [odczyt: 27.06.2011].

[12] Fundacja Kronenberga jest postrzegana jako jedna z największych i najważniejszych instytucji sektora organizacji pozarządowych w Polsce oraz niekwestionowany lider wśród darczyńców korpora-

Współpracy Polsko-Niemieckiej, Polsko-Amerykańska Fundacja Wolności, Fundacja Polskiej Biblioteki Internetowej i Fundacja im. Ericha Brosta. Aktywnością na tym polu wykazują się także firmy prywatne (np. PKN Orlen, Samsung, Yes, Rentier. pl, ZETO S.A. z Tarnowa, itp. [Kowalska 2007, s. 256]), polskie banki (PKO Bank Polski, Bank Gospodarski Żywnościowej, Bank Millenium S.A., Bank Przemysłowo-Handlowy, Bank Zachodni WBK, City Bank) oraz instytucje działające na rzecz rozwoju kultury i nauki (towarzystwa naukowe i szkoły wyższe). Niestety, zwyczaj sponsorowania cyfryzacji jednego, konkretnego dzieła przez osobę prywatną należy nadal do rzadkości Znacznie częściej spotykaną praktyką jest dotowanie określonych obszarów digitalizacji (np. zakupu sprzętu i oprogramowania, procesów skanowania, mikrofilmowania i elektronicznej archiwizacji) czy przeznaczanie funduszy na konwersję cyfrową określonych kompleksów tematycznych (np. danego tytułu wydawnictwa ciągłego).

*

Ochrona dziedzictwa kulturowego w postaci cyfrowej to bez wątpienia przedsięwzięcie kosztowne. Według prognoz Komisji Europejskiej digitalizacja wszystkich zbiorów europejskich muzeów, archiwów i bibliotek wymagać będzie nakładów finansowych w wysokości około 100 miliardów euro (zbiorów samych bibliotek ponad 19 mld) [Niggermann; Decker; Levy 2011]. Całkowite koszty polskiej cyfryzacji do roku 2020 wyniosą z kolei blisko 3 miliardy zł (ok. 720 mln euro), z czego ponad 2 miliardy (ok. 594 mln euro) pokryje budżet państwa (por. tab. 2, znajdująca się w aneksie). Mimo, że wyliczenia kosztów – zarówno europejskiej, jak i polskiej digitalizacji, opierają się na szeregu hipotez i ekstrapolacji, wskazują one prawdopodobny rząd wielkości wymaganych środków. Niestety, w wielu krajach, w tym także w Polsce, potrzeby związane z systematycznym finansowaniem procesu cyfryzacji nie są jeszcze dostatecznie odzwierciedlone w budżetach instytucji kultury. Wprawdzie z pomocą przychodzą tu alternatywne źródła finansowania (dotacje celowe, środki unijne, donacje, sponsoring, partnerstwa publiczno-prywatne, reklama, usługi digitalizacji na żądanie[13]), ale nawet ich pozyskanie nie zawsze jest gwarantem ciągłości i trwałości projektów digitalizacyjnych.

Z całą pewnością cyfryzacji dorobku piśmienniczego nie należy rozpatrywać tylko w kategoriach kosztów – bo to przecież proces bez ściśle rynkowej orientacji. Jeśli jednak

cyjnych. Wsparła ona finansowo m.in. digitalizację zbiorów Biblioteki Uniwersyteckiej we Wrocławiu, Biblioteki Centralnej Polskiego Związku Niewidomych w Warszawie, Centralnej Biblioteki Rolniczej w Warszawie, Wojewódzkiej i Miejskiej Biblioteki Publicznej w Zielonej Górze, Wojewódzkiej i Miejskiej Biblioteki Publicznej w Łodzi i Miejskiej Biblioteki Publicznej w Tarnowie. W roku 2010 na realizację projektów związanych z ochroną dziedzictwa kulturowego, w tym digitalizację, fundacja przekazała kwotę w wysokości 361 840 zł. [Malitka-Górska 2010].

[13] W przypadku bibliotek posiadających własne pracownie digitalizacji zbiorów dodatkowym źródłem finansowania cyfryzacji są środki pochodzące z opłat za usługi reprograficzne. Ich ceny w zasadzie nie odbiegają od tych proponowanych przez firmy komercyjne. W wielu bibliotekach za skanowanie obiektów powyżej formatu A2 oraz pozycji pochodzących ze zbiorów specjalnych wysokość opłat uzgadnia są jednak odrębnie dla każdego obiektu. Ponadto w wielu przypadkach za digitalizację materiałów powierzonych pobiera się kwoty o około 25% wyższe. Przykładowe cenniki usług reprograficznych zob. *Biblioteka Uniwersytecka – Opłaty* 2011; *Załącznik do Zarządzenia nr 13 Rektora UMK* 2011.

w przyszłości mają z niego wyrastać najważniejsze elementy wzrostu gospodarczego: nowe miejsca pracy, wysokie kwalifikacje, kreatywność i zdolność do współpracy, to zapewnienie długofalowych i stabilnych źródeł finansowania musi stać się przedmiotem szczególnej troski. Kultura – a tak postrzegać należy również zasoby cyfrowe – tak długo jest bowiem stymulatorem rozwoju społeczno-ekonomicznego, jak długo jest zadaniem wszystkich[14].

Aneks

Tabela 1. Szacunkowe koszty elementów składowych procesu digitalizacji[15]

Rodzaj inwestycji	Zadania szczegółowe	Przykładowe koszty
zakup sprzętu i wdrażanie nowych technologii	zakup skanerów	– skanery płaskie: 200 zł-10 tys. zł/szt., – skanery planetarne: 100 zł-350 tys. zł/szt., – skanery przelotowe: od 1 tys. do kilkudziesięciu tys. zł/szt., – skanery bębnowe: od kilku do kilkudziesięciu tys. zł/szt., – skanery do mikroform: 140 zł-300 tys. zł (automatyczne) i 50 zł-100 tys. zł (z ręczną obsługą)/szt., – skanery automatyczne: kilkaset tys. zł/szt., – urządzenia hybrydowe: od 500 tys. zł/szt.
	zakup aparatów fotograficznych	– aparaty cyfrowe: od 1 tys. do kilkuset tys. zł/szt., – aparaty skanujące: od kilku do kilkudziesięciu tys. dolarów/szt.
	zakup urządzeń wspomagających digitalizację	– specjalistyczne stoły reprodukcyjne: 1-6. tys. euro/szt., – specjalistyczne zestawy oświetleniowe: 300 zł-800 zł/zestaw, – specjalistyczne statywy i podesty fotograficzne: 200 zł-1 300 zł/szt.,

[14] W Polsce uruchomienie programu *Kultura+* stało się pierwszym wymiernym krokiem w kierunku realizacji idei przeznaczania na kulturę 1% z budżetu państwa. Kolejnym było podpisanie dnia 14 maja 2011 r. przez Prezesa Rady Ministrów Donalda Tuska oraz przedstawicieli ruchu społecznego Obywatele Kultury *Paktu dla Kultury*, w którym znalazły się postulaty podwyższenia finansowania kultury z budżetu państwa do co najmniej 1% wszystkich wydatków budżetowych (poczynając od roku 2012), zapewnienia równego dostępu do kultury wszystkim obywatelom, wypracowania systemu umożliwiającego przeznaczanie 1% podatku dochodowego od osób prawnych (CIT) na wsparcie przedsięwzięć kulturalnych oraz wprowadzenia nowoczesnej ustawy medialnej. Więcej: *Pakt dla Kultury* 2010.

[15] Stan na koniec czerwca 20011 r. Poszczególne koszty oszacowano na podstawie korespondencji z producentami specjalistycznego sprzętu i oprogramowania, rozmów z pracownikami bibliotek, danych zamieszczonych na stronach WWW firm komercyjnych oraz informacji zamieszczonych w publikacji *Digitalizacja piśmiennictwa* w oprac. i red. Dariusza Paradowskiego [2010]. Jeśli nie zaznaczono inaczej, podana cena jest ceną brutto.

zakup sprzętu i wdrażanie nowych technologii	zakup komputerów, monitorów	– komputer do obróbki digitalizatów np. INTEL PENTIUM DUAL CORE E5400, 4096 MB, 500 GB, DVDRW: od 1 tys. zł/zestaw, – **serwer do archiwizacji plików np.** HP ProLiant ML110 G6 G6950 1P 2GB-U z dyskami SATA 160 GB LFF: od 2 tys. zł/zestaw, – monitor np. Eizo FlexScan S2231W: od 2 700 zł/szt.

	zakup oprogramowania	– system operacyjny:
		  – Microsoft Windows 7: 300-600 zł,
		  – OpenSolaris firmy Sun Microsystems: darmowy,
		  – Linux: darmowy.
		– oprogramowanie do skanerów i obróbki plików cyfrowych:
		  – SilverFast: 35-400 euro,
		  – VueScan: 40-80 dolarów,
		  – Adobe Photoshop: 40-2.200 euro,
		  – Acrobat X Pro: 280-700 euro,
		  – Creative Suite:170-3.500 euro,
		  – Sane: darmowe,
		  – Xsane: darmowe,
		  – GIMP: darmowe,
		  – IfranView: darmowe (do użytku non-profit).
		– oprogramowanie do budowy repozytorium:
		– DigiTool:
		  – koszt systemu zależy od wielkości projektu, tj. liczby obiektów, które mają być składowane w systemie, liczby pracowników zarządzających systemem, liczby jednoczesnych dostępów do interfejsu użytkownika oraz modułów dodatkowych,
		  – koszt przykładowej biblioteki cyfrowej do 15 tys. obiektów, dla 3 osób jednocześnie pracujących w systemie i dla 10 osób równocześnie korzystających z interfejsu użytkownika, z modułem zarządzania kolekcjami, z wdrożeniem i szkoleniami wynosi około 180 tys. zł netto [Dziubecki 2011].
		– dLibra:
		  – podstawowe wdrożenie: 1 200 zł netto za zakup licencji wraz z instalacją, szkoleniami redaktorów, pomocą techniczną, migracją oprogramowania do nowych wersji,
		  – pełne wdrożenie: 1 200 zł netto za zakup licencji + zdalna instalacja 1 100 zł netto + szkolenie administratorów systemu na miejscu i w siedzibie PCSS 4 800 zł netto + szkolenie redaktorów 4 680 zł netto [dlibra-cennik 2011].
		– Dspace: darmowe,
		– Fedora: darmowe,
		– Greenstone: darmowe.

szkolenie i zatrudnienie pracowników	udział w szkoleniach, konferencjach, warsztatach, itp.	– 2-dniowe szkolenie (14 ha) przygotowujące do realizacji projektu digitalizacyjnego organizowane przez Bibliotekę Narodową: 330 zł/os., – 1-dniowa konferencja Centrum Promocji Informatyki nt. bibliotek cyfrowych lub digitalizacji: 580 zł netto/os., – szkolenia i konferencje Poznańskiego Centrum Superkomputerowo-Sieciowego: – 1-dniowe szkolenie z tematyki bibliotek cyfrowych: 250 zł/os., – 1-dniowe warsztaty „Biblioteki cyfrowe": 200 zł/os., – 2-dniowa konferencja „Polskie Biblioteki Cyfrowe": 400 zł/os., – 2-dniowe szkolenie redaktorów biblioteki cyfrowej na miejscu u licencjobiorcy: 4 680 zł netto/grupę szkoleniową (kilkanaście osób), – 1-dniowe szkolenie administratora systemu dLibra w siedzibie PCSS: 1 900 zł netto/2-4 os., – 1-dniowe szkolenie administratora systemu dLibra na miejscu u licencjobiorcy: 2 900 zł netto/2-4 os. [Werla 2011], – 1-dniowe (4-5 ha) szkolenie z zakresu prawa autorskiego w środowisku cyfrowym, realizowane przez firmę prywatną: 1 700 zł/os., – pozostałe konferencje krajowe: 250-450 zł/os.
	etaty pracownicze (selekcjoner, kataloger, skanerzysta, korektor, redaktor cyfrowy)	od 1 200 do 1 800 zł/etat
konsulting prawny	wykonanie analizy/ekspertyzy prawnej	– stawka godzinowa: 120-250 zł/ha, – stawka godzinowa z reprezentacją (pełnomocnictwo): 240-420 zł/ha, – prowadzenie sprawy w postępowaniu sądowym, administracyjnym lub podatkowym: według stawek ustawowych - min. 1 200 zł lub stawek godzinowych, – stała obsługa prawna: według indywidualnie uzgadnianych stawek godzinowych. Stawki minimalne w sprawach cywilnych, ze stosunku pracy i ubezpieczeń społecznych wynoszą przy wartości przedmiotu sprawy: 1) do 500 zł – 60 zł, 2) powyżej 500 zł do 1 500 zł – 180 zł, 3) powyżej 1 500 zł do 5 000 zł – 600 zł, 4) powyżej 5 000 zł do 10 000 zł – 1 200 zł, 5) powyżej 10 000 zł do 50 000 zł – 2 400 zł, 6) powyżej 50 000 zł do 200 000 zł – 3 600 zł, 7) powyżej 200 000 zł – 7 200 zł [Rozporządzenie Ministra Sprawiedliwości 2002]. Stawka minimalna w sprawach o ochronę dóbr osobistych i ochronę praw autorskich wynosi 360 zł.

pozyskiwanie praw autorskich	nabycie licencji autorskiej	od kilkuset nawet do kilkudziesięciu tys. zł (np. za zgodę na wykorzystanie fragmentu mapy jpeg wielkości znaczka pocztowego z jednego z serwisów szukacz.pl firma żąda 700 zł + VAT), przy sprzedaży na wyłączność majątkowych praw autorskich niektórzy twórcy stosują zasadę: cena razy 50, przy sprzedaży niewyłącznej majątkowych praw autorskich najczęściej cena pomnożona od 5 do 10 razy.
archiwizacja i konserwacja	archiwizacja danych cyfrowych	– zakup nośników: – płyty CD np. CD-R 700MB 52x: od 18 zł/25 szt., – płyty DVD np. DVD+R 4,7GB 16x: od 30 zł/50 szt., – pamięci flash: od 150 zł za 32 GB. – utrzymanie serwera: – dedykowanego (np. 8 GB DDR3 RAM, 2 x 750 GB SATA-II HDD): od 300 zł netto wzwyż/msc. dla transferu 5 tys. GB/msc. [Serwery dedykowane 2011a; 2011b], – własnego: koszty amortyzacji sprzętu, prądu, dostępu do sieci Internet, chłodzenia itd. + pensja dla informatyka-administratora (4-5 tys. zł netto), koszty licencji na oprogramowanie i pomocy technicznej. – zdalna archiwizacja plików: – utrzymanie archiwum online w serwisie internetowym polskiej firmy: od 100 zł netto/rok (szacunkowa cena za archiwum składające się z ok. 1 tys. dokumentów, dostęp 24h/dobę) [Cennik 2011], – Amazon Simple Storage: – do 1 TB/msc.: 0,140 dolara/GB, – każde kolejne 49 TB/msc.: 0,125 dolara/GB, – import danych: 0.100 dolara/GB, – eksport danych: – do 10 TB/msc.: 0,150/GB, – powyżej 40 TB/msc.: 0,110/GB, – więcej niż 150 TB/msc.: 0,080 dolara/GB [Cennik 2011].
	usługi informatyczne	– naprawa/modernizacja sprzętu: od 50 zł netto/ha, – analizy informatyczne/raportowanie: od 75 zł netto/ha, – dLibra: – e-mailowa pomoc techniczna dla administratorów systemu biblioteki cyfrowej przez okres 1 roku wraz z prawem do nowych wersji oprogramowania dLibra udostępnionych w okresie objętym pomocą techniczną: 3 120 zł netto/rok, – telefoniczna pomoc techniczna: 2 tys./10 dni, – zdalna migracja oprogramowania dLibra z wersji 4.0.24 do wersji 5: 1 600 zł netto [dLibra – cennik 2011].
budowanie sieci współpracy	konferencje, warsztaty	cena uzależniona m.in. od miejsca i czasu trwania spotkania: – Europa: od 70 euro/os., – Stany Zjednoczone: od 60 dolarów/os.

promocja zasobów cyfrowych	reklama	– jeden billboard w centrum miasta (3m x 6m): – wynajem: od 300 zł/msc., – sprzedaż: od 5 tys., – standardowa reklama w prasie (strony ogłoszeń): – od 120 zł za 1 moduł (46,8 x 57,5 mm) czarno-biały, – od 200 zł za moduł w kolorze, – mailing do 1000 osób: – w serwisach aukcyjnych: 10-50 zł, – w wyspecjalizowanych agencjach: 50-200 zł, – baner w popularnym portalu (1000 odsłon): od 15 zł netto za reklamę w wymiarze 468 x 60 pikseli, wielkości do 15 KB, w formacie jpeg, gif, flash lub JavaScript.

Źródło: Opracowanie własne.

Tabela 2. Koszty realizacji polskiego programu digitalizacji na lata 2009-2020

Zadanie	Całkowite koszty w mln zł	W tym:		
		środki z budżetu państwa	środki UE	środki z innych źródeł finansowania
Digitalizacja zasobów audialnych i audiowizualnych	680	520	60	100
Digitalizacja materiałów drukowanych	420	370	50	0
Działalność Narodowego Instytutu Audiowizualnego	398	358	20	20
Przechowywanie polskich zasobów cyfrowych	300	250	0	50
Udostępnianie polskich zasobów cyfrowych	272	212	40	20
Budowa pracowni digitalizacyjnych	238	138	50	50
Budowa repozytoriów cyfrowych	230	180	50	0
Organizacja Krajowych Centów Kompetencji	155	150	5	0
Rejestracja polskich zasobów sieciowych	100	90	0	10
Edukacja	90	70	10	10
SUMA	**2883**	**2338**	**285**	**260**

Źródło: Ministerstwo Kultury i Dziedzictwa Narodowego (2009).

Bibliografia:

1. *A Framework of Guidance for Building Good Digital Collections – the 3rd edition* (2007) [dok. elektr.] Institute for Museum and Library Services http://framework.niso.org/ [odczyt: 27.06.2011].
2. *Aktion Buchpatenschaft* (2007-2011). [dok. elektr.] Österreichische Nationalbibliothek http://www.onb.ac.at/about/buchpatenschaften.htm [odczyt: 27.06.2011].
3. Amazon S3 Pricing (2011). [dok. elektr.] *Amazon Simple Storage Service (Amazon S3)* http://aws.amazon.com/s3/#pricing [odczyt: 27.06.2011].
4. Bednarek-Michalska, Bożena (2010). *Pozyskiwanie finansowania na budowanie cyfrowych zasobów.* [W:] Hollender, Henryk red. (2010). *Cyfrowy świat dokumentu: wydawnictwa, biblioteki, muzea, archiwa.* Warszawa: Centrum Promocji Informatyki, s. 238-262.
5. *Biblioteka Uniwersytecka – Opłaty* (2011). [dok. elektr.] Biblioteka Uniwersytecka w Poznaniu http://lib.amu.edu.pl/index.php?option=com_kontent&task=view&id=533&Itemid=79#4 [odczyt: 27.06.2011].
6. Buczek, J.M. (2011). Fwd: *Cennik usług* [dok. elektr.] Do: M. Kowalska. 13 czerwca 2011, 16:07. Korespondencja osobista [odczyt: 27.06.2011].
7. Burska, Joanna red. (2004). *Biblioteki publiczne w erze cyfrowej: poradnik Pulmana.* Warszawa, s. 166-168 [Wersja elektroniczna: http://www.pulmanweb.org/countries/Poland.htm].
8. *Cennik* (2011). [dok. elektr.] *Skanowanie dokumentów (NeoSITech)* http://skanowanie.info/cennik [odczyt: 27.06.2011].
9. *dLibra-cennik aktualny* (2011). [dok. elektr.] Poznań: Poznańskie Centrum Superkomputerowo-Sieciowe, 1 s. http://dlibra.psnc.pl/docs/dlibra-cennik-aktualny.pdf [odczyt: 27.06.2011].
10. Dziubecki, Maciej (2011). Fwd: *Koszt implementacji DigiTool* [dok. elektr.] Do: M. Kowalska. 10 czerwca 2011, 08:56. Korespondencja osobista [odczyt: 27.06.2011].
11. Gębołyś, Zdzisław (2009). *Sponsoring i fundraising biblioteczny – między teorią a praktyką* [dok. elektr.] „Biuletyn EBIB" nr 8 (108) http://www.ebib.info/2009/108/a.php?gebolys [odczyt: 27.06.2011].
12. Gordon, Christopher (2011). *Sponsoring i alternatywne finansowanie kultury* [dok. elektr.] *LabforCulture* http://www.labforculture.org/pl/home/zawarto%C5%9B%C4%87/floating-pages/articles/sponsoring-i-alternatywne-finansowanie-kultury [odczyt: 27.06.2011].
13. Komarnicki, M. (2011). Fwd: *Cennik usług* [dok. elektr.] Do: M. Kowalska. 8 czerwca 2011, 12:38. Korespondencja osobista [odczyt: 27.06.2011].
14. Kowalska, Małgorzata (2007). *Dygitalizacja zbiorów bibliotek polskich.* Warszawa: Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich, 293, [2] s.
15. *Lista sponsorów* (2008-2011). [dok. elektr.] Miejska Biblioteka Publiczna im. J. Słowackiego w Tarnowie http://www.biblioteka.tarnow.pl/index.php?plik=sponsor [odczyt: 27.06.2011].
16. Malitka-Górska, D. (2010). Fwd: *Kwota dotacji za rok 2010* [online]. Do: M. Kowalska. 8 czerwca 2011, 17:08 [odczyt: 27.06.2011]. Korespondencja osobista.
17. Ministerstwo Kultury i Dziedzictwa Narodowego (2009). *Program digitalizacji dóbr kultury oraz gromadzenia, przechowywania i udostępniania obiektów cyfrowych w Polsce na lata 2009-2020* (2009) [dok. elektr.] Warszawa: Ministerstwo Kultury i Dziedzictwa Na-

rodowego, 67 s. http://www.kongreskultury.pl/library/File/RaportDigitalizacja/Program digitalizacji 2009-2020.pdf [odczyt: 27.06.2011].

18. Ministerstwo Kultury i Dziedzictwa Narodowego (2011a). *Programy Ministra* [dok. elektr.] http://www.mkidn.gov.pl/pages/strona-glowna/finanse/programy-ministra/programy-mkidn-2011.php [odczyt: 27.06.2011];
19. Ministerstwo Kultury i Dziedzictwa Narodowego (2011b). *Program Wieloletni KULTURA+* [dok. elektr.] http://www.mkidn.gov.pl/pages/strona-glowna/finanse/program-wieloletni-kultura.php [odczyt: 27.06.2011].
20. Ministerstwo Kultury i Dziedzictwa Narodowego (2011c). *Program Wieloletni KULTURA+. Priorytet: „Digitalizacja"* [dok. elektr.] Narodowy Instytut Audiowizualny http://www.nina.gov.pl/files/Regulamin_WPR_ Kultura_Priorytet_Digitalizacja.pdf [odczyt: 27.06.2011].
21. Niggermann, Elisabeth; Decker, Jacques de.; Levy, Maurice (2011). *Nowy renesans. Sprawozdanie „Comité des Sages"* [dok. elektr.] Bruksela: Komisja Europejska, Comité des Sages, 49 s., http://ec.europa.eu/information_society/activities/digital_libraries/doc/executivesummery/ final_renaissace_pl.pdf [odczyt: 27.06.2011].
22. *O programach regionalnych* (2008). [dok. elektr.] Portal Funduszy Europejskich http://www.funduszeeuropejskie.gov.pl/RPO/Strony/O_Programach_Regionalnych.aspx [odczyt: 27.06.2011].
23. *Outsourcing marketingu* (2007). [dok. elektr.] *Encyklopedia Zarządzania* http://mfiles.pl/pl/index.php/Outsourcing_marketingu [odczyt: 27.06.2011].
24. *Pakt dla Kultury* (2010). [dok. elektr.] Obywatele Kultury http://obywatelekultury.pl/2010/12/pakt-dla-kultury-tekst-jednolity/ [odczyt: 27.06.2011].
25. Paradowski, Dariusz oprac. i red. (2010). *Digitalizacja piśmiennictwa.* [Aut. Maciej Hnydka i in.; red. merytoryczna Agnieszka Konopka]. Warszawa: Biblioteka Narodowa, 212, [1] s.: il.
26. Potrzebnicka, Ewa (2005). *Digitalizacja nowoczesną formą tworzenia kopii dokumentów bibliotecznych.* [W:] Drewniewska-Idziak, Barbara red. (2005). *Trwałe zbiory – źródłem historii.* Warszawa: Biblioteka Narodowa, s. 66-77.
27. Potrzebnicka, Ewa (2006). *Stan dygitalizacji w polskich bibliotekach.* [W:] Stefańczyk, Elżbieta red. (2006). *Dygitalizacja zbiorów bibliotecznych: materiały z ogólnopolskiej konferencji pt. Digitalizacja zbiorów bibliotecznych. Warszawa 3-4 czerwca 2005 r.* Warszawa: Wydawnictwo SBP, s. 16-21;
28. *Program digitalizacji dóbr kultury oraz gromadzenia, przechowywania i udostępniania obiektów cyfrowych w Polsce na lata 2009-2020* (2009) [dok. elektr.] Warszawa: Ministerstwo Kultury i Dziedzictwa Narodowego, 67 s. http://www.kongreskultury.pl/library/File/Raport Digitalizacja/Program digitalizacji 2009-2020.pdf [odczyt: 27.06.2011].
29. *Program Operacyjny Infrastruktura i Środowisko. Narodowe Strategiczne Ramy Odniesienia 2007–2013* (2007). [dok. elektr.] Warszawa: Ministerstwo Rozwoju Regionalnego, 277 s. http://www.pois.gov.pl/Dokumenty/Lists/Dokumenty programowe/Attachments/91/POIiS0 51208. pdf [odczyt: 27.06.2011].
30. *Program Operacyjny Innowacyjna Gospodarka, 2007-2013. Narodowe Strategiczne Ramy Odniesienia 2007–2013* (2007). [dok. elektr.] Warszawa: Ministerstwo Rozwoju Regionalnego, 192 s. http://www.poig.gov.pl/ Dokumenty/Lists/Dokumenty programowe/Attachments/89/POIG_01102008.pdf [odczyt: 27.06.2011].
31. *Program Operacyjny Kapitał Ludzki. Narodowe Strategiczne Ramy Odniesienia 2007-2013* (2007). Warszawa: Ministerstwo Rozwoju Regionalnego, 280 s. [dok. elektr.] http://

www.efs.gov.pl/Dokumenty/Lists/Dokumenty programowe/Attachments/87/POKL_ zatwierdzony_7092007.pdf [odczyt: 27.06.2011].

32. *Program rozwoju infrastruktury informatycznej nauki na lata 2007–2013* (2007). Warszawa: Ministerstwo Nauki i Szkolnictwa Wyższego, 13 s. [dok. elektr.] http://www.nauka.gov.pl/fileadmin/user_upload/28/81/28816/20070628_Program_rozwoju_infrastruktury _informatycznej_nauki_na_lata_2007_-_2013.pdf [odczyt: 27.06.2011].
33. *Przewodnik programowy. Program Kultura (2007-2013)* (2010). [dok. elektr.] Bruksela: Agencja Wykonawcza ds. Edukacji, Kultury i Sektora Audiowizualnego Komisji Europejskiej, 108 s. http://eacea.ec.europa.eu/culture/programme/documents/2010/may/PL.pdf. [odczyt: 27.06.2011].
34. *Rozporządzenie Ministra Sprawiedliwości z dnia 28 września 2002 r. w sprawie w sprawie opłat za czynności radców prawnych oraz ponoszenia przez Skarb Państwa kosztów pomocy prawnej udzielonej przez radcę prawnego ustanowionego z urzędu* (2002). Dziennik Ustaw nr 164, poz. 1349.
35. *Serwery dedykowane* (2011a). [dok. elektr.] *Serveradmin.pl* http://serveradmin.pl/Serwery-dedykowane_f_6 [odczyt: 27.06.2011].
36. *Serwery dedykowane* (2011b). [dok. elektr.] *home.pl* http://www.home.pl/serwery/dedykowane [odczyt: 27.06.2011].
37. *Sponsorzy* (2009). [dok. elektr.] Biblioteka Uniwersytecka w Poznaniu http://lib.amu.edu.pl/index.php?option=com_content&task=view&id=7&Itemid=10 [odczyt: 27.06.2011].
38. Vollmer, Renate (2001). *Fundraising an wissenschaftlichen Bibliotheken*. „Bibliothek" nr. 2, s. 192–213.
39. Werla, Marcin (2011). Fwd: *Koszty w digitalizacji* [dok. elektr.] Do: M. Kowalska. 21 czerwca 2011, 17:35. Korespondencja osobista [odczyt: 27.06.2011].
40. *Załącznik do Zarządzenia nr 13 Rektora UMK z dnia 3 lutego 2011 r. w sprawie wprowadzenia cennika usług i opłat Biblioteki Uniwersyteckiej w Toruniu* (2011). Biuletyn Prawny UMK [dok. elektr.] (akty oczekujące na opublikowanie) http://www.umk.pl/uczelnia/ dokumenty/biuletyn/prawo/inc/zalaczniki/Z_Rektora_13_2011_z1.pdf [odczyt: 27.06.2011].

Zagadnienia estetyczne, a biblioteki cyfrowe

Małgorzata Janiak

Wstęp

W pracy przedstawiono rozważania na temat powiązań interdyscyplinarnych pomiędzy zagadnieniami związanymi z estetyką i z bibliotekami cyfrowymi. Celem pracy jest ukazanie różnorodności interakcji pomiędzy tymi dwoma obszarami tzn. określenie elementów estetycznych odnoszących się do systemu oraz zawartości bibliotek cyfrowych (np. w związku z ucyfrowianiem dzieł sztuki), zagadnień estetycznych branych pod uwagę w trakcie ewaluacji bibliotek cyfrowych oraz zaznaczenie, iż struktura bazodanowa tworząca biblioteki cyfrowe może także pomagać tworzyć (czasem na nowo) dzieła artystyczne. Biblioteki cyfrowe, jako systemy informacyjne o konkretnych strukturach są bowiem bardzo odpowiednie dla przechowywania tzw. informacji artystycznej. Poza tym bazy danych w jakiś sposób mogą charakteryzować współczesną kulturę, zwłaszcza jeśli w centrum zainteresowań badaczy oraz odbiorców znajdą się nowe media (systemy informacyjne z obiektu same stać się mogą kryterium) – zob. prace Lva Manovica [2006].

Na początek przedstawić należy zakresy kilku podstawowych, używanych w pracy terminów. Estetyka: *[gr. aisthētikē – od aísthēsis 'wrażenie zmysłowe (nauka o wrażeniach)'] rozumiana jest jako dyscyplina filozoficzna – nauka o pięknie, jego kryteriach oraz o przeżyciach z nim związanych.* Natomiast wartości estetyczne, o których będzie jeszcze mowa, to *powszechnie akceptowane sądy egzystencjonalno-normatywne – orientacje wartościujące poszczególne obiekty poprzez piękno* [czasem różnorodnie rozumiane] [Estetyka cop. 1997-2011]. Trzeba jednak zaznaczyć, iż są to definicje adekwatne dla estetyki klasycznej, gdyż dla współczesnej sztuki i jej teorii piękno nie jest głównym kryterium wyróżniającym sztukę ze wszystkich innych dzieł ludzkiej kreatywności. Główny akcent w najnowszej sztuce położony jest bowiem na sam akt kreacji i decy-

zję artysty, które z jego dzieł są dziełami sztuki. Tak więc czasem element zwyczajności czy odpowiedź na zapotrzebowanie społeczne są charakterystyczne dla wielu nurtów. *Sztuka stała się konstrukcją psychospołeczną, definiowaną przez swoją tożsamość instytucjonalną, wartość rozrywkową i komercyjny rozmach* [Kuspit 2004, s. 30].

Natomiast biblioteka cyfrowa rozumiana jest jak w manifeście z 2011 r. *The Digital Library Reference Model* [DL.org: Coordination Action on Digital Library Interoperability, Best Practices and Modelling Foundations w ramach Network of Excellence on Digital Libraries], tzn. jako: *potencjalna wirtualna organizacja, która w sposób kompleksowy gromadzi, zarządza i zachowuje na długi czas bogate treści cyfrowe oraz oferuje swoim docelowym społecznościom użytkowników wyspecjalizowane funkcje (umożliwiające także personalizację systemu), określoną jakość oraz zgodność z polityką tworzenia całego systemu* [Candela; Nardi 2011].

Kolejny ważny termin to: informacja artystyczna. Nie pojawia się on zbyt często w piśmiennictwie. Jednym z nielicznych cytatów, które można tu przytoczyć jest tekst Witolda Dederko, fotografa, aktora, teoretyka fotografii i pedagoga. Napisał on, iż *informacja artystyczna w przeciwieństwie do informacji realnej działa* [także, a może przede wszystkim – uzup. autorki] *na uczucia odbiorcy. Sztuka ma ogromne aspiracje, ma pobudzać człowieka do przemyśleń, a potem ewentualnych działań w myśl intencji artysty. Informacja przez sztukę, mocniej i głębiej działa na odbiorcę od informacji suchej* [Dederko 1985].

W polskiej literaturze z zakresu informacji naukowej brak jest szczegółowych rozważań na temat informacyjnych aspektów sztuki, czy interdyscyplinarnych powiązań pomiędzy estetyką a informacją. Nieliczne prace dotyczące informacyjnych walorów ilustracji oraz publikacji elektronicznych są głównie autorstwa Małgorzaty Komzy [2002, 2004, 2009]. Inny aspekt, który można by było zaliczyć do tematu: informacja artystyczna, poruszają prace Krystyny Bednarskiej-Ruszajowej [1998, 2003, 2006], mówiące o wizji biblioteki w dziełach literackich (nie pojawia się tu jednak biblioteka cyfrowa). Pozostałe prace, już z zakresu szeroko rozumianej bibliologii, odnoszą się do ilustracji dzieł, opraw itp., a więc do zasobów tworzących także biblioteki cyfrowe. Prace z pogranicza design'u i informacji odnoszą aspekt informacyjny funkcjonujący w dziełach sztuki do reklam, plansz informacyjnych np. w przestrzeni miasta itp. Dodatkowo estetyka pojawia się w pracach o architekturze informacji, które jednak najbardziej skupione są na funkcjonalności systemów. Natomiast w wielu podręcznikach dotyczących teorii sztuki, czy estetyki podkreślany jest informacyjny charakter dzieła sztuki [np. Gutowska 2006]. W większości prac brak jednak odwołań do bibliotek cyfrowych [zob. Janiak 2009a, 2009b].

Natomiast w języku angielskim funkcjonuje termin: 'information art' (lub 'informatism'). Termin ten został ukuty dla powstającego pola sztuki elektronicznej, syntetyzującej informatykę, technologię informatyczną z bardziej klasycznymi formami sztuki, włączając w to sztukę konceptualną, sztukę nowych mediów, sztukę wizualną czy performance. 'Information art' związane jest oczywiście z interakcją z komputerem, gdyż zawartość dzieł sztuki w dużym stopniu bazuje na wielkiej ilości przetwa-

rzanych danych [Wilson 2001]. Z drugiej strony 'informatism' bywa rozumiany jako informacja o dziełach sztuki.

1. Estetyka systemów informacyjnych – tworzenie i wykorzystywanie systemów

Przedstawienie zagadnień związanych z estetyką rozpocząć należy od estetyki systemów informacyjnych, jako zagadnienia szerszego oraz częściej analizowanego niż estetyka bibliotek cyfrowych. Omówione zostaną tutaj tylko niektóre prace w układzie chronologicznym, wybrane ze względu na różnorodność ich tematyki. Wybór ten ma też wskazać zakresy podejmowanych analiz.

Jedną z pierwszych prac, w których tytule pojawia się estetyka systemów informacyjnych był artykuł z 1994 r. Erika Stoltermana (*The aesthetics of information systems*), informatyka zajmującego się teorią i projektowaniem systemów interakcyjnych (interakcje: człowiek-komputer). Odwołał się on do książki Rafaela Ramireza z 1991 r. *The beauty of social organization*, w której zaznaczono, iż trzeba wypełnić teoretyczną dziurę związaną z brakiem badań nad tym, co piękne w organizacji, nad odbiorem estetycznym i intuicyjnym informacji. Stolterman przedstawił pokrótce projekt badań w zakresie estetyki systemów informacyjnych, w których to powinno się przeanalizować: 1) jak ludzie doświadczają systemów informatycznych – w aspekcie odbioru, doświadczenia, piękna oraz ogólnego charakteru systemu (doświadczenie estetyczne odbiorcy), 2) jak systemy są ocenianie, 3) czy istnieją relacje między cechami systemu informacyjnego oraz pewnych szczególnych doświadczeń odbiorców, a także określić 4) ramy tych doświadczeń.

Oczywiście odnieść należy się do badanych ludzi (w tym też projektantów, którzy w pewnym momencie tracą 'panowanie' nad systemem – użytkownicy odbierają go często bez analizy zamierzeń designerskich) w różnych kontekstach społecznych, w odniesieniu do doświadczeń także estetycznych. Przydatne stają się w tych analizach dwa pojęcia: styl (w tym relacja między elementami a całością) i dobry przykład (np. lista informacji, przykład ogólnego charakteru, relacji między szczegółami a całością oraz między funkcją a formą).

W kolejnych swoich wypowiedziach Erik Stolterman także odwoływał się do projektowania, podkreślając ważne dla niego cechy, skojarzone z estetyką, a mianowicie: kreatywność, element zaskoczenia, innowacyjność, zaznaczając, iż design to nie sztuka i nie nauka – ma swoją własna tradycję, zwłaszcza wykorzystywania już przetartych ścieżek, jak i podejścia do zmieniania świata [Stolterman 2008].

W 1997 r. pojawił się następny artykuł z estetyką w tytule: *Aesthetics and apparent usability: empirically assessing cultural and methodological issues* autorstwa Noama Tractinsky'ego. Przedstawił on m.in. eksperymenty Masaaki'ego Kurosu i Kaori Kasimura'ego [1995], którzy badali funkcjonalność interfejsów wśród użytkowników z różnych kręgów kulturowych (Izrael, Japonia) oraz odniósł ich wyniki do aspektów estetycznych. Zaakcentował wysoką korelację (różną w tych dwóch krajach) pomię-

dzy oceną estetyki interfejsu oraz łatwości obsługi systemu o także wskazał, iż różnice międzykulturowe widoczne są w wartościach korelacji.

Odwołania do estetyki w badaniach systemów informacyjnych znalazły się także w pracy *Assessing dimensions of perceived visual aesthetics of web sites* Talii Laviea i wymienionego wcześniej Noama Tractinsky'ego z 2004 r. Zaznaczyli oni, iż estetyka interfejsów jest istotnym czynnikiem wpływającym na zadowolenie użytkowników i ich odczucia przyjemności, związane z wykorzystywaniem systemów informacyjnych. Przy czym brak jest wypracowanych ogólnie koncepcji oraz technik badawczych, które umożliwiłyby szersze badania w tej dziedzinie. Stąd autorzy przeprowadzili badania w celu opracowania instrumentów do pomiaru postrzegania estetyki stron www. Wykazali oni, iż użytkownicy odnoszą się do 2 wymiarów, które nazywane zostały przez nich: estetyką klasyczną i ekspresyjną. Wymiar klasyczny podkreśla uporządkowanie i przejrzysty wygląd, co jest ściśle związane z zasadami projektowania zalecanymi przez ekspertów design'u i użyteczności. Ekspresyjny wymiar przejawia się w kreatywności projektantów, oryginalności oraz zdolności do łamania konwencji projektowania. Wymienili także najczęściej wykorzystywane określniki estetyczne, a mianowicie: kreatywny, artystyczny, nowoczesny, wyrafinowany, efektowny, kolorowy, oryginalny, energetyczny, piękny, wymagający, fascynujący, intrygujący, zaprojektowany z myślą o umiejętnościach użytkowników, profesjonalny, zawierający zdjęcia, wywołujący dobre samopoczucie, estetyczny, jasny, czysty (w formie), przyjemny, miły, zorganizowany, wywołujący podziw, ekscytujący, symetryczny. Uporządkowane zostały one wraz z dodanymi ostatecznie elementami w 5 grup czynników: 1) klasyczna estetyka z wyglądami: estetycznym, przyjemnym, jasnym, czystym (w formie) oraz symetrycznym, 2) ekspresyjna estetyka z designe'm: kreatywnym fascynującym, oryginalnym i wyrafinowanym, a także z używaniem efektów specjalnych, 3) użyteczność: wygoda użycia (dogodność), łatwa orientacja, łatwość użycia, łatwa nawigacja oraz zrozumiały (wyrazisty) design, 4) przyjemność płynąca z interakcji: poczucie radości, przyjemności i gratyfikacji, 5) jakość serwisu: bezbłędność, niezawodność oraz możliwość dokonywania obliczeń na stronach.

W 2006 r. pojawił się artykuł rzeźbiarza i profesora sztuk wizualnych Jacka W. Burnhama *The aesthetics of intelligent systems,* w którym podkreślono inne spojrzenie na dzieła sztuki w systemach informacyjnych. Autor zaznaczył, iż do tej pory, dzieła sztuki były traktowane jako samodzielne i skończone źródła informacji: gdy obiekt sztuki został utworzony, mógł *tylko przekazywać własną obecność*. Obecnie sztuka może przekazywać inny rodzaj informacji, jeśli odejdzie się od jednokierunkowej komunikacji od artysty do odbiorcy. Stąd estetyka inteligentnych systemów musi odnosić się do dialogu, gdzie dwa systemy gromadzenia i wymiany informacji (zarówno komputery, jak i ludzie) wpływają na siebie.

Kolejnym tekstem, o którym należy tu wspomnieć, jest praca Johna D. Haynesa oraz Davida B. Paradice'a *Information technology and aesthetics: passive and active* z 2007 r. Autorzy także odwołali się do teorii przeżycia estetycznego, która powinna wpływać na wszystkie procesy tworzenia systemów informacyjnych – nie tylko samego interfejsu. Estetyka systemu odbierana być może w wymiarach pasywnych

– zmysłowych, ale i aktywnych – w umyśle. Inny podział na pasywne i aktywne to odniesienie do definicji i tematów (bardziej kreatywne działania – skojarzeniowe). Ostateczną tezą było stwierdzenie, iż ludzka zdolność do rozpoznawania i oceniania piękna odnosi się do wymiaru szczęścia (a więc np. zadowolenia użytkownika), jak i skutecznej kreatywności (np. wykorzystywania zdobytych informacji), stąd 'piękny system informacyjny' powinien zachęcać użytkownika do wchodzenia z nim w relacje oraz umożliwiać bardziej twórcze i wręcz wizjonerskie zarządzanie informacją[1].

W tym samym 2007 r. ukazał się także artykuł *Towards a model of information aesthetics in information visualization* Andrei Lau i Andrew Vande Moere'ra, w którym starano się przedstawić propozycję nowego modelu łączącego estetykę informacji z wizualizacją informacji. Model powstał na podstawie coraz większej liczby pojawiających się właśnie projektów wizualizacyjnych, łączących techniki wizualizacji informacji z zasadami kreatywnego wzornictwa (design'u). Autorzy skoncentrowali się na doświadczeniu estetycznym, jako jednym z podstawowych doświadczeń, które powstają w trakcie interakcji z systemem informacyjnym. Połączyli więc estetykę z badaniami użytkowników oraz wykazali artystyczne wpływy na techniczną realizację systemów, wskazując znaczenie estetyki dla procesu wizualizacji, a co za tym idzie jej wpływ na projektantów oraz wpływ 'stylów', w których funkcjonują projektanci na systemy i użytkowników.

*

W polskim piśmiennictwie zagadnienia związane z tworzeniem systemów informacyjnych związane są z terminami: typografia, design, architektura systemu czy wizualizacja informacji. Rzadko pojawia się w tytułach sama estetyka, czy przeżycie estetyczne lub piękno.

Typografia oraz design omówione zostaną przy analizach związanych z zasobami. Natomiast kolejny termin (architektura informacji), pojawia się w piśmiennictwie informatologicznym, informatycznym oraz związanym z wzornictwem, rozumianym jako branża oraz jako dziedzina sztuki (tu także występuje: design).

Architektura systemu to zarówno dyscyplina, jak i sam sposób *uporządkowania i wzajemnego powiązania informacji w systemach zawierających informacje*. Jej celem jest ułatwienie użytkownikowi dostępu do informacji, a następnie jej wykorzystania. Zgodnie z definicją zawartą w książce Petera Morville'a i Lou Rosenfelda wymienia się systemy: 1) organizacji (grupowania), 2) nazewnictwa (etykietowania), 3) nawigacji, 4) wyszukiwania [*Architektura...* dok. elektr.]. Elementy związane z estetyką pojawią się głównie w systemach odnoszących się do struktury i nawigowania. Kryją się one pod takimi cechami, jak np.: logiczność działania, prostota, 'uroda', czy elegancja systemu. Szersza analiza estetycznych aspektów systemów informacyjnych oraz bibliotek cyfrowych przeniesiona została do podrozdziału o ocenie jakości.

[1] Zob. także np. estetyka komplementarna, w której obecne są tendencje subiektywistyczne i obiektywistyczne. Twórcą nurtu jest Roman Ingarden. Kluczowe pojęcie to sytuacja estetyczna, w której zakres wchodzą akty twórcze artysty i odtwórcze miłośnika sztuki oraz dzieło sztuki: jego struktura, sposób istnienia i wartość. Wymienione fenomeny są ujmowane zarówno w swej istocie, jak i we wzajemnych związkach i zależnościach natury teoriopoznawczej, ontologicznej i aksjologicznej.

Wizualizacja jako termin odnoszony jest w Polsce głównie do prac związanych z psychologią oraz marketingiem. Wizualizacja informacji także się pojawia, jednakże przede wszystkim w artykułach odnoszących się do tworzenia grafów, wykresów, czy np. przedstawiania informacji geograficznej itp. (zob. np. *Wizualizacja wyników analizy syntaktycznej* Janusza S. Bienia). Spojrzenie informatologiczne zaprezentowane zostało dopiero w książce *Wizualizacja i wyszukiwanie dokumentów* Veslavy Osińskiej. Autorka opisała w niej sposoby prezentacji danych w *czytelnej i kognitywnej formie graficznej* [SBP 2010], analizując także sposoby ich porządkowania i wyszukiwania.

*

Podsumowując przedstawienie publikacji zachodnich, odnoszących się do systemów informacyjnych i zagadnień estetycznych, można określić zakresy podejmowanych analiz. Są to tematy odnoszące się do: estetyki i projektowania systemów informacyjnych (np. prace Erika Stoltermana), interakcji człowieka z systemem – wejścia w sytuację przeżycia estetycznego (np. rozważania Johna D. Haynesa oraz Davida B. Paradice'a), odbioru dzieł sztuki funkcjonujących lub posiadających swoje kopie w systemach informacyjnych (np. analizy estetyczne Jacka W. Burnhama), analiz samego interfejsu (np. publikacje Talii Laviea i Noama Tractinsky'ego) oraz wizualizacji danych (np. artykuły Andrei Lau i Andrew Vande Moere'ra).

Piśmiennictwo polskie skoncentrowane zostało na tematyce, odnoszącej się przede wszystkim do funkcjonalności systemów, ich interfejsów oraz design'u, odbioru systemu przez użytkowników oraz wizualizacji informacji.

2. Estetyka systemów informacyjnych, w tym bibliotek cyfrowych – ocena jakości

Pewne elementy związane z estetyką odnoszące się do oceny systemów informacyjnych zostały już opracowane. Wielu twórców kryteriów ocen systemów informacyjnych zawarło w swoich analizach elementy odnoszące się do estetyki.

Analizując ogólne kryteria ocen systemów informacyjnych przygotowane np. przez Carolyn Caywood (1995), T. Matthew Ciołka (1994 – popr. 1997; 1996), Jane Alexander i Marshę Tate (1996), Alistaira Smitha (1996), Hope N. Tillman (1996, popr. 2003) i Esther Grassian (1996), Kathy Schrock (2008), czy instytucjonalne takie, jak: Sosig [A list of quality... 2007] (dziś Intute), New Mexico State University Library (autorstwa Susan E. Beck – 2009), UNESCO [ICT for Library... 2001], American Library Association [Great Web Sites – 1997], Cornell University Library (2004), The Johns Hopkins University, The Sheridian Libraries (2010) itp. można stwierdzić, iż dość rzadko pojawia się termin podstawowy dla estetyki, a mianowicie: 'piękno'. W zestawach ocen pojawiają się częściej orzeczniki, synonimy, czy dopełnienia tego terminu. Autorzy wymieniają 'estetyczny wygląd', 'ładny układ typograficzny', 'elegancję' systemu. Cechy estetyczne, które jeszcze są wymieniane to: *logiczność i porządek całego systemu, rozumiane jako porządek semantyczny wraz z czystą, formalną składnią, czystość formy, jasność oraz jednoznaczność struktury, założeń i procedur.*

Do tego dochodzi minimalizm środków, czyli elegancja i dążenie do prostoty [...] oraz *uniwersalność, czyli wykorzystanie zasady korespondencji.* Estetyka odnoszona jest także do cech informacji, przechowywanej w systemach. Wymienić można tu: *spójność (elementy, struktura współgrają ze sobą), odpowiedniość formy (poprawna prezentacja danych), czy kompletność (odpowiednia liczba danych). Do tego dochodzą oczywiste kryteria, takie, jak: zastosowana grafika i multimedia, czy piękno typograficzne.* Podsumowując te kryteria stwierdzić można, iż *idea piękna ukrywa się za: 1) 'estetycznym wyglądem', 'układem typograficznym', 2) 'logicznością działania', 'prostotą', 'elegancją', 'uniwersalnością', 'rozszerzalnością', 'spójnością' oraz 3) 'bogactwem systemu' w elementy graficzne i multimedialne* [Janiak 2009b].

Podobnie przeanalizować można kryteria ocen opracowane dla bibliotek cyfrowych takie, jak: przygotowane przez Tefko Saracevica, a wykorzystywane także w projekcie *DELOS Network of Excellence* (Saracevic, 2000; 2004), *A Framework of Guidance for Building Good Digital Collections – the 3rd edition* (2007), czy opracowania: Ying Zhang, Roberta J. Sandusky'iego (2004) i Hong Iris Xie (2006; 2008).

Cechy estetyczne brane pod uwagę w wymienionych zestawach to po raz kolejny 'rozszerzenia' lub dopełnienia pojęcia piękna. Główne, które należy tutaj wymienić dla całych systemów bibliotek cyfrowych to: porządek, czystość formy, uniwersalność, minimalizm środków, stosowane uogólnienia, zastosowane abstrakcje, jasność oraz jednoznaczność struktury, założeń, procedur; porządek semantyczny; czysta, formalna składnia; wykorzystanie zasady korespondencji; modularność, rozszerzalność i modyfikowalność. Natomiast dla zasobów bibliotek cyfrowych najważniejsze atrybuty to: spójność (elementy, struktura współgrające ze sobą), odpowiedniość formy (poprawna prezentacja danych), czy kompletność (odpowiednia liczba danych) [Janiak 2009a].

3. Estetyka bibliotek cyfrowych – zasoby

Zasoby bibliotek cyfrowych, zarówno pełnotekstowe, jak i cyfrowe kopie konkretnych wydań książek, pocztówek, grafik, czy innych obiektów sztuki także wymagają analiz estetycznych. Najszerzej zostały już omówione same obiekty sztuki, gdyż to 'na nich tworzyła się' cała myśl estetyczna [zob. np. Tatarkiewicz 2009]. Oczywiste jest jednak to, iż nie wszystkie gotowe analizy da się zastosować do cyfrowych obiektów, gdyż nośnik elektroniczny oraz 'odbieralność' dzieła sztuki na ekranie komputera implikuje nowe zagadnienia: znaczenia kopii oraz wpływu jednej 'drogi' odbioru sztuki (poprzez ekran komputera) na tak różnorodne prace: malarskie, rzeźbiarskie, fotograficzne, literackie, czy widowiska itp. W lepszej sytuacji są badacze współczesnej sztuki tworzonej od razu na nośniku elektronicznym – np. projektów wideo, chociaż dla nich trudnym problemem jest oryginalność dzieła, jego unikatowość oraz możliwość tworzenia dużej ilości nierozróżnialnych kopii.

Jeśli chodzi o zasoby książkowe i czasopiśmiennicze, najczęściej jeszcze kojarzone z zasobami bibliotek cyfrowych, to ponownie można stwierdzić, iż estetyczne analizy

oryginalnych dzieł pojawiały się na przestrzeni wieków, co oznacza, że zagadnienie to zostało już w pewien sposób rozpoznane i opracowane [Kuźmina; Tobera 2006]. W tym miejscu podkreślić tylko należy, iż wielu teoretyków bibliologii wskazywało na połączenie w pięknej książce wartości duchowych z materialnymi, 'wnętrza i zewnętrza', formy i treści (Przecław Smolik, Michał Arct, Stanisław Lam, Bonawentura Lenart, Joanna Wiercińska, Karol Głombiowski i inni). Analizując więc estetykę prac tworzących bibliotekę cyfrową odnosić się trzeba także do jej zawartości, treści, znaczenia merytorycznego, naukowego, czy kulturowego.

Sama sztuka książki została natomiast opisana jako połączenie działań edytorskich i typograficznych z elementami estetycznymi, które w efekcie tworzą z książki kolejne dzieło sztuki (Józef Wojakowski, Radosław Cybulski, Piotr Rypson, Dorota Folga-Januszewska i inni). Wynika z tego konkluzja, iż można łączyć analizy estetyczne tworzone na przestrzeni wieków także z pewną zawartością bibliotek cyfrowych.

Kolejne dwa terminy bliskoznaczne: typografia i design najczęściej są kojarzone z zasobami bibliotek cyfrowych ze względu na ich odniesienie do wszystkich dzieł. Książki piękne, artystyczne, literatura totalna – liberatura (Katarzyna Bazarnik, Zenon Fajfer), wydania bibliofilskie odnoszą się bowiem tylko do części zasobów tworzących biblioteki cyfrowe. Typografia przedstawiana jest, poza opisem jej jako techniku druku, przede wszystkim jako dziedzina grafiki użytkowej obejmująca tworzenie szaty graficznej [zob. np. antologię: Kuźmina; Tobera 2006] lub sztuka nadawania ozdobnej szaty, a więc jako sztuka użytkowa [*Typografia* – dok. elektr. Gazeta.pl]. Pojawiają się też w różnych pracach, czy encyklopediach określenia: *ogół zagadnień dotyczących projektowania drukowanych liter i innych znaków pisarskich oraz wzajemnych relacji pomiędzy tymi znakami i grupami znaków, układ graficzny drukowanej strony, sztuka użytkowa zajmująca się estetyką szaty graficznej publikacji* lub *układ graficzny strony na witrynie WWW* [*Typografia* 2011]. Podobnie jest z design'em, który opisywany jest bardziej generalnie, jako wygląd, ogólna sztuka projektowania itp. Analizy typograficzno-design'owskie, akcentujące elementy związane z działalnością artystyczną oraz podkreślające aspekty estetyczne można więc odnosić zarówno, a właściwie przede wszystkim do zawartości, chociaż ich części odnieść się mogą także do samego systemu biblioteki cyfrowej (układ graficzny strony internetowej, ocena biblioteki cyfrowej, architektura informacji w bibliotece cyfrowej).

Z bardzo nielicznych prac o estetyce bibliotek cyfrowych, które można przedstawić w tym miejscu, wymienić należy publikację odnoszącą się do zasobów książek specjalnego typu, tzn. książek dla dzieci, które w bibliotekach cyfrowych odnaleźć można. Artykuł ten autorstwa Kary Reuter *Assessing aesthetic relevance: children's book selection in a digital library* ukazał się w 2007 r. Autorka opracowała zagadnienie dokonywania przez dzieci ze szkół podstawowych w Stanach Zjednoczonych wyboru własnych lektur z dostępnych bibliotek cyfrowych. W badaniach zidentyfikowała w sumie 46 czynników zorganizowanych w siedmiu wymiarach, które wpływają na oceny dzieci, w tym właśnie aspekty estetyczne. Te siedem wymiarów i ich czynniki to: 1) metadane oraz elementy fizyczne książki (tytuł, okładka, strona tytułowa, format, autor / ilustrator, obecność ilustracji, niedawność, spis treści, noty, współpracownik, uwagi),

2) dostępność (język, długość tekstu, czytelność, wielkość, poziom trudności, gęstość tekstu), 3) zawartość (ilustracje, streszczenie, temat, fabuła, rodzaj, nagłówki z hasłami przedmiotowymi, styl), 4) zaangażowanie (gust, moda - '*coolness*', nastrój, jakość, zainteresowanie, doświadczenia czytelnicze, atrakcyjność wizualna, wzajemne oddziaływanie), 5) nowość (ciekawość, ciekawość językowa, nowość, niezwykłość, różnorodność), 6) aspekty społeczno-kulturowe (pochodzenie, dzielenie się wiedzą, więzi społeczne, osobiste związku, intertekstualność, wartości, popularność), 7) znajomość (znane pozycje, dotychczasowe doświadczenia). Jak widać elementy estetyczne rozrzucone zostały po wielu kategoriach.

4. Bazy danych jako kryterium

Ostatnim zagadnieniem, omawianym w tym artykule jest przedstawienie bazy danych, a więc podstawowego także dla bibliotek cyfrowych elementu, nie tylko jako części systemu, ale jako tworu, który sam umożliwia tworzenie oraz rozszerzoną interpretację nowych dzieł sztuki. Każdy odbiorca zostaje wciągnięty w krąg kolejnych twórców nowych bytów sztuki. Wybory użytkownika w trakcie interakcji z systemem informacyjnym nadają istniejącym w zasobach dziełom nową jakość. Za każdym razem dzieło odtwarza się na nowo. Wpływ na niego ma także sytuacja estetycznego przeżywania go za każdym razem w trochę odmienny sposób, gdyż to użytkownik ze swoim nastawieniem, kryteriami estetycznymi, zakorzenieniem we własnej kulturze, ale też ze swoim otwarciem się na nowe sytuacje zmienia go.

Baza danych jako określona struktura, z praktycznie nieograniczoną możliwością wyborów wydaje się być najbardziej odpowiednia dla informacji, zwłaszcza artystycznej. Nie jest to bowiem struktura ani zbyt otwarta (jak Internet – rozumiany jako system), nie do ogarnięcia dla odbiorcy, ani zbyt zamknięta – model rozszerzać można o kolejne podmioty, grupy i relacje.

Twórcą takiego rozumienia baz jest Lev Manovich, który opisuje je jako podstawową strukturę przede wszystkim dla tzw. nowych mediów, czyli mediów zdigitalizowanych lub elektronicznych (przynajmniej częściowo). Filozofia baz danych – jej strukturalność oraz relacyjność to nowy sposób indeksowania: *rzeczywistości (sztuki nowych mediów), który znakomicie nadaje się do jej krytycznej (re)prezentacji*. Bazy danych są formami symbolicznymi, posiadającymi strukturę paradygmatyczną (asocjacyjną, oparta na związkach). Są także eksplicytne, czyli charakteryzują się wyraźną, możliwą do uchwycenia strukturą. Manovich przeciwstawia je tradycyjnej narracji, którą opisuje jako strukturę o porządku syntagmatycznym (z ciągiem tekstowym) i implicytnym. Narracja jest linearna, w bazach można wyszukiwać, nawigować, oglądać. Baza danych przedstawia świat w postaci listy elementów do wyboru, narracja tworzy ciągi przyczynowo-skutkowe [Manovich 2006; Zawojski, 2007].

Są dzieła, które te dwa porządki łączą. Wedle Manovicha udało się to już w filmie *Człowiek z kamerą* Dzigi Wiertowa [z 1929 r., przechowywany w Internet Archive - www.archive.org]. Reżyser przedstawił w nim montażownię, w której 'życie' po-

dzielone jest na kolejne ujęcia. Dopiero montaż czyni z nich historię, przedstawia ład świata – poprzez film na trzech poziomach: film, tworzenie filmu, film o widowni oglądającej film [[Manovich 2006].

Na poparcie tezy, iż baza może stać się podstawową strukturą dla informacji artystycznej, przytoczyć można także rozważania o tekście, jako o sieci połączeń i odniesień, *rozpatrywanym w kategoriach otwartej i żywej sieci* z pism Gillesa Deleuze`a i Felixa Gautarriego. Ich 'kłącze' to otwarty system, którego struktura zbudowana jest z punktów i maksymalnej liczby połączeń między nimi. Wydaje się jednak, że myślenie o relacyjności i formacie tej otwartej przestrzeni można zamknąć w bazie, w jakiś sposób skończonej – przestrzeń zbudowana jest bowiem z punktów (a więc obiektów „skończonych"). Inne cechy 'kłącza', czyli heterogeniczność danych oraz różne możliwości dotarcia do elementów przestrzeni (bez wyraźnego centrum) są immanentną cechą baz relacyjnych, czy obiektowych. Kolejne cechy to zerwanie połączeń i odbudowa kłącza – a więc informacja o strukturze jest zapisana w częściach systemu – choć może się on odrodzić w zmienionej formie. Piąta i szósta zasada to zasada *kartografii i przekalkowania*. Kłącze jest mapą, które można *rozrysować na nowo, [...] przystosowywać do nowych warunków* [Pisarski, 2001-2008]. Podobnie poprawiamy bazy, także te tworzone dla bibliotek cyfrowych, przystosowujemy je do wymagań użytkowników, nowych możliwości technologicznych oraz obiektów, w tym dzieł sztuki, które coraz częściej 'goszczą' w zasobach bibliotek cyfrowych.

Współczesnych przykładów struktur bazodanowych w odniesieniu do informacji artystycznej jest już bardzo dużo. Są nimi chociażby rozbudowane platformy do tworzenia nowych obiektów artystycznych, które można podciągną także pod ogólną definicję bibliotek cyfrowych, gdyż są także galeriami dzieł. Wymienić tu można: Database of Virtual Art. (http://www.virtualart.at/), ArtBase (w ramach: Rhizome.org – http://rhizome.org/artbase/featured/), netzspannung.org (http://netzspannung.org/), czy Media Art Net (http://www.mediaartnet.org/). Wszystkie one oparte są o bazy danych. Database of Virtual Art zawiera bazę z danymi bio- i bibliograficznymi, które osobie oglądającej pozwalają na lepsze orientowanie się w specyfice, własnościach i powiązaniach danych obiektów. ArtBase to sieciowe archiwum nowych mediów, lokalizujące dzieła sztuki poprzez relacje, linki, informacje o „klonach" dzieł itp. netzspannung.org to dostępna dla wszystkich użytkowników baza danych o sztuce współczesnej o kilku ciekawych interfejsach: a) „klasycznym" interfejsie bazodanowym – wyszukiwanie informacji odbywa się poprzez spis treści oraz wybór spośród kryteriów: daty, projektu, użytych narzędzi; b) przeglądarce archiwów: przeszukiwanie według autorów, konkretnych dzieł, artykułów, wykładów; c) „randomizera" automatycznie generującego przypadkowych trzydzieści obrazów, z których każdy, po kliknięciu, przenosi użytkownika do wybranego obiektu w bazie danych; d) mapie semantycznej. Natomiast Media Art Net to platforma o najwyższym poziomie i zakresie opracowań o sztuce mediów, której twórcy: Dieters Daniels i Rudolf Frieling preferują zamieszczanie materiałów wyselekcjonowanych, połączonych relacjami. One to zarówno w sposób intuicyjny, jak i intelektualny pozwalają użytkownikowi zapoznać się ze zgromadzonymi danymi. Do tych przykładowych projektów dodać jeszcze nale-

ży platformy komunikacyjne, takie chociażby jak: Rhizome [www.rhizome.org], The Thing [bbs.thing.net], czy Furtherfield [www.furtherfield.org], które także wykorzystują immanentną cechę baz tzn. strukturalność [Zawojski 2006; Janiak 2008].

Inne przykłady to np. próba wizualizacji wybranych klasycznych tekstów: TextArc (Tekst w łuku) – W. Bradforda Paley'a (http://www.TextArc.org). *Alicja w krainie czarów* czy *Hamlet* zostały przedstawione w postaci połączonych relacjami słów, z zaznaczeniem tych najczęściej używanych. Przestrzenią fazową dla dzieła jest elipsa, a zwykłą narrację uzupełnia relacyjność. Wizualizacja taka otwiera więcej możliwości interpretacyjnych, np. inne spojrzenie na zrozumienie tekstu z odniesieniami do statystyki.

Olbrzymie możliwości przechowywania oraz porządkowania danych zostały także wykorzystane w projekcie SemaSpace (http://residence.aec.at/didi/FLweb/) [2006 – Ars Electronica – Gerhard Dirmoser i Dietmar Offenhuber]. Wykorzystując grafikę 2-D i 3-D, zgromadzono w bazie danych zasoby ponad dziesięciu tysięcy dynamicznych rekordów odnoszących się do cyberkultury, *które mogą być manualnie, za pomocą prostego interfejsu, wyświetlane, aranżowane w postaci przestrzennej wizualizacji danych tekstowych, obrazowych, graficznych, dźwiękowych*. Projekt ten to rodzaj paradygmatycznego modelu prezentacji wiedzy w sieci, oparty o możliwości wizualizowania sieci semantycznych. Ma ułatwić wyszukiwanie informacji: *w świecie stale powiększającej się panbibliotece sieciowej* [Zawojski 2007]. Innym przykładem jest projekt edukacyjny, który historię sztuki przedstawia w postaci diagramów relacyjnych, opartych o struktury baz. Projekt Ars (http://www.servus.at/kontext/ARS/), pozwala poza oglądem zmian historycznych, przeanalizować relacje pomiędzy poszczególnymi nurtami i artystami.

Przykłady te dokładnie pokazują, jaka jest 'information art'. Platformy bazodanowe dają bowiem poza informacją o dziełach sztuki także możliwości wykorzystywania części z nich do tworzenia własnych obiektów. Piotr Zawojski napisał wręcz: *współczesna kultura audiowizualna ma charakter totalnego palimpsestu* [Zawojski 1996, s. 162] a Wolfgang Ernest stwierdził, iż obecnie: *jesteśmy w trakcie indukowanego przez media przejścia od kultury zorientowanej na magazynowanie do kultury transmisji* [Zawojski 2007]. Nasza współczesna kultura złożona jest z różnych dzieł, które artysta wybiera – przetwarzając i tworząc nowe wartości, a więc 'produkując' nowe doznania estetyczne. Podobnie czyni odbiorca – wybierając różne obiekty, łącząc je oraz nadając im dodatkowe wartości poprzez umieszczenie ich we własnym zewnętrznym modelu relacyjnym rzeczywistości.

Zakończenie

W pracy starano się przedstawić wzajemne oddziaływania pomiędzy zagadnieniami związanymi z estetyką oraz z bibliotekami cyfrowymi. Jak widać z przeprowadzonych analiz nie są to jeszcze dopracowane przez badaczy tematy, głównie ze względu na krótką historię bibliotek cyfrowych. Kwestie, które już doczekały się wstępnych

badań to przede wszystkim estetyka systemów oraz zawartości bibliotek cyfrowych, głównie w odniesieniu do zagadnień branych pod uwagę w trakcie ewaluacji bibliotek cyfrowych. Możliwe także są translokacje myśli badawczej odnoszące się do dzieł drukowanych oraz dzieł sztuki z zaznaczeniem nowej drogi ich odbioru (ze względu na nowy nośnik) – myśl estetyczna, bibliologiczna, informatologiczna, design'erska, typograficzna i informatyczna. Także popatrzenie na biblioteki cyfrowe jako na platformy informujące o sztuce, zwłaszcza o nowej, tworzy pole badań. I ostatnim zagadnieniem, związanym z bibliotekami cyfrowymi i estetyką jest propozycja Lva Manovicha, który przedstawił bazy danych jako kryterium i strukturę sztuki.

Na zakończenie już stwierdzić należy, iż pomimo tego, że czasem, gdy *nasi najlepsi naukowcy i technicy w swym działaniu kierują się estetyką, rodzi* [się] *poczucie zażenowania* [Gelernter 1999, s. 23], to jednak powinniśmy zwrócić estetyce i pięknu należne im miejsce na jakie *zasługuje i jakiej domaga się w czasach powszechnej estetyzacji* [Zawojski 2006b], zwłaszcza ze względu na konstrukcję psycho-fizyczną ludzi tworzących i wykorzystujących różnego rodzaju systemy informacyjne, w tym biblioteki cyfrowe. Za Subrahmanyan'em Chandrasekhar'em, astrofizykiem i matematykiem powtórzyć można: *dążmy do piękna w nauce* [1999, s. 98].

Bibliografia

1. *A Framework of Guidance for Building Good Digital Collections – the 3rd edition* (2007) [dok. elektr.] *Institute for Museum and Library Services* http://framework.niso.org/ [odczyt: 2.02.2011].
2. *A list of quality selection criteria: a reference tool for Internet subject gateways.* [dok. elektr.] UKOLN Metadata Group http://www.ukoln.ac.uk/metadata/desire/quality/report-2.html [odczyt: 2.02.2009].
3. American Library Association (1997-2012). *Great Web Sites for Kids Selection Criteria.* [dok. elektr.] http://www.ala.org/ala/mgrps/divs/alsc/greatwebsites/greatwebsitesforkids/greatwebsites.cfm [odczyt: 2.02.2009].
4. *Architektura informacji* (2011). [dok. elektr.] http://pl.wikipedia.org/wiki/Architektura_informacji [odczyt: 2.12.2011].
5. Balcerak-Wawrzaszek, Agnieszka (2002). David Gelernter: *Mechaniczne piękno. Kryterium estetyczne w informatyce* [dok. elektr.] *Biuletyn Informacyjny. Informatyka dla Szkoły* 2002, nr 1 (36) http://www.vulcan.pl/biuletyn/bi36/11czytaniekrzepi.pdf [odczyt: 2.02.2009].
6. Beck, Susan E. (2009). *The Good, The Bad & The Ugly: or, why it's a good idea to evaluate web sources* [dok. elektr.] Collection Development, New Mexico State University Library http://lib.nmsu.edu/instruction/evalcrit.html [odczyt: 2.02.2009].
7. Bednarek-Michalska, Bożena (2002). *Ocena jakości bibliotekarskich serwisów informacyjnych udostępnianych w Internecie* [dok. elektr.] „EBIB Elektroniczny Biuletyn Informacyjny Bibliotekarzy" nr 31 http://ebib.oss.wroc.pl/2002/31/michalska.php [odczyt: 2.02.2009].
8. Bednarek-Michalska, Bożena (2007a). *Informacja w Internecie: podręcznik studenta.* [dok. elektr.] Kujawsko-Pomorska Biblioteka Cyfrowa http://kpbc.umk.pl/dlibra/docmetadata?id=32425&from=pubindex&dirids=79&lp=39 [odczyt: 2.02.2009].

9. Bednarek-Michalska, Bożena (2007b). *Ocena jakości informacji elektronicznej: pułapki sieci* [dok. elektr.] „EBIB Elektroniczny Biuletyn Informacyjny Bibliotekarzy" nr 5(86) http://www.ebib.info/2007/86/a.php?bednarek [odczyt: 2.02.2009].
10. Bednarek-Michalska, Bożena (2007c). *Ocena jakości serwisu internetowego* [dok. elektr.] http://moodle.weinoe.us.edu.pl/file.php?file=%2F1%2FTechnologia_informacyjna_w_edukacji_wczesnoszkolnej%2FKryteria_Oceny_Stron_internetowych-1.doc [odczyt: 2.02.2009].
11. Bednarek-Michalska, Bożena (2007d). *Oceń jakość informacji elektronicznej!* [dok. elektr.] http://www.biblos.pk.edu.pl/files/File/BazTOL-Ocen%20jakosc%20informacji.ppt oraz http://www.bu.uni.torun.pl/BE/ocenjakosc.html [odczyt: 2.02.2009].
12. Bednarska-Ruszajowa, Krystyna (2003). *Biblioteki i książki w pamiętnikach polskich XVIII--XX wieku: rekonesans źródłowy.* Kraków: Wydaw. Uniw. Jagiel., 200 s.
13. Bednarska-Ruszajowa, Krystyna (2006). *Biblioteki w literaturze polskiej.* Kraków: Wydaw. Uniw. Jagiel., 254, [1] s.
14. Bednarska-Ruszajowa, Krystyna red. (1998). *Biblioteki i książki w literaturze.* Kraków: Wydaw. Uniw. Jagiel., 166 s.
15. Beynon-Davies, Paul (2003). *Systemy baz danych.* Warszawa: Wydawnictwa Naukowo--Techniczne, s. 547.
16. Bień, Janusz S. (2006). *Wizualizacja wyników analizy syntaktycznej.* „Poradnik Językowy" z. 9, s. 24-29.
17. Bojar, Bożenna oprac. (2002). *Słownik encyklopedyczny informacji, języków i systemów informacyjno wyszukiwawczych.* Warszawa: Wydaw. Stowarzyszenia Bibliotekarzy Polskich, s. 261.
18. Burnham, Jack W. (2006). *The Aesthetics of Intelligent Systems* [dok. elektr.] http://www.volny.cz/horvitz/burnham/aesth-sys.pdf [odczyt: 2.06.2011].
19. Candela, Leonardo; Nardi, A. red. (2011). *The Digital Library Reference Model* [dok. elektr.] http://www.dlorg.eu/index.php/outcomes/reference-model [odczyt: 2.06.2011].
20. Caywood, Carolyn (1995). *Library Selection Criteria for WWW Resources* [dok. elektr.] http://www.keele.ac.uk/depts/aa/landt/lt/Internet/criteria.htm [odczyt: 2.02.2009].
21. Chandrasekhar, Subrahmanyan (1999). *Prawda i piękno: estetyka i motywacja w nauce.* Warszawa: Prószyński i Sp., 239 s.
22. Chmielecki, Konrad (2008). *Kwestionariusz kultury wizualnej* [dok. elektr.] *KW: O wizualności i jej społecznych aspektach.* http://www.kulturawizualna.pl/archives/154 [odczyt: 2.02.2011].
23. Ciołek, Matthew T. (1994). *Information Quality – catalogue of potent truisms* [dok. elektr.] http://www.ciolek.com/WWWVLPages/QltyPages/QltyTruisms.html [odczyt: 2.02.2009].
24. Ciołek, Matthew T. (1996a) *Information Quality – some definitions* [dok. elektr.] http://www.ciolek.com/WWWVLPages/QltyPages/QltyDefinitions.html [odczyt: 2.02.2009].
25. Ciołek, Matthew T. (1996b) *The six questions for the electronic grail: current approaches to information quality in WWW resources* [dok. elektr.] http://www.ciolek.com/PAPERS/six--quests1996.html [odczyt: 2.02.2009].
26. Ciołek, Matthew T. (1997) *Information Quality WWW Virtual Library: the Internet guide to construction of quality online resources.* [dok. elektr.] http://www.ciolek.com/WWWVL--InfoQuality.html [odczyt: 2.02.2009].
27. Cornell University Library (2004). *Critically analyzing information sources.* [dok. elektr.] http://www.library.cornell.edu/olinuris/ref/research/skill26.htm [odczyt: 2.02.2009].
28. Dederko, Witold (1985). *Warsztat techniczny artysty fotografa.* Warszawa: COK, 239 s.

29. Drabek, Marcin (2010). *Dominacja wzroku: kryzys pisma po zwrocie wizualnym* [dok. elektr.] „Kultura, historia, globalizacja", nr 8 http://www.khg.uni.wroc.pl/files/5khg8drabek.pdf [odczyt: 2.02.2011].
30. *Estetyka* (cop. 1997-2011) [dok. elektr.] *Encyklopedia PWN* http://encyklopedia.pwn.pl/haslo.php?id=3898755 [odczyt: 2.02.2011].
31. Fuhr, Norbert [i in.] (2007). *Evaluation of digital libraries* [dok. elektr.] „The International Journal on Digital Libraries" http://www.scribd.com/doc/185523/Evaluation-of-Digital--Libraries-Fulltext [odczyt: 3.03.2011].
32. Gelernter, David (1999). *Mechaniczne piękno: kryterium estetyczne w informatyce*. Warszawa: CIS; „W.A.B.", 206 s.
33. Gołaszewska, Maria (1990). *Portret piękna: studium o społecznym wymiarze wartości estetycznych*. Kraków: UJ, 247 s.
34. Gombrich, E.H. (1996). *W poszukiwaniu historii kultury.* [W:] Białostocki, Jan wybór i wstęp (1996). *Pojęcia, problemy, metody współczesnej nauki o sztuce*. Warszawa: PWN, s. 302-345.
35. Gombrich, E.H. (2009). *Zmysł porządku: o psychologii sztuki dekoracyjnej*. Kraków: Universitas, XII, 409, [2] s., [8] s.
36. Grassian, Esther (1996). *Thinking critically about world wide web resources* [dok. elektr.] http://www.wartburg.edu/library/infolit/Handouts/ThinkingCritically.doc oraz http://www.uwm.edu/~maryb/evalnet.htm [odczyt: 2.02.2009].
37. *Great Web Sites for Kids Selection Criteria* (1997). [dok. elektr.] American Library Association [odczyt: 2.02.2009] http://www.ala.org/ala/mgrps/divs/alsc/greatwebsites/greatwebsitesforkids/greatwebsites.cfm [2009].
38. Gutowska, Krystyna (2006). *Zrozumieć sztukę: wybrane zagadnienia wiedzy o sztuce*. Wyd. 3. Wocław: Oficyna Wydaw. Politechniki Wrocławskiej, 100 s.
39. Haynes, John D.; Paradice, David B. *Information technology and aesthetics: passive and active.* [dok. elektr.] http://www.google.pl/url?sa=t&rct=j&q=haynes%2C%20john%20d.%3B%20paradice%2C%20david%20b.%20information%20technology%20and%20aesthetics%3A%20passive%20and%20active&source=web&cd=3&ved=0CC4QFjAC&url=http%3A%2F%2Fdl.acs.org.au%2Findex.php%2Fajis%2Farticle%2Fdownload%2F30%2F411&ei=WsAOT6HcGsee-waDktHMAg&usg=AFQjCNHpaL59jdKIPHhwAeksHVOB2sURTA [odczyt: 2.12.2011].
40. Hernandez Michael J. (1998). *Bazy danych dla zwykłych śmiertelników*. Warszawa: "Mikom", 329 s.
41. *ICT for Library and Information Professionals. A Training Package for Developing Countries* (2001) [dok. elektr.] Unesco. Module 4: *Database design, and information storage and retrieval.* http://dlist.sir.arizona.edu/1537/01/127825e.pdf [odczyt: 2.02.2009]; Module 5: *The Internet as the Information resource.* http://dlist.sir.arizona.edu/1538/01/127901e.pdf [odczyt: 2.02.2009] oraz http://greenstonesupport.iimk.ac.in/greenstone2010-nepal/pdf/UNESCO-Publications/Introduction%20to%20information%20and%20communication%20technologies.pdf [odczyt: 12.12.2011].
42. Janiak, Małgorzata (2008). *Bazy danych – kryteria estetyczne.* [W:] Nowakowski, Antoni red. (2008). *Infobazy'2008 „Systemy – Aplikacje – Usługi"*, Materiały V Konferencji Naukowej, Sopot 15-17 października 2008 r. Gdańsk: Wydawnictwo Centrum Informatyczne TASK, Politechnika Gdańska, s. 174-179.
43. Janiak, Małgorzata (2009a). *Kryteria estetyczne w ocenie jakości bibliotek cyfrowych.* Referat wygłoszony na konferencji: 04-05.04.2011 – Warszawa „Nauka o informacji (informacja

naukowa) w okresie zmian". Organizator: Instytut Informacji Naukowej i Studiów Bibliologicznych Uniwersytetu Warszawskiego.

44. Janiak, Małgorzata (2009b). *Kryteria estetyczne w ocenie jakości systemów informacyjnych.* [W:] Kamińska, Joanna; Żołędowska-Król, Beata red. (2009). *Jakość usług bibliotecznych w społeczeństwie informacyjnym: praca zbiorowa.* Warszawa: Wydaw. SBP, s. 136-144.
45. Janiak, Małgorzata; Krakowska, Monika (2010). *Ocena bibliotek cyfrowych – kryteria jakości.* [W:] Mazurek, Cezary, Stroiński, Maciej; Węglarz, Jan red. (2010). *Polskie biblioteki cyfrowe 2010: materiały z konferencji zorganizowanej w dniach 20-21 października 2010 roku przez: Bibliotekę Kórnicką PAN, Poznańską Fundację Bibliotek Narodowych, Poznańskie centrum Superkomputerowo-Sieciowe.* Poznań: Ośrodek Wydawnictw Naukowych, s. 101-108.
46. Johns Hopkins University, The Sheridian Libraries (2010). *Evaluating Information Found on the Internet.* [dok. elektr.] http://guides.library.jhu.edu/evaluatinginformation [odczyt: 2.02.2011].
47. Kluszczyński, Ryszard W. (2001). *Społeczeństwo informacyjne. Cyberkultura. Sztuka multimediów.* Warszawa: Rabid, 219 s.
48. Knosala, Bartłomiej (2009). *Projekt nauki nowej Marshalla McLuhana: filozoficzne konsekwencje zmian form komunikacji* [dok. elektr.] *Śląska Biblioteka Cyfrowa* http://www.sbc.org.pl/Content/20927/doktorat2974.pdf [odczyt: 5.03.2011].
49. Komza, Małgorzata (2002). *Estetyczne walory publikacji elektronicznych.* [W:] Kocójowa, Maria red. nauk. (2002). *Elektroniczne publikacje w bibliotekach.* Kraków: Wydaw. Uniw. Jagiel., s. 36-44.
50. Komza, Małgorzata (2004). *Ewolucja informacyjnych walorów ilustracji naukowych.* [W:] Kocójowa, Maria red. (2004). *Przestrzeń informacji i komunikacji społecznej.* Kraków: Wydaw. Uniw. Jagiel., s. 349-356.
51. Komza, Małgorzata (2009). *Zbiory graficzne we współczesnej cywilizacji obrazu.* [W:] Komza, Małgorzata red. (2009). *Świat w obrazach: zbiory graficzne w instytucjach kultury – ich typologia, organizacja i funkcje.* Wrocław: Wydaw. Uniw. Wrocł., s. 9-18.
52. Kurosu, Masaaki; Kashimura, Kaori (1995). *Apparent usability vs. inherent usability: experimental analysis on the determinants of the apparent usability.* CHI ,95 Conference companion on Human factors in computing systems [dok. elektr.] http://dl.acm.org/citation.cfm?id=223680&CFID=59988674&CFTOKEN=37606404 [odczyt: 10.10.2011].
53. Kuspit, Donald B. (2004). *Koniec sztuki.* Gdańsk: Muzeum Narodowe, XVII, [3], 208 s.
54. Kuźmina, Dariusz; Tobera, Marek red. (2006). *Nauka o książce: antologia tekstów : praca zbiorowa.* Warszawa: Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich, 254, [1] s.
55. Lau, Andrea; Moere, Andrew Vande (2007). *Towards a model of information aesthetics in information visualization. 2007 11th International Conference Information Visualization IV 07,* s. 87-92.
56. Laviea, Talia; Tractinsky, Noam (2004). *Assessing dimensions of perceived visual aesthetics of web sites. Int. J. Human-Computer Studies* 60, s. 269–298.
57. Manovich, Lev (2006). *Język nowych mediów.* Warszawa: Wydaw. Akademickie i Profesjonalne, strony
58. Muraszkiewicz Mieczysław Ryszard; Rybiński, Henryk (1993). *Bazy danych.* Warszawa: Akademicka Oficyna Wydawnicza RM, VI, [2], 163 s.
59. Osińska, Veslava (2010). *Wizualizacja i wyszukiwanie dokumentów.* Warszawa: Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich, 183, [1] s.

60. *Piękno* (cop. 1997-2011) [dok. elektr.] *Encyklopedia PWN* http://encyklopedia.pwn.pl/haslo.php?id=3957085 [odczyt: 2.02.2011].
61. Pisarski, Mariusz (2001-2008). *Postmodernizm – Foucault i Derrida* [dok. elektr.] http://www.techsty.art.pl/hipertekst/teoria/postmodernizm/post_533.htm [odczyt: 1.09.2008].
62. Potocka, Maria Anna (2007). *Estetyka kontra sztuka: kompromitacja założeń estetycznych w konfrontacji ze sztuką nowoczesną.* Warszawa: Wydawnictwo Fundacja Aletheia, 254 s.
63. *Quality Selection Criteria for Subject Gateways* [dok. elektr.] Produced by the European Union-funded DESIRE Project http://sosig.ac.uk/desire/qindex.html [odczyt: 2.02.2009].
64. Ramirez, Rafael (1991*). The beauty of social organization.* Munich: Accedo Verlagsgesellschaft, 167 s.: il.
65. Reuter, Kara (2007). *Assessing aesthetic relevance: children's book selection in a digital library* [dok. elektr.] http://onlinelibrary.wiley.com/doi/10.1002/asi.20657/pdf [odczyt: 2.02.2011].
66. Sandusky, Robert J. (2004). *Digital Library Attributes: Framing Usability Research* [dok. elektr.] http://www.uclic.ucl.ac.uk/annb/docs/Sandusky35.pdf [odczyt: 5.03.2011].
67. Sapa, Remigiusz (2005). *Benchmarking w doskonaleniu serwisów WWW bibliotek akademickich.* Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego, 254 s.: il.
68. Saracevic, Tefko (2000). *Digital library evaluation: toward an evolution of concepts– 1– evaluation criteria for design and management of digital libraries.* [dok. elektr.] *Library Trends* http://findarticles.com/p/articles/mi_m1387/is_2_49/ai_72274400/pg_18/?tag=content;col1 [odczyt: 3.03.2011].
69. Saracevic, Tefko (2004). *Evaluation of digital libraries: an overview, paper presented at DELOS Workshop on the Evaluation of Digital Libraries* [dok. elektr.] *DELOS* http://dlib.ionio.gr/wp7/WS2004_Saracevic.pdf [odczyt: 2.02.2009].
70. SBP – recenzja książki *Wizualizacja i wyszukiwanie dokumentów* Veslavy Osińskiej (2010). [dok. elektr.] http://www.sbp.pl/wydawnictwa/produkt?produkt_id=27 [odczyt: 2.12.2011].
71. Schrock, Kathy (2008). [dok. elektr.] www.kathyschrock.net [odczyt: 2.02.2009].
72. Smith, Alistair (1996). *Criteria for Evaluation of Internet Information Resources.* [dok. elektr.] http://www.vuw.ac.nz/~agsmith/evaln/index.htm [odczyt: 2.02.2009].
73. Smith, Alistair (1996; modyf. 2011). *Evaluation of Information Resources.* [dok. elektr.] http://www.vuw.ac.nz/~agsmith/evaln/evaln.htm [odczyt: 2.02.2009].
74. Stolterman, Eric (1994). *The aesthetics of information systems.* [dok. elektr.] http://www.mendeley.com/research/users-some-theorles-modeling-users-information-systems-some--theories-methods/ [odczyt: 2.02.2009].
75. *Erik Stolterman talks Design at CMU* (2008). [dok. elektr.] http://jamin.org/archives/2008/erik-stolterman-talks-design-at-cmu/ [odczyt: 2.02.2009].
76. Subieta, Kazimierz (2009) *Ideologia, nauka, wynalazczość i komercja w informatyce (na przykładach z obiektowości i baz danych)* [dok. elektr.] Strona prof. dr hab. Kazimierza Subiety http://www.ipipan.waw.pl/~subieta/prezentacje/ ideologia%20nauka%20wynalazczosc%20i%20komercja%20w%20informatyce%201998.ppt [odczyt: 2.02.2009].
77. Tatarkiewicz, Władysław (1988). *Dzieje sześciu pojęć: sztuka, piękno, forma, twórczość, odtwórczość, przeżycie estetyczne.* Warszawa: Państ. Wydaw. Naukowe, 437, [3] s.
78. Tate, Marsha; Alexander, Jane (1996). *Teaching critical evaluation skills for world wide web resources.* „Computers in Libraries" nr 16(10), s. 49-54.
79. Tillman, Hope N. (1996). *Evaluating quality on the net.* [dok. elektr.] http://www.hope-tillman.com/findqual.html [odczyt: 2.02.2009].

80. Tractinsky, Noam (1997). *Aesthetics and apparent usability: empirically assessing cultural and methodological issues.* „Proceedings of the ACM Conference on Human Factors in Computing Systems" (March), s. 115-122.
81. *Typografia* [dok. elektr.] Gazeta.pl Słowniki http://slowniki.gazeta.pl/pl/typografia [odczyt: 2.02.2011].
82. *Typografia* (2011) [dok. elektr.] http://pl.wikipedia.org/wiki/Typografia [odczyt: 2.02.2011].
83. UC Berkeley – Teaching Library Internet Workshops (2008). *Evaluating Web Pages: Techniques to Apply & Questions to Ask.* [dok. elektr.] http://www.lib.berkeley.edu/TeachingLib/Guides/Internet/Evaluate.html oraz http://www.lib.berkeley.edu/TeachingLib/Guides/Internet/evaluation_checklist_2008_spring.pdf [odczyt: 2.02.2009].
84. Ullman, Jeffrey D.; Widom, Jennifer (2000). *Podstawowy wykład z systemów baz danych.* Warszawa: Wydawnictwa Naukowo-Techniczne, 525, [1] s.
85. Wilson, Steve (2001). *Information Arts: intersections of art, science and technology information arts: intersections of art, science, and technology.* Cambridge: The MIT Press, A Leonardo Book, 1024 s.
86. Woelfflin, Heinrich (2006). *Podstawowe pojęcia historii sztuki.* Gdańsk: Słowo/Obraz Terytoria, 216, [4] s., [123] s.
87. Xie, Hong (2006). *Evaluation of digital libraries: criteria and problems from users' perspective.* „Library and Information Science Research" 28, s. 433-452.
88. Xie, Hong (2008). *Users' evaluation of digital libraries: their uses, their criteria, and their assessment.* „Information Processing & Management" 44(3), s. 1346-1373.
89. Zawojski, Piotr (2006a). *Archiwizacja, prezentacja i dyseminacja cybersztuki w sieci* [dok. elektr.] http://www.zawojski.com/2006/11/15/archiwizacja-prezentacja-i-dyseminacja-cybersztuki-w-sieci/ [odczyt: 17.06.2008]
90. Zawojski, Piotr (2006b). *Ciało jako interfejs: o kilku (nie)przypadkowych sytuacjach, w których staję się (współ)twórcą.* [dok. elektr.] Blog Piotra Zawojskiego http://www.zawojski.com/2006/04/19/cialo-jako-interfejs-o-kilku-nieprzypadkowych-sytuacjach-w-ktorych--staje-sie-wspoltworca/ [odczyt: 2.02.2009].
91. Zawojski, Piotr (1996). *„Stara" estetyka w konfrontacji z nowymi mediami.* [W:] Zeidler--Janiszewska, Anna red. (1996). *Estetyczne przestrzenie współczesności.* Warszawa: Instytut Kultury, s. 162.
92. Zawojski, Piotr (2007). *Teoretyczne konteksty dokumentacji sztuki nowych mediów* [dok. elektr.] Blog Piotra Zawojskiego http://www.zawojski.com/2007/07/09/teoretyczne-konteksty-dokumentacji-sztuki-nowych-mediow/ [odczyt: 2.02.2011].
93. Zawojski, Piotr (2007). *Teoretyczne konteksty dokumentacji sztuki nowych mediów* [dok. elektr.] http://www.zawojski.com/2007/07/09/teoretyczne-konteksty-dokumentacji-sztuki-nowych-mediow/ [odczyt: 17.06.2008].
94. Zawojski, Piotr (2010). *Cyberkultura: syntopia sztuki, nauki i technologii.* Warszawa: Poltext, 328 s.
95. Zhang, Ying (2010). *Developing a holistic model for digital library evaluation* [dok. elektr.] *Wiley Online Library* http://onlinelibrary.wiley.com/doi/10.1002/asi.21220/pdf [odczyt: 2.02.2011].

Standardy i dobre praktyki budowania kolekcji cyfrowych

Bożena Bednarek-Michalska

W pracy przedstawiono główne zasady tworzenia poprawnej kolekcji cyfrowej określone w dokumencie amerykańskiej organizacji NISO (National Information Standards Organization) pt.: *A Framework of Guidance for Building Good Digital Collections* [2007]. Dotyczą one elementów planowania projektowego i zarządzania kolekcją, zasobów, obiektów i metadanych. Jest to próba przyswojenia amerykańskich i europejskich dobrych praktyk dla polskiej pragmatyki zawodowej. Tworząc kolekcję cyfrową należy stosować podobne podejście do zasad, które zostały opisane poniżej.

WSTĘP

Światowe zasoby elektroniczne organizowane w duże kolekcje powstają mniej więcej od 1995 r. i jest w tej chwili bardzo dużo zarówno tych komercyjnych, jak i budowanych przez instytucje publiczne czy fundacje. Ich różnorodność jest ogromna od fachowych po amatorskie, od wąskich tematycznie przez uniwersalne, od prowadzonych zgodnie ze standardami międzynarodowymi do odbiegających od wszelkich zasad.

Budowanie kolekcji cyfrowych jest dość dobrze rozpoznanym zadaniem i opisanym w miarę szczegółowo, nie tylko w zagranicznych opracowaniach, ale także w Polsce. Biblioteka Narodowa, która jest od niedawna Narodowym Centrum Kompetencji w zakresie digitalizacji ma za zadanie świadczyć pomoc wszystkim, którzy chcą się podjąć cyfryzacji. W 2010 r. ukazał się podręcznik Biblioteki Narodowej zatytułowany *Digitalizacja piśmiennictwa* [Paradowski 2010], ponadto odbywają się w niej praktyczne szkolenia, warsztaty i seminaria powiązane z różnymi aspektami digitalizacji, także prawnymi. W 2008 r. dzięki pracy Zespołu ds. digitalizacji pracującego przy MKiDN opracowano także pierwsze w Polsce *Standardy w procesie digitalizacji obiektów dziedzictwa kulturowego* pod redakcją i we współpracy Grzegorza Płoszajskiego,

Tomasza Kaloty, Dariusza Paradowskiego oraz Kazimierza Schmidta. Opublikowano je ze wsparciem Politechniki Warszawskiej.

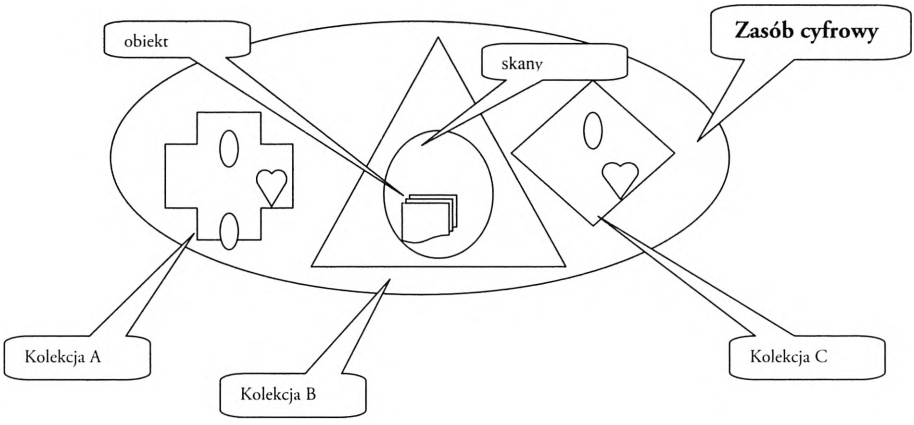
Mamy ponadto w kraju kilka ośrodków, które mogą służyć radą i konsultacją, ponieważ mają długoletnie doświadczenie zawodowe w tym zakresie. Są to: Uniwersytet Mikołaja Kopernika w Toruniu, Uniwersytet Wrocławski, Uniwersytet Poznański, Zakład Narodowy Ossolineum, Biblioteka Kórnicka czy Biblioteka Śląska. Znajdują się w nich pracownie digitalizacji i zespoły fachowców, którzy poradzą, jak zbudować poprawnie kolekcje, jak je opisać czy nimi zarządzać. Mamy także polskie oprogramowanie dLibra i dMuseion dla obsługi tego typu zadań i istotne wsparcie technologiczne instytucji naukowej PAN – Poznańskiego Centrum Superkomuterowo-Sieciowego. Zespół młodych informatyków tego ośrodka prowadzi bardzo ambitne programy rozwojowe, dzięki którym powstała Federacja Polskich Bibliotek Cyfrowych skupiająca w 2010 r. ponad 60 odrębnych bibliotek cyfrowych gromadzących pół miliona obiektów i udostępnionych od 2009 r. w Europeanie.

Skala digitalizacji w Polsce nie jest duża, ale stale się zmienia. W 2006 r. mieliśmy zaledwie kilka bibliotek cyfrowych, w 2008 r. już tworzyliśmy niewielką sieć 17 z nich, która udostępniła online ponad 120 tysięcy obiektów. Był to dość dynamiczny proces pokazujący i ambicje bibliotekarzy i przekonanie, że zadanie to jest standardowe dla bibliotek i innych instytucji posiadających ważne dla Polski kolekcje. Można powiedzieć, że na owe lata było to niemałe osiągnięcie, zważywszy na to, że instytucje państwowe dotujące kulturę czy naukę włączyły się finansowo do cyfryzacji dopiero w roku 2006, wcześniej mieliśmy do czynienia z oddolnymi inicjatywami.

Dlaczego standardy są ważne?

Wartościowe kolekcje, jeśli mają przetrwać kolejne rewolucje technologiczne, stanowić wspólne europejskie dziedzictwo, muszą oprzeć się o kilka zasad, pozwalających stworzyć profesjonalne archiwum cyfrowe[1] z zapewnioną interoperacyjnością, stabilnością, możliwością powtórnego użycia i weryfikacji zasobów, zabezpieczenia obiektów oraz zgodnością z prawem autorskim czy prawem o ochronie danych osobowych. Archiwa cyfrowe (AC) nie powinny być tymczasowymi, przypadkowymi projektami poszczególnych instytucji, działającymi na niesprawdzonym oprogramowaniu, z byle jak opisanymi obiektami, ale muszą być budowane tak, by nie trzeba było powtarzać kosztownego procesu skanowania czy opisywania kolekcji.

[1] Będą stosowała kilka terminów na określenie zasobów cyfrowych, wszystkie można traktować zamiennie, choć ich szczegółowe znaczenie nieco się różni, a zatem będziemy mówić o: *archiwum cyfrowym, bibliotece cyfrowej, kolekcji cyfrowej, zasobie cyfrowym,* można także użyć określenia: *archiwum elektroniczne,* ale to raczej podkreśla techniczny charakter tegoż. *Zasób cyfrowy, kolekcja cyfrowa* rozumiane tu są jako zbiory obiektów cyfrowych (zdigitalizowanych, elektronicznych) i metadanych (informacji tekstowych), które je opisują. *Archiwum cyfrowe* i *bibliotekę cyfrową* można rozumieć szerzej – razem z platformą technologiczną (oprogramowaniem), na której zasób i metadane są posadowione.



Rys. 1. Schemat elementów składowych zasobu cyfrowego KPBC: zasób (KPBC), kolekcje (Baltica, Pomeranica), obiekty (książka, czasopismo) i skan (1 strona z książki).

Dziś z perspektywy czasu warto już pokusić się o dokładną analizę na ile zbudowane w Polsce kolekcje zachowują standardy, jakie powinny być zachowane, by przetrwały 100 lat. Bez takiej analizy ciężko nam będzie naprawić błędy i wejść na ścieżkę naprawy. Brak koordynacji wielu działań sprawia, że nadal mamy kłopoty z jednolitym stosowaniem metadanych opisowych, nie rozumiemy prawa autorskiego w związku z tym źle opisujemy status prawy obiektu cyfrowego, dramatycznie brakuje nam wiedzy z zakresu konserwacji i zachowaniu cyfrowego zasobu. W Polsce nie ma profesjonalnych ośrodków kształcenia w tym konkretnym zakresie. Wiedza, jaką mamy, wynika głównie z doświadczeń praktycznych oraz literatury zagranicznej. To za mało, by podjąć się tak trudnych wyzwań.

Mamy chyba jednak świadomość, że wszystkie procesy związane z digitalizacją powinny respektować międzynarodowe standardy tak, by nasze zasoby stały się częścią światowych kolekcji, dały się wyszukiwać, rozpoznawać, by były wiarygodne i wreszcie trwale przechowywane. Zachowanie dla przyszłości obiektów cyfrowych jest jednym z ważniejszych zadań kultury, nauki czy przemysłu mediów. Powtórzmy zatem te najważniejsze zasady, bez których nie da się zbudować poprawnej biblioteki cyfrowej i zastanówmy się, na ile ich przestrzegamy. Zasady o których będzie mowa w następnej części tego opracowania są określone w dokumencie amerykańskiej organizacji NISO (National Information Standards Organization) zatytułowanym: *A Framework of Guidance for Building Good Dygital Collections* [2007].

PLANOWANIE I PROJEKTOWANIE

Zasada 1: Dobre archiwum cyfrowe musi mieć plan działania od samego początku jego tworzenia.

Planowanie jest najistotniejszą sprawą, jeśli chcemy wyprodukować zasób dobrej jakości. Na początek należy określić grupę docelową, dla której chcemy stworzyć kolekcję i jej potrzeby w tym zakresie. Następnie należy napisać szczegółowy plan działania, który zawiera istotne elementy takie jak: cel działania, polityczne priorytety, czas trwania projektu, rodzaj kolekcji, jaki obejmuje, selekcja, proces digitalizacji, sprawy praw autorskich, dostęp, zarządzanie zasobami, technologie, rozpowszechnianie i ocena jakości. Dopiero z tak opracowanym planem można przystąpić do poszukiwania pieniędzy i realizacji projektu.

Zasada 2: Dobry pomysł potrzebuje fachowców i ekspertów.

Rola zespołu, który będzie projektem zarządzał jest bardzo istotna. Od wiedzy i doświadczenia ludzi, ich zapału zależy sukces projektu. Do realizacji projektu powinno się także zatrudniać ekspertów, którym zleca się szczególne analizy, np. bibliotekarskie, archiwalne, ekonomiczne czy prawnicze, po to, by samemu nie rozwiązywać problemów, na których się nie znamy. Daje to dużą oszczędność czasu i kosztów. Prace organizacyjne i koordynacyjne z takim zespołem powinny być także prowadzone profesjonalnie z wykorzystaniem wszystkich zasobów ludzkich i zapału, jaki ludzie mają na początku tworzenia. Możemy wybrać trzy strategie doboru ludzi: ze swojej własnej instytucji, zatrudnić firmę zewnętrzną do wykonania zadania (outsourcing) lub podjąć współpracę z innym partnerem i wykorzystać jego wiedzę oraz doświadczenie. Każda ze strategii ma wady i zalety.

Zasada 3: Dobra inicjatywa na rzecz budowania biblioteki cyfrowej (BC) powinna korzystać z dobrych praktyk z zakresu zarządzania projektem.

Zarządzanie projektem jest sztuką, która ma swoją tradycję i jest szeroko opisana, szczególnie w praktyce biznesowej. Warto czerpać z tych doświadczeń. Metody opracowywania projektów są różne, ale są pewne stałe komponenty, które trzeba uwzględnić w swoim projekcie.

Planowanie:

a) opisujemy cele i zadania oraz punkt, jaki chcemy osiągnąć kończąc projekt;
b) określamy wskaźniki oceny projektu, które musimy osiągnąć;
c) określamy potrzeby, jakie powinny być spełnione, by zadanie się udało;
d) szacujemy czas potrzebny na wykonanie zadania;
e) opisujemy szczegółowe zadania i rozpisujemy odpowiedzialność dla każdego z nich;

f) piszemy harmonogram projektu.

Implementacja:

a) kiedy zaczniemy realizować projekt musimy na bieżąco śledzić zadania i ich wykonanie;
b) regularnie oceniamy postępy prac i modyfikujemy je, jeśli trzeba;
c) gdyby pojawiły się istotne przeszkody, to zmieniamy plan.

Ocena wykonania zadania:

a) sprawdzamy, czy osiągnięte zostały główne cele;
b) sprawdzamy założone w planie wskaźniki;
c) identyfikujemy zmiany;
d) podsumowujmy zadania i wyciągamy wnioski z jego realizacji, szczególnie badamy te elementy, które nam nie wyszły;

Modyfikacja:

a) poprawiamy błędne założenia;
b) wprowadzamy zmiany do procesów pracy;
c) robimy monitoring w określonym czasie.

Zasada 4: Dobry projekt posiada plan oceny.

Ocena powinna wynikać z realizacji wskaźników, które określi się podczas planowania. Wskaźniki oceny w przypadku zasobu cyfrowego można określić następująco:

1. liczba kolekcji zbudowanych w ramach jednego zasobu;
2. liczba obiektów zeskanowanych (obiektem jest tu np. 100 stronicowa książka) w ramach poszczególnych kolekcji;
3. liczba skanów wykonanych (skan jest to jedna strona książki) dla wszystkich kolekcji;
4. liczba potencjalnych użytkowników;
5. planowany wzrost liczby użytkowników po implementacji archiwum cyfrowego w %.

Ocena może skupiać się na procesach pracy zespołu i wynikach, jakie się uzyska w określonym czasie. W ocenie można też uwzględnić: zarządzanie zespołem, zasobami oraz procedury działania, jakość produktu. Ważne są wyniki: ile zeskanowano stron, ile to jest obiektów, czy działają wdrożone systemy oraz czy sprzęt i wyposażenie zostały zakupione zgodnie z założeniami i są optymalnie wykorzystane, czy liczba użytkowników biblioteki cyfrowej rośnie, itp. Metody oceny mogą być różne, np. badania grup docelowych, robienie wywiadów, badania logów transakcji lub studia przypadków. Wyniki ocen przedstawia się decydentom, donatorom, pomagają nam one także w rozwijaniu projektu.

Zasada 5: Dobry projekt jest promowany, szeroko rozpowszechnia się informację o pracach projektowych i jego realizacji.

Realizatorzy AC powinni zrobić wszystko, żeby dobrze opracowana dokumentacja projektowa była przejrzysta i dostępna dla każdego, zwłaszcza tam, gdzie projekt realizowany jest z pieniędzy podatników. Jeśli projekt może być modelowym powinno się opublikować podręcznik dobrych praktyk, by każdy mógł się z nimi zapoznać i omijać błędy, które były naszym udziałem. Sprawozdania z realizacji projektu powinny być publiczne dostępne. Ten zwyczaj szerokiego dzielenia się wiedzą praktyczną ma wpływ na rozwój i mnożenie się podobnych realizacji w poszczególnych instytucjach. Jeżeli projekt ma znaczenie lokalne, regionalne lub narodowe, powinien być opisane w publikacjach, prezentacjach, anonsowany w mediach. Skuteczne promowanie kolekcji cyfrowej przyniesie pożądane rezultaty, wzrost zainteresowania i wykorzystania. Nowoczesny plan marketingowy zakłada, że promocja projektu podąża tam, gdzie są jego potencjalni użytkownicy.

Zasada 6: Dobry projekt dotyczy całego cyklu życia kolekcji cyfrowej i związanych z nią usług.

Ważne jest, by przy planowaniu pamiętać o tym, że projekt ma przetrwać nie tylko do zakończenia harmonogramu związanego z wnioskowaniem o pieniądze, ale i dłużej. W przypadku biblioteki cyfrowej ważnym jest, by dokonać dobrego wyboru sprzętu, platformy cyfrowej, zapewnić dalsze finansowanie, opracować taką organizację pracy zespołu, żeby po roku zespół się nie rozpadł. Należy także zaprojektować wykorzystywanie zasobów w długim okresie i niezbędne aktualizacje zasobu i stron www. Technologie się zmieniają, musimy założyć ich nieustanne modyfikacje i odnawianie. Niezbędny jest więc także plan zabezpieczenia kolekcji dla przyszłości. Należy obserwować trendy światowe, standardy, patrzeć, w jakim kierunku zmierzają inne podobne inicjatywy i „być zawsze na fali". Współorganizować szersze inicjatywy, narodowe i międzynarodowe, by wspólnie zmieniać to, co stworzy się z wielkim trudem.

ZASOBY – kolekcja

Zasada 1: Dobry zasób cyfrowy jest tworzony na podstawie jasno określonej polityki rozwoju zasobu, uzgodnionej i udokumentowanej przed rozpoczęciem digitalizacji.

Każda biblioteka cyfrowa jest inna, jej rozwój zależy od polityki instytucji, która ją tworzy, jej misji i określonej grupy użytkowników. Warto jednak mieć także na uwadze użytkowników potencjalnych, o których na początku nic nie wiemy, a którzy z pewnością się pojawią. Nie działamy w społeczności lokalnej, ale w sieci glo-

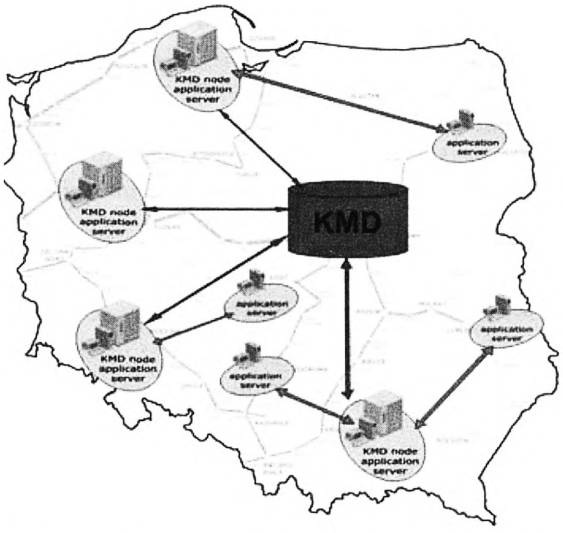
balnej, taka jest różnica. Jeśli instytucja, która tworzy archiwum cyfrowe, ma swoje trwałe, stare kolekcje, to jej cyfrowy zasób powinien być zgodny z gromadzeniem kolekcji tradycyjnej. Zasada ta nie obowiązuje przy masowej digitalizacji i przy digitalizacji na życzenie. Kryteria selekcji materiału do digitalizacji powinny być opisane i publicznie dostępne. Politykę można spisać w paru punktach na jednej kartce. Zob. politykę KPBC: http://kpbc.umk.pl/dlibra/text?id=polityka.

Zasada 2: Zasoby powinny być opisane w sposób pozwalający użytkownikowi uzyskać charakterystyki zasobów, w tym ich zakres, ograniczenia dostępu, własność i inne informacje istotne dla określenia autentyczności, integralności i interpretacji zasobów.

Opis kolekcji powinien pozwalać użytkownikowi zrozumieć, z jakimi zasobami i obiektami ma do czynienia (zakres i zasięg), jakim ograniczeniom związanym z wykorzystaniem podlega kolekcja, do kogo należy, jaki jest jej historyczny kontekst czy pochodzenie, w jakim celu i dla kogo została stworzona itp. Należy zadbać o podanie wszelkich kontaktów, by czytelnik mógł dopytać się o status kolekcji i prawa własności. Powinno się wyjaśnić użytkownikom, jaki software jest potrzebny, by dobrze wykorzystać zasoby. Zob. opis projektu KPBC: http://kpbc.umk.pl/dlibra/text?id=library-desc.

Zasada 3: Nad kolekcją cyfrową trzeba stale sprawować pieczę, dobrze nią zarządzać w trakcie całego cyklu życia i dbać o jej zabezpieczenie.

Zarządzanie kolekcją polega na stałym jej redagowaniu, meliorowaniu, dodawaniu nowych obiektów, usuwaniu zbytecznych, czuwaniu nad pozyskiwaniem praw autorskich, archiwizowaniu, linkowaniu do podobnych obiektów oraz zabezpieczaniu całego zasobu na płytach lub serwerach (zaleca się to drugie rozwiązanie). Zabezpieczenie to także budowanie polityki przyszłej migracji obiektów i emulacji softwaru oraz hardwaru. Warto wiedzieć, że w Polsce powstaje Krajowy Magazyn Danych, który będzie zabezpieczał wartościowe zasoby: zob. strony PCSS https://kmd.pcss.pl/.



Rys. 2. Rozmieszczenie węzłów KMD. Rysunek ze stron PCSS.

Zasada 4: Dobry zasób jest szeroko dostępny bez zbędnych przeszkód w jego użytkowaniu. Zasoby powinny być dostępne dla osób niepełnosprawnych i efektywnie wykorzystywane w połączeniu z technologiami adaptacyjnymi.

Dostępność oznacza, że kolekcja jest w sieci 24/h, co związane jest z implementacją odpowiedniego systemu zabezpieczenia i administrowania. Ktoś musi stale czuwać na biblioteką cyfrową i reagować na wszystkie techniczne problemy. Dostępność nie oznacza, ze wszystkie materiały są za darmo w sieci i bez logowania, to już jest odmienny problem prawny. Kolekcja powinna być także łatwa w użyciu (interface, formaty, wyszukiwanie), szczególnie dla użytkowników niepełnosprawnych pod względem ruchowym, czy przy niedowidzeniu. Podstawowe zasady użytkowania (help) powinny być opisane i dostępne online. Biblioteka cyfrowa powinna być widoczna przez wszystkie przeglądarki, działać zarówno pod Windows, jak i Linux czy Mac, trzeba ją testować pod tym kątem. Powinna mieć wersję językową zrozumiałą dla wszystkich. Zob. wersje językowe Śląskiej Biblioteki Cyfrowej: http://www.sbc.org.pl/dlibra.

Zasada 5: Dobra kolekcja cyfrowa respektuje wszystkie prawa autorskie.

Zasady respektowania prawa autorskiego powinny być opisane i udostępnione online na stronach AC na poziomie całej kolekcji oraz na poziomie obiektu [zob. zasady udostępniania, gromadzenia i zabezpieczania kolekcji w KPBC http://kpbc.umk.pl/dlibra/text?id=polityka]. Należy podać informację o tym, jakie prawa do kolekcji ma

instytucja sprawcza i jej partnerzy, co jest elementem ochrony (baza danych, obiekty cyfrowe) i jaki jest jej zakres. Należy zacytować przepisy prawne, które chronią bibliotekę cyfrową jako całość. Należy brać tu pod uwagę fakt, że w Polsce biblioteki cyfrowe często powstają dzięki funduszom publicznym, zatem roszczenie sobie wyłącznych do nich praw i nakładanie restrykcji wobec obywateli jest niestosowne.

Sprawy prawne są trudne, przepisów prawa trzeba się uczyć, warto mieć bibliotekarza, który będzie się specjalizował w zakresie wyjaśniania praw majątkowych dla naszej kolekcji. Decyzje prawne generują ryzyko, nie zawsze możemy ustalić właściciela praw majątkowych, dlatego zaleca się, by opracować strategię ryzyka, która pozwoli kolekcję szeroko prezentować, ale zminimalizuje ewentualne roszczenia. Należy w takiej strategii brać pod uwagę dobro publiczne oraz ograniczenie strat majątkowych ewentualnych właścicieli praw. Strategię trzeba udostępnić publicznie, by zachować przejrzystość swoich działań.

Zasada 6: Dobre biblioteki cyfrowe posiadają mechanizmy pozwalające na zbieranie informacji o użyteczności zasobów.

Zasoby cyfrowe muszą być cyklicznie oceniane, należy badać efektywność ich użycia, stan dostępności, to pomaga potem w podejmowaniu decyzji związanych z rozwojem kolekcji. Kryteria, metody i miary w ocenie powinny być dobrane zgodnie z polityką prowadzenia biblioteki cyfrowej. Jeśli mamy do czynienia z kolekcją naukową, to będziemy mierzyli jej wpływ na rozwój edukacji, jeśli z kolekcją regionalną, to wpływ na rozwój turystyki w regionie, itp. Metody badań są znane, trzeba je tak dobrać, by sobie odpowiedzieć na pytanie*: Kto korzysta z BC, czego używa, jak i dlaczego?* Takie badania można zlecić firmie fachowej. Są także dostępne analizatory, technologie, które wspomagają takie obserwacje.

Zasada 7: Dobra biblioteka cyfrowa współdziała z innymi systemami, jest interoperacyjna[2].

Oznacza to, że oprogramowanie posiada zdolność komunikowania się z innym oprogramowaniem, wymienia metadane, daje możliwości importowania i eksportowania danych, pobierania zasobów, przeglądania przez roboty wyszukujące. Opiera się o standardy przesyłania danych.

[2] Zgodnie z ISO/IEC 2382-01 *Information Technology Vocabulary, Fundamental Terms (3.3.1.10)*, definicja interoperacyjności brzmi: „The capability to communicate, execute programs, or transfer data among various functional units in a manner that requires the user to have little or no knowledge of the unique characteristics of those units".

Zasada 8: Dobra kolekcja daje się integrować z zasobami użytkowników i jest wsparciem dla indywidualnej pracy.

Każdy końcowy użytkownik ma swój warsztat pracy, w którym musi często wykorzystać zasoby cyfrowe, jeśli BC jest przyjazna, to wykorzystanie jej obiektów nie jest trudne. Użytkownicy końcowi mogą bardzo wspomagać rozwój bibliotek, ponieważ dostarczają informacji zwrotnej i metodach pracy z zasobami. Niektóre BC dają dodatkowe możliwości użytkownikom, takie jak dodawanie słów kluczowych do obiektów (tagowanie), przesyłanie swoich zasobów cyfrowych do kolekcji, komentowania ich czy eksportowanie opisów bibliograficznych do własnych bibliografii załącznikowych. Oprogramowanie BC powinno zawsze mieć nowe funkcjonalności wspomagające i indywidualizujące prace użytkowników.

Zasada 9: Dobra kolekcja cyfrowa jest zawsze stabilna.

Instytucja zarządzająca kolekcją musi zapewnić nieustanny, wieloletni dostęp do zasobów, co oznacza zabezpieczenie kolekcji na każdym jej poziomie. Stabilizację AC musi zapewnić należy zarówno od strony organizacyjnej, technicznej, jak i finansowej. Jeśli zleca się wykonanie kolekcji cyfrowej instytucji zewnętrznej, należy sobie zapewnić wszelkie prawa do dysponowania tą kolekcją (plikami i metadanymi) w przyszłości. Zapewnienie stabilności, to także zapewnienie przechowywania długoterminowego nie tylko danych, ale i technicznych, merytorycznych oraz administracyjnych informacji, dzięki którym kolekcja istnieje (know-how).

OBIEKTY CYFROWE

Zasada 1: Dobry obiekt zapisywany jest w formacie, który zapewni mu powtórne wykorzystanie.

Istnieje ścisłe powiązanie między jakością, z jaką produkuje się obiekt, a jego czytelnością i elastycznością, dzięki którym można go potem powtórnie użyć lub wyeksportować na inną platformę. Przy czym format rozumiemy tu także jako sposób kodowania transmisji bitów. Zalecanych formatów jest tak wiele, jak wiele jest typów dokumentów (text, fotografia, audio, video)[3]. Dla każdego typu dokumentu zaleca się inny format. Trzeba w tym kontekście także widzieć inaczej obiekt cyfrowy uzyskany z digitalizacji a inaczej taki, który urodził się w postaci elektronicznej (born digital). W przypadku tych drugich należy pamiętać, żeby starać się przechowywać te materiały nie w formatach lokalnych, ale takich, które w przyszłości będą mogły być czytelne. Im bardziej popularny format, tym większe prawdopodobieństwo jego łatwej migracji do innego. Pliki matki powstałe w wyniku skanowania zaleca się dziś zabezpieczać w formacie tiff.

[3] Listę formatów zob. *Framework* 2007; Paradowski 2010; Płoszajski 2008.

Zasada 2: Dobry obiekt cyfrowy musi się dać zachować dla przyszłości. Musi być dostępny przez długi okres[4].

W tej chwili pracuje się nad dwiema strategiami zachowania obiektów: emulacją i migracją, które są ściśle od siebie zależne. Emulacja polega na bezkolizyjnym przechodzeniu do bardziej współczesnych systemów i urządzeń komputerowych. Migracja jest przenoszeniem obiektów między formatami czy „przepisywaniem" na nowszy format. Strategie te opracowywane są:

a) dla nośników informacji, które się zmieniają od dysków miękkich przez CD, DVD czy pamięci flash (pen-drive),
b) dla oprogramowania, które nieustannie się rozwija,
c) dla formatów i metod kodowania (od SGML do XML), które wybiera się jako standardy w danym momencie.

Im bardziej popularne są formaty czy oprogramowanie, tym większa szansa na opracowanie dla nich strategii przetrwania. Kopie matki zaś im bliższe są źródłowego formatu, tym lepiej, wszelkie przetwarzanie powoduje bowiem straty.

Zasada 3: Dobry obiekt jest zrozumiały, użyteczny na zewnątrz poza swoim lokalnym środowiskiem.

Taki obiekt można nazwać interoperacyjnym, wielokrotnie wykorzystywanym, przenoszonym między platformami. Taki obiekt musi być dobrze opisany za pomocą metadanych opisowych (z możliwością przełączania języków), technicznych oraz musi mieć jasny status, co do możliwości jego wykorzystania. Wszystkie techniczne parametry pozwalające skorzystać z obiektu powinny być dostępne i zrozumiałe dla innych. Obiekt powinien być łatwo odczytywalny przez meta wyszukiwarki [http://en.wikipedia.org/wiki/Metasearch_engine] czy przenoszony między archiwami cyfrowymi.

Zasada 4: Dobry obiekt powinien mieć stały, unikalny identyfikator, który pozwoli bezbłędnie rozpoznać obiekt i trafić do niego.

Identyfikator musi mieć taką strukturę, by nie można go było automatycznie duplikować, musi zawierać jakiś kod rozpoznawalny w systemie globalnym, musi być stały i nie może się zmieniać, nawet jeśli obiekt zmieni swoje miejsce w sieci [Digital Object Identifier: DOI, zob. Wikipedia.pl: http://pl.wikipedia.org/wiki/DOI_%28identyfikator_cyfrowy%29 lub Frequently Asked Questions about the DOI® System http://www.doi.org/faq.html#1]. Musi pozwalać na szybkie odnalezienie dokumentu, najlepiej przez jedno kliknięcie. KPBC ma identyfikator o składni: Identyfikator zasobu: *oai:kpbc.umk.pl:28732,* BCUWr Identyfikator zasobu: *oai:www.bibliotekacyfrowa.pl:21944.*

[4] Szczegóły zobacz w dokumencie elektronicznym NISO: http://framework.niso.org/node/19.

Zasada 5: Dobry obiekt powinien mieć potwierdzoną autentyczność.

Autentyczność dobrego obiektu może być określona na różne sposoby. Po pierwsze powinien mieć określone pochodzenie, nienaruszoną strukturę i wyraźnie oznaczone kolejne wersje. Obiekt nie może być uszkodzony lub zmieniony w sposób niedozwolony. W metadanych technicznych można określić pochodzenie obiektu (jest taka możliwość w schemacie METS). Do sfotografowanego obiektu można dodać znak wodny określający prawo własności do fizycznego dokumentu, ale nigdy tak, by zakłócał normalne korzystanie z niego. Zobacz znak wodny w KPBC na grafice: *Rzeczka w zimowej szacie* – Stanisława Rolicza (1913-1997) http://kpbc.umk.pl/dlibra/doccontent?id=27266&dirids=1. Jeżeli do obiektu upublicznionego online (w zmienionym formacie) dodamy element obcy, np. brakującą kartę, to należy to wyraźnie zaznaczyć w opisie bibliograficznym i podać źródło pochodzenia dodatku. W pliku matce nie wolno dokonywać takich ingerencji, bo nie odzwierciedla to stanu faktycznego.

Zasada 6: Dobry obiekt ma powiązane (wszyte) metadane.

Każdy obiekt powinien być dobrze opisany zarówno od strony merytorycznej (opis bibliograficzny), jak i administracyjnej czy technicznej. Ale szczególnie ważne jest by obiekt miał wszyte (zagnieżdżone, wbudowane w plik) metadane już w momencie tworzenia. Takie dane mają np. strony html – są to META tags (znaczniki meta)[5]. Są one istotne z punktu widzenia zarządzania obiektami oraz długotrwałego ich przechowywania. Najważniejsze dane to data utworzenia, pojemność itp.

METADANE

Zasada 1: Dobre metadane[6] mają odniesienie do standardów międzynarodowych związanych z konkretnymi rodzajami danych. Powinny być dobrane do użytkowników tych zasobów oraz zamierzonego, bieżącego i prawdopodobnego przyszłego wykorzystania obiektu cyfrowego.

Nie powinno się stosować własnych rozwiązań, które zamykają nam drogę do międzynarodowej wymiany danych. Powinno się własne dane opisać zgodnie ze standardami

[5] Znaczniki meta (ang. meta elements) – zbiór znaczników w sekcji nagłówkowej dokumentu używanych do opisu ich zawartości, stosowany przy tworzeniu stron internetowych w formacie HTML oraz XHTML. Za ich pomocą można przypisać danemu dokumentowi takie wartości, jak adres e-mail, imię i nazwisko autora strony, datę jej powstania słowa kluczowe itp. Można również zadeklarować kodowanie strony. Hasło z Wikipedia.pl: http://pl.wikipedia.org/wiki/Znaczniki_meta.

[6] Metadane – czyli „dane o danych", ich przykładem są np. klasyczne katalogi biblioteczne są metadanymi. Przy pomocy metadanych opisywane są dokumenty elektroniczne, w szczególności dokumenty dostępne poprzez sieci komputerowe, np. strony World Wide Web, a także dokumenty tworzące nowoczesne biblioteki cyfrowe. Jednym ze standardów metadanych jest Dublin Core Metadata Element Set (DCMES). Hasło z Wikipedia.pl: http://pl.wikipedia.org/wiki/Metadane.

przyjętymi powszechnie, by opis był zrozumiany przez każdego użytkownika. Słowniki, tezaurusy także powinny być dostosowane do naszych użytkowników, ale mieć odniesienie do międzynarodowej rodziny słowników, szczególnie jeśli chodzi o strukturę takich danych. Jeśli kolekcja jest mieszana, to stosuje się różne schematy metadanych, potem można je „opakować" za pomocą kontenera METS. Własne rozwiązania mogą być dodatkiem do standardowych. Polskie biblioteki cyfrowe przy tworzeniu opisów bibliograficznych dla obiektów stosują standard metadanych Dublin Core.

Zasada 2: Dobre metadane wspomagają interoperacyjność.

Oznacza to, że trzeba tak budować metadane, by mogły być łatwo importowane, eksportowane, wykorzystanie przy pobieraniu ich do metawyszukiwarek, zgodne z protokołami przesyłania danych (OAI PMH, Z39.50). Dobre metadane muszą być użyteczne w globalnym kontekście. To, co nie wydaje się ważne w lokalnym kontekście, bo jest dla wszystkich jasne, czasem musi być opisane w szerszym. To należy przewidywać. Ważnym jest także pamiętać o tym, że twórcy zasobów cyfrowych mogą współpracować w przyszłości z innymi dostawcami usług, którzy muszą dane standaryzować, implementować do własnych systemów czy łączyć z trzecimi zasobami.

Zasada 3: Dobre metadane stosują kontrolę autorytarną i narzędzia kontroli treści takie jak słowniki kontrolowane, spełniające wymagania użytkowników, co do opisu treści obiektów i łączenia ze sobą obiektów podobnych.

Słowniki haseł wzorcowych są bardzo potrzebne przy tworzeniu wszelkich metadanych, dzięki nim słownictwo jest spójne a opis kontrolowany, co do zgodności z innym takim samym. Klasyfikacje systematyczne porządkują wyszukiwanie danych, grupują je w struktury hierarchiczne, które są bardzo użyteczne w dostępie online. Słowniki dobiera się oczywiście do kolekcji i użytkowników, jacy z nich korzystają. Nie ma jednolitych słowników dla wszystkich na świecie, są tylko próby unifikacji niektórych. Wiele polskich bibliotek cyfrowych stosuje (ale niestety nie wszystkie) słownictwo kontrolowane dla nazw osobowych, korporatywnych, nazw geograficznych oraz haseł przedmiotowych. Wynika to z tego, że eksportują opisy ze swoich katalogów bibliotecznych, a te pobierają dane bądź to z NUKAT bądź Biblioteki Narodowej. Niestety nie jest to standard powszechnie stosowany.

Zasada 4: Dobre metadane zawierają jasne określenie warunków i zasad wykorzystywania obiektu cyfrowego.

Przy obiekcie należy umieścić w metadanych opisowych status prawny obiektu, ale nie digitalizatu a utworu intelektualnego czy artystycznego. Pomocny w tym jest poszerzony *Schemat Dublin Core (DC Terms),* który wprowadza wszystkie istotne elementy opisu dla pola PRAWA. W opisie powinno się zaznaczyć, czy obiekt jest pod

ochroną prawa (coprighted, not coprighted), czy należy już do domeny publicznej (public domain), kto jest właścicielem praw majątkowych do utworu, na jakiej licencji je udostępnił i na jaki zakres wykorzystania może liczyć użytkownik. Jeśli kolekcja jest jednolita, taka informacja może być zawarta na poziomie opisu kolekcji, jeśli nie, to na poziomie każdego obiektu.

Zasada 5: Dobre metadane techniczne i administracyjne ułatwiają długoterminowe zarządzanie obiektami w zasobach oraz ich zachowanie dla przyszłości.

Metadane techniczne pozwalają na zarządzanie zasobami, dostępem do nich, porządkowaniem ich, strukturalizowaniem, odświeżaniem, opisują pliki, ich format, objętość, wersjonowanie – są one niezbędne na przykład do rekonstrukcji zniszczonego obiektu, przesyłania go, czy przetwarzania. Administracyjne mówią, kto jest odpowiedzialny za obiekty, kto ma do nich dostęp, kto zrobił ich opis i ewentualne musi wnieść poprawki. Wszystkie one muszą być zawsze powiązane z obiektem, gdyż są niezbędne do zarządzania i zabezpieczania kolekcji.

Zasada 6: Dobre rekordy metadanych same w sobie są obiektami, powinny trzymać jakość, powinny być tworzone zgodnie z zasadami tworzenia dobrych obiektów, takimi jak autorytarność, autentyczność, archiwizacja, niezmienność, unikatowość oraz identyfikowalność.

Dlatego powinno się dokładnie opisać, kto, jaka instytucja jest odpowiedzialna za tworzenie metadanych, jakich używa standardów i czy zapewnie długoterminowe ich przechowywanie oraz kontrolę i melioracje. Instytucja musi stworzyć warunki do komunikowania się z użytkownikami, w razie gdyby chcieli skomentować jakość metadanych. Często wnoszą oni istotne poprawki do opisów, zauważają braki lub błędy. Wiele schematów metadanych zawiera elementy je opisujące w nagłówkach. Te metadane są niezwykle ważne wobec takich zagrożeń, jak pozycjonowanie stron, fałszerstwa czy spamowanie.

ZAKOŃCZENIE

Nie jest łatwo stworzyć dobry jakościowo produkt, ale o ile łatwiej, kiedy ma się gotowe recepty? Wypracowane standardy i zasady budowania dobrych zasobów cyfrowych są łatwo dostępne. Literatura zagraniczna na ten temat jest ogromna (część z niej można przejrzeć w pozycji bibliograficznej nr 1), wystarczy po nią sięgnąć.

Bibliografia

1. *A Framework of Guidance for Building Good Digital Collections – the 3rd edition* (2007) [dok. elektr.] Institute for Museum and Library Services http://framework.niso.org/ [odczyt: 2.02.2011].
2. Bednarek-Michalska, Bożena (2006). *Kujawsko-Pomorska Biblioteka Cyfrowa a standardy* [dok. elektr.] „EBIB – Elektroniczny Biuletyn Bibliotekarzy" nr 4(74) http://www.ebib.info/2006/74/michalska.php [odczyt: 12.03.2011].
3. *Digital Library Standards and Practices* (2008). Digital Library Federation Washington DC [dok. elektr.] http://www.diglib.org/standards.htm [odczyt: 28.08.2008].
4. *Good Practice Guide for Developers of Cultural Heritage Web Services* (2004). [dok. elektr.] UKOLN http://www.ukoln.ac.uk/interop-focus/gpg/ [odczyt: 10.03.2008].
5. *Metadata* (2007). Ukoln [dok. elektr.] http://www.ukoln.ac.uk/metadata/ [odczyt: 10.03.2008].
6. Paradowski, Dariusz oprac. i red. (2010). *Digitalizacja piśmiennictwa.* [Aut. Maciej Hnydka i in.; red. merytoryczna Agnieszka Konopka]. Warszawa: Biblioteka Narodowa, 212, [1] s.
7. Płoszajski, Grzegorz red. (2008). *Standardy w procesie digitalizacji obiektów dziedzictwa kulturowego* [aut. Grzegorz Płoszajski i Kazimierz Schmidt i in.]. Warszawa: Biblioteka Główna Politechniki Warszawskiej; Oficyna Wydawnicza Politechniki Warszawskiej, 247 s. [Wersja elektroniczna: http://bcpw.bg.pw.edu.pl/Content/1262/BG_Stand_w_proc_digit.pdf].
8. *The DOI System* (2008). The International DOI Foundation [dok. elektr.] http://www.doi.org/ [odczyt: 10.03.2008].
9. *The NINCH Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materiale* (2002). National Initiative for a Networked Cultural Heritage UK http://www.nyu.edu/its/humanities/ninchguide/index.html [odczyt: 10.03.2008].

Literatura na temat bibliotek cyfrowych – analiza bibliometryczna

Magdalena Wójcik

Wstęp

W pracy przedstawiono analizę bibliometryczną artykułów naukowych na temat bibliotek cyfrowych, opublikowanych w latach 2000-2010. Głównym celem jest ustalenie stanu badań w zakresie problematyki bibliotek cyfrowych w polskiej i zagranicznej literaturze przedmiotu. Wśród szczegółowych celów można wymienić:

- rozpoznanie głównych kierunków badań nad bibliotekami cyfrowymi,
- określenie dynamiki rozwoju publikacji na temat bibliotek cyfrowych,
- scharakteryzowanie podobieństw i różnic w sposobie przedstawiania problematyki bibliotek cyfrowych w polskiej i anglojęzycznej literaturze przedmiotu.

W celu rozpoznania głównych kierunków badań nad bibliotekami cyfrowymi oraz określenia dynamiki ich rozwoju posłużono się metodą analizy piśmiennictwa. Analizie poddano prace publikowane w wytypowanych polskich i zagranicznych czasopismach naukowych z dziedziny informacji naukowej i bibliotekoznawstwa. Podstawowym kryterium doboru czasopism był ich profil zgodny z przyjętymi celami pracy. Wśród wytypowanych czasopism znalazły się: *Journal of Documentation, D-Lib Magazine, Journal of the American Society for Information Science and Technology (JASIST), ARIADNE Magazine, International Journal of Digital Libraries, Journal of Digital Information, Przegląd Biblioteczny, Zagadnienia Informacji Naukowej* oraz *Praktyka i Teoria Informacji Naukowej i Technicznej*. W obrębie wytypowanych czasopism pod uwagę zostały następnie wzięte:

- ze względu na formę:
  - artykuły naukowe (w tym artykuły badawcze, teoretyczne i metodologiczne, problemowe itp.),
- ze względu na treść:

- prace, w których problematyka digitalizacji zasobów oraz ich udostępniania za pośrednictwem bibliotek cyfrowych stanowiła główny przedmiot opracowania lub stanowiła jego istotną część,
- wszystkie artykuły opublikowane w *D-Lib Magazine* (jako czołowym czasopiśmie związanym z problematyką bibliotek cyfrowych),

– ze względu na zasięg chronologiczny:

- publikacje z lat 2000-2010,

Wykluczone zostały:

– ze względu na formę:

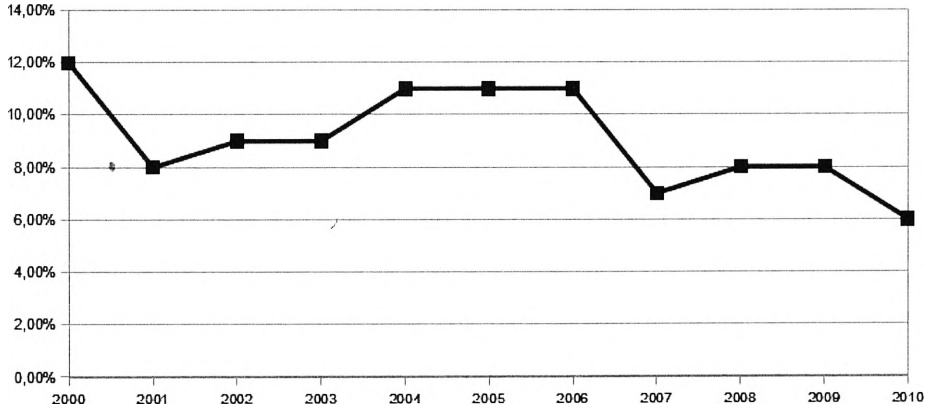
- ogłoszenia o konferencjach, sprawozdania z konferencji, recenzje książek, kroniki itp.,

Ogółem analizie poddano 690 artykułów z 9 czasopism naukowych, z czego 4% stanowiły publikacje polskie, 96% zagraniczne.

Dynamika rozwoju publikacji na temat bibliotek cyfrowych

Ogólna analiza częstotliwości publikowania artykułów poświęconych bibliotekom cyfrowym wykazała, że najwięcej publikacji z omawianego zakresu tematycznego ukazało się w roku 2000 oraz w latach 2004-2006, najmniej zaś w roku 2007 i 2010.

Wykres 1. Częstotliwość publikowania prac dotyczących bibliotek cyfrowych w wytypowanych polskich i zagranicznych czasopismach naukowych w latach 2000-2010



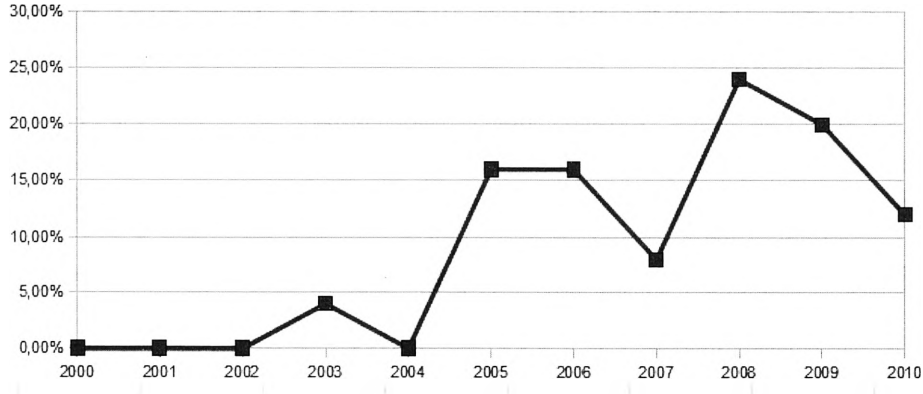
Źródło: opracowanie własne, 2011

W toku analizy dały się jednak zaobserwować duże różnice między częstotliwością i intensywnością publikowania w czasopismach polskich i zagranicznych.

Przeprowadzona analiza wytypowanych polskich czasopism naukowych pokazała, że problematyka bibliotek cyfrowych nie była tematem często poruszanym.

Charakterystyczne dla polskich czasopism okazało się naprzemienne występowanie okresów zwiększonego i zmniejszonego zainteresowania tematyką bibliotek cyfrowych. Problematyka zasobów cyfrowych zaczęła pojawiać się począwszy od 2003 r., jednak pierwszy, znaczący wzrost zainteresowania bibliotekami cyfrowymi nastąpił w dopiero latach 2005-2006. W 2007 r. miał miejsce widoczny spadek zainteresowania tą problematyką, by w 2008 r. osiągnąć najwyższy poziom. Począwszy od 2008 r. liczba publikacji zaczęła znów stopniowo spadać, osiągając w 2010 r. poziom nieco wyższy niż ten z 2007 r.

Wykres 2. Częstotliwość publikowania prac dotyczących bibliotek cyfrowych w wytypowanych polskich czasopismach naukowych w latach 2000-2010

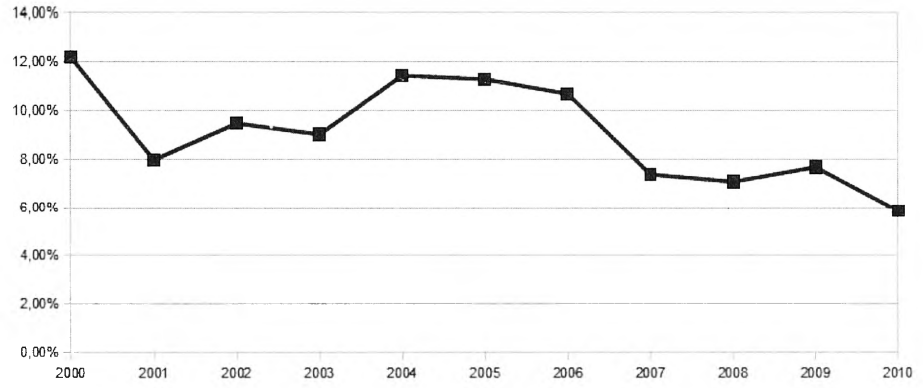


Źródło: opracowanie własne, 2011

W czasopismach anglojęzycznych podobnie jak w czasopismach polskich dały się zauważyć okresy zwiększonego i zmniejszonego zainteresowania problematyką bibliotek cyfrowych. Zaobserwowane różnice między okresami największej i najmniejszej intensywności publikowania nie były jednak tak duże jak w przypadku czasopism polskich.

W czasopismach zagranicznych problematyka bibliotek cyfrowych cieszyła się największą popularnością w 2000 r., najmniejszą zaś w 2010 r. W pozostałych latach intensywność publikowania była różna, wahania nie przekraczały jednak 4%.

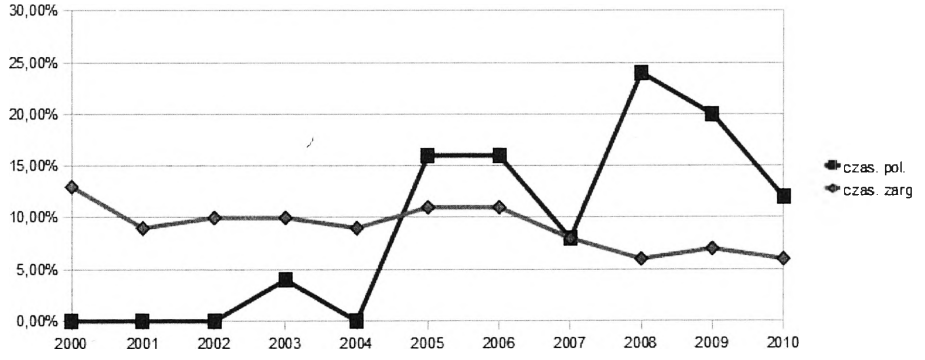
Wykres 3. Częstotliwość publikowania prac dotyczących bibliotek cyfrowych w wytypowanych anglojęzycznych czasopismach naukowych w latach 2000-2010



Źródło: opracowanie własne, 2011

Zestawienie danych dotyczących intensywności pojawiania się publikacji z omawianego zakresu tematycznego w polskich i anglojęzycznych czasopismach naukowych ujawniło różnice w występowaniu okresów największego i najmniejszego zainteresowania problematyką bibliotek cyfrowych.

Wykres 4. Częstotliwość publikowania prac dotyczących bibliotek cyfrowych w wytypowanych polskich i anglojęzycznych czasopismach naukowych w latach 2000-2010



Źródło: opracowanie własne, 2011

W okresie największego zainteresowania problematyką bibliotek cyfrowych na łamach czasopism zagranicznych, w analizowanych czasopismach polskich publikacje na ten temat pojawiały się sporadycznie lub wcale. Wspólny okres zainteresowania przypada na lata 2005-2006, wspólny jest również znaczący spadek liczby publikacji na temat bibliotek cyfrowych w 2007 r.

W 2008 r. w Polsce następuje okres największej intensywności publikowania (w 2008 r. powstaje ¼ wszystkich analizowanych polskich publikacji), podczas gdy w czasopismach anglojęzycznych liczba publikacji utrzymuje się w latach 2008-2009 na stałym, niskim poziomie. W 2010 r. zarówno w czasopismach polskich jak i zagranicznych intensywność publikowania widocznie spada.

Główne kierunki badań nad bibliotekami cyfrowymi

Przeprowadzona analiza pozwoliła stwierdzić, że zakresy tematyczne analizowanych publikacji można zaklasyfikować do dziewięciu głównych kategorii:

I: artykuły o charakterze ogólnym omawiające ideę i historię bibliotek cyfrowych,
II: użytkownicy bibliotek cyfrowych i ich zachowania informacyjne,
III: standardy, formaty danych, techniczne aspekty funkcjonowania bibliotek cyfrowych,
IV: doświadczenia, wdrożenia, projekty,
V: raporty o stanie digitalizacji w poszczególnych krajach, regionach itp.,
VI: zagadnienia prawne funkcjonowania bibliotek cyfrowych,
VII: zastosowanie bibliotek cyfrowych i ich wpływ na poszczególne obszary działalności,
VIII: projektowanie i ocena bibliotek cyfrowych,
IX: marketing, zarządzanie, aspekty polityczne i organizacyjne, strategie rozwoju bibliotek cyfrowych.

Najliczniej reprezentowane były prace omawiające techniczne aspekty funkcjonowania bibliotek cyfrowych (30%). Najczęściej poruszano problematykę narzędzi, platform i technologii wykorzystywanych przy tworzeniu nowoczesnych bibliotek cyfrowych oraz analizowano standardy opisu zasobów cyfrowych. Przykładem mogą być publikacje takie jak: *RSLP Collection Description* [Powell, Heaney, Dempsey 2000], *Mapping and Converting Essential Federal Geographic Data Committee (FGDC) Metadata into MARC21 and Dublin Core: Towards an Alternative to the FGDC Clearinghouse* [Chandler, Foley, Hafez 2000], *Making digital libraries effective: Automatic generation of links for similarity search across hyper-textbooks* [Melucci 2004], *Implementing Metadata in Digital Preservation Systems: The PREMIS Activity* [Lavoie 2004] czy *Metadane w Polsce na tle inicjatyw zagranicznych* [Ajdukiewicz-Tarkowska 2008].

Często pojawiały się również publikacje prezentujące doświadczenia, wdrożenia i projekty związane z tworzeniem bibliotek cyfrowych (21%). Za przykład mogą

posłużyć artykuły takie jak: *Lessons Learned: Digitization of Special Collections at The University of Iowa Libraries* [Hughes 2000], *How the Use of Standards Is Transforming Australian Digital Libraries* [Campbell 2004], *Large Scale Digitization of Oral History: A Case Study* [Weig, Kopana, Lybarger 2007], i *Doświadczenia poznańskie w budowie biblioteki cyfrowej na przykładzie prac Biblioteki Kórnickiej Polskiej Akademii Nauk* [Jastrząb 2008].

Stosunkowo często (14%), szczególnie w ostatnich latach, publikowano prace dotyczące metod projektowania i ewaluacji bibliotek cyfrowych. Przykładem są tutaj prace, takie jak: *Evaluation of digital libraries* [Fuhr 2008], *Techniques to audit and certify the long-term integrity of digital archives* [Song, JaJa 2009] czy *Usability Inspection of Digital Libraries* [Paterson, Low 2010].

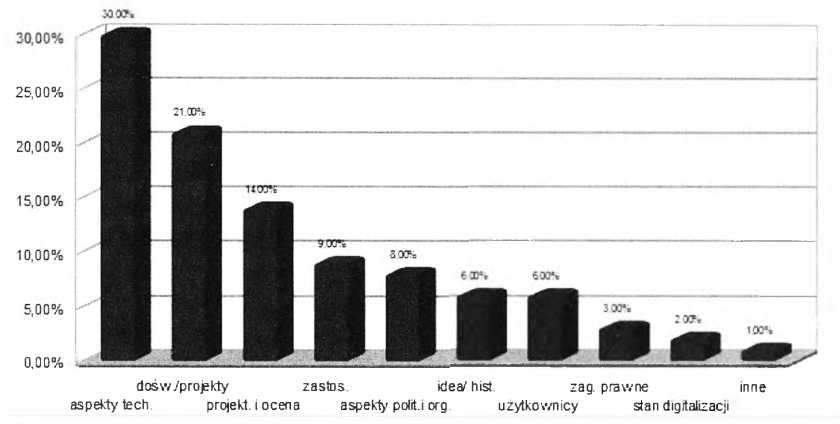
Nieco rzadziej (9%) pojawiały się prace związane z wykorzystaniem bibliotek cyfrowych i ich rolą w różnych obszarach działalności. W tej grupie, najczęściej pojawiały się artykuły omawiające zastosowanie zasobów bibliotek cyfrowych w edukacji, kulturze i nauce oraz w odniesieniu do osób niepełnosprawnych. Przykładem mogą być prace, takie jak: *Digital library resources as a basis for collaborative work* [Wilensky 2000], *Digital Libraries and Education: Trends and Opportunities* [Roes 2001], *One digital library, two undergraduate classes, and four learning modules: Uses of a digital library in classrooms* [Pan, 2006] czy *Cyfrowe książki i bilioteki – próba oceny przydatności dla osób z dysfunkcją wzroku* [Durska 2008].

Interesującą, choć niezbyt liczną (8%) grupę stanowiły prace prezentujące polityczne, organizacyjne i marketingowe aspekty prowadzenia bibliotek cyfrowych. Przykładem takiego podejścia są prace, takie jak: *Information Wants to Be Free... Then Who's Going to Pay for It?* [Kaser 2000], *Virtual Remote Control: Building a Preservation Risk Management Toolbox for Web Resources* [McGovern 2004] czy *Systematic planning for digital preservation: evaluating potential strategies and building preservation plans* [Becker 2009].

Zdecydowanie rzadko pojawiały się prace o charakterze czysto teoretycznym, prezentujące ogólną ideę bibliotek cyfrowych, ich historię, perspektywy, takie jak na przykład: *Digital Libraries and the Problem of Purpose* [Levy 2000], *Digital libraries and reference services: present and future* [Chowdhury 2002] czy *The digital library as place* [Pomerantz, Marchionini 2007]. Rzadko poruszano również problematykę użytkowników bibliotek cyfrowych i ich zachowań informacyjnych. Wśród przykładowych artykułów można wymienić: *A library or just another information resource? A case study of users' mental models of traditional and digital libraries* [Makri 2007], *Effects of interaction design in digital libraries on user interactions* [Zhang 2008], oraz *Student digital information-seeking behaviour in context* [Nicholas 2009].

Najmniej licznie reprezentowane były prace dotyczące aspektów prawnych tworzenia kolekcji cyfrowych (3%) oraz teksty prezentujące w sposób przeglądowy stan digitalizacji poszczególnych krajów, regionów, typów instytucji (2%). Niewielka grupa tekstów (1%) została przyporządkowana do kategorii „inne".

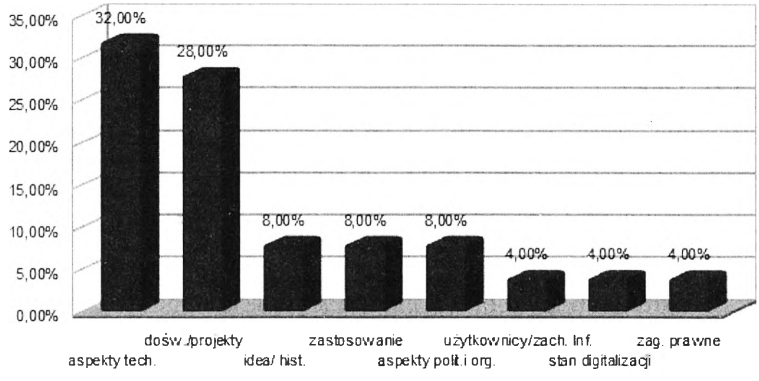
Wykres 5. Główne kierunki analizy problematyki bibliotek cyfrowych na łamach wytypowanych polskich i zagranicznych czasopismach naukowych w latach 2000-2010



Źródło: opracowanie własne, 2011

Na łamach polskich czasopism dominowały prace omawiające techniczne aspekty funkcjonowania bibliotek cyfrowych (32%) oraz opisujące projekty, wdrożenia i doświadczenia twórców cyfrowych kolekcji (28%). Zdecydowanie rzadziej pojawiały sie prace o charakterze ogólnym, związane z ideą i historią bibliotek cyfrowych (8%), artykuły dotyczące ich roli i zastosowania (8%) oraz prace poruszające aspekty polityczne i organizacyjne (8%). Sporadycznie publikowano prace dotyczące użytkowników bibliotek cyfrowych (4%), teksty omawiające zagadnienia prawne związane z umieszczaniem dokumentów w Internecie (4%) oraz raporty o postępach w digitalizacji zbiorów polskich instytucji (4%).

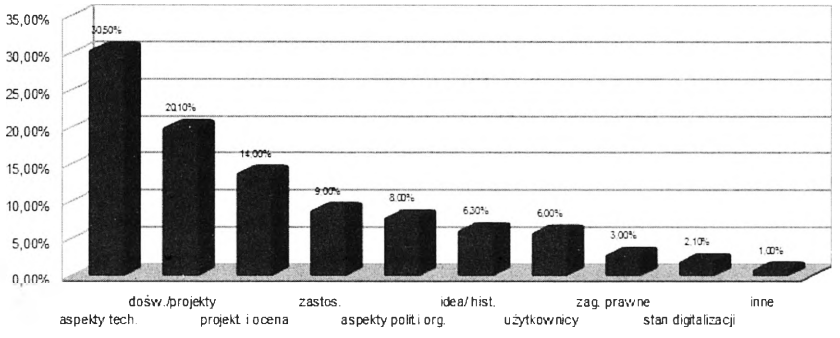
Wykres 6. Główne kierunki analizy problematyki bibliotek cyfrowych na łamach wytypowanych polskich czasopismach naukowych w latach 2000-2010



Źródło: opracowanie własne, 2011

Na łamach analizowanych czasopism zagranicznych dominowały prace dotyczące technicznych aspektów tworzenia zasobów cyfrowych (30,50%) oraz opisujące wdrożenia, zrealizowane projekty i doświadczenia twórców bibliotek (20%). Często pojawiały się również artykuły poruszające problematykę projektowania i oceny kolekcji cyfrowych (14%). Rzadziej publikowano prace na temat różnych obszarów wykorzystania zdigitalizoawanych zasobów (9%), artykuły omawiające polityczne, organizacyjne i marketingowe aspekty działania bibliotek cyfrowych (8%) oraz teksty poświęcone użytkownikom informacji (6%). Sporadycznie publikowano prace omawiające zagadnienia prawne związane z umieszczaniem dokumentów w Internecie (3%) oraz raporty o postępach w digitalizacji (2%).

Wykres 7. Główne kierunki analizy problematyki bibliotek cyfrowych na łamach wytypowanych zagranicznych czasopismach naukowych w latach 2000-2010

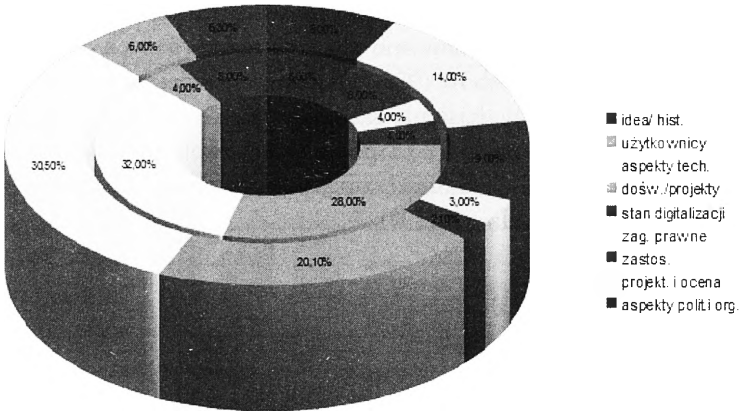


Źródło: opracowanie własne, 2011

Zestawienie danych dotyczących częstości poruszania wytypowanych zakresów tematycznych na łamach analizowanych czasopism polskich (wewnętrzna część diagramu) i zagranicznych (zewnętrzna cześć diagramu) ujawniło pewne różnice.

W czasopismach polskich procentowo częściej (choć nieporównywalnie rzadziej biorąc pod uwagę liczby bezwzględne) publikowano prace referujące praktyczne doświadczenia twórców bibliotek cyfrowych oraz analizujące aspekty techniczne. Nieco większą popularnością cieszyły się w Polsce również teksty o charakterze ogólnym oraz prace omawiające zagadnienia prawne i postępy w digitalizacji. W polskich czasopismach nieco rzadziej niż w zagranicznych publikowano natomiast prace dotyczące użytkowników bibliotek cyfrowych oraz prace dotyczące praktycznego wykorzystania cyfrowych kolekcji. Stosunkowo popularne w czasopismach zagranicznych publikacje dotyczące projektowania i oceny bibliotek cyfrowych (14%) nie pojawiały się w czasopismach polskich w ogóle.

Wykres 8. Główne kierunki analizy problematyki bibliotek cyfrowych na łamach wytypowanych polskich i zagranicznych czasopismach naukowych w latach 2000-2010 (zestawienie)



Źródło: opracowanie własne, 2011

Tendencje w rozwoju badań nad bibliotekami cyfrowymi

Uogólniając wyniki przeprowadzonej analizy można sformułować wstępne wnioski dotyczące głównych tendencji w rozwoju piśmiennictwa na temat bibliotek cyfowych.

Analiza ujawniła, że zarówno w czasopismach polskich, jak i zagranicznych dominuje techniczne i praktyczne podejście do tematu tworzenia zasobów cyfrowych. Podejście to jest charakterystyczne szczególnie dla pierwszych okresów zainteresowania tą problematyką, jednak utrzymuje się nadal a liczba publikacji omawiających techniczne aspekty tworzenia bibliotek cyfrowych nie maleje. Nadal popularna jest również praktyka opisywania przez poszczególne instytucje czy konsorcja realizowanych projektów, wdrożeń i pomysłów.

Na łamach zagranicznych czasopism, coraz częściej pojawiają się artykuły z zakresu projektowania i oceny bibliotek cyfrowych. Szczegółowej analizie poddaje się zarówno użyteczność i przyjazność bibliotek, jak i narzędzia ich ewaluacji. Na razie nieliczną, ale rozwijającą się grupę publikacji stanowią również artykuły poruszające ekonomiczne, polityczne i organizacyjne aspekty tworzenia kolekcji cyfrowych.

Przeprowadzona analiza ujawniła niedobór publikacji o charakterze przeglądowym, ogólnym i teoretycznym. Niewystarczająca wydaje się również liczba publikacji poświęconych użytkownikom bibliotek cyfrowych i ich zachowaniom informacyjnym, choć pewna grupa zagadnień z tego zakresu jest poruszana przy okazji omawiania innych tematów np. obszarów wykorzystania zasobów cyfrowych. Zdziwienie

może budzić mała liczba publikacji poruszających istotny problem regulacji prawnych, lecz podobnie jak w przypadku użytkowników informacji, część zagadnień prawnych jest omawiana na marginesie rozważań o ekonomicznych i organizacyjnych aspektach działania bibliotek cyfrowych.

Wnioski

Wyniki przeprowadzonej analizy pokazały, że w latach 2000-2010 problematyka bibliotek cyfrowych była tematem stosunkowo często poruszanym na łamach wybranych polskich i zagranicznych czasopism bibliotekoznawczych.

Ogólna analiza częstotliwości publikowania artykułów poświęconych bibliotekom cyfrowym wykazała, że najwięcej publikacji z omawianego zakresu tematycznego ukazało się w 2000 r. oraz w latach 2004-2006, najmniej zaś w roku 2007 i 2010. Przeprowadzona analiza ujawniła również, szczególnie w odniesieniu do czasopism polskich, cykliczność okresów zwiększonego i zmniejszonego zainteresowania tym tematem.

Zakres poruszanych tematów obejmuje przede wszystkim techniczne i praktyczne aspekty tworzenia kolekcji cyfrowych. Szczegółowym analizom poddawane są narzędzia tworzenia bibliotek cyfrowych i metody opisu ich zasobów. Często pojawiają się również relacje z realizacji projektów, opisy wdrożeń oraz refleksje twórców bibliotek cyfrowych. Przeprowadzona analiza ujawniła jednak niedobór opracowań na temat użytkowników bibliotek cyfrowych oraz niedostatek prac o charakterze teoretycznym i przeglądowym.

Bibliografia:

1. Ajdukiewicz-Tarkowska, Anna (2008). *Metadane w Polsce na tle inicjatyw zagranicznych.* „Zagadnienia Informacji Naukowej" nr 1, s. 59-84.
2. Becker, Christoph [i in.] (2009). *Systematic planning for digital preservation: evaluating potential strategies and building preservation plans.* „International Journal of Digital Libraries" vol. 10 nr 4, s. 133-157.
3. Campbell, Debbie (2004). *How the use of standards is transforming australian digital libraries.* „ARIADNE Magazine issue" 41 [dok. elektr.] http://www.ariadne.ac.uk/issue41/campbell/ [odczyt: 30.11.2011].
4. Chandler, Adam; Foley, Dan; Hafez, Alaaeldin (2000). *Mapping and converting essential Federal Geographic Data Committee (FGDC) Metadata into MARC21 and Dublin Core: towards an alternative to the FGDC clearinghouse.* „D-Lib Magazine" vol. 6 nr 1 [dok. elektr.] http://www.dlib.org/dlib/january00/chandler/01chandler.html [odczyt: 30.11.2011].
5. Chowdhury, Gobinda (2002). *Digital libraries and reference services: present and future.* „Journal of Documentation" vol. 58 nr 3, s. 258-283.
6. Durska, Anna (2008). *Cyfrowe książki i biblioteki – próba oceny przydatności dla osób z dysfunkcją wzroku.* „Zagadnienia Informacji Naukowej" nr 2, s. 29-43.
7. Fuhr, Norbert [i in.] (2008). *Evaluation of digital libraries.* „International Journal of Digital Libraries" vol 8 nr 1, s. 21-38.

8. Hughes, Carol Ann (2000). *Lessons learned: digitization of special collections at the University of Iowa Libraries.* „D-Lib Magazine" vol. 6 nr 6 [dok. elektr.] http://www.dlib.org/dlib/june00/hughes/06hughes.html [odczyt: 30.11.2011].
9. Jastrząb, Łukasz (2008). *Doświadczenia poznańskie w budowie biblioteki cyfrowej na przykładzie prac Biblioteki Kórnickiej Polskiej Akademii Nauk.* „Przegląd Biblioteczny" R. 76 z. 1, s. 79-85.
10. Kaser, Richard (2000). *If information wants to be free... then who's going to pay for it?* „D-Lib Magazine" vol. 6 nr 5 [dok. elektr.] http://www.dlib.org/dlib/may00/kaser/05kaser.html [odczyt: 30.11.2011].
11. Lavoie, Brian F. (2004). *Implementing metadata in digital preservation systems: the PREMIS activity.* „D-Lib Magazine" vol. 10 nr 4 [dok. elektr.] http://www.dlib.org/dlib/april04/lavoie/04lavoie.html [odczyt: 30.11.2011].
12. Levy, David (2000). *Digital libraries and the problem of purpose.* „D-Lib Magazine"vol. 6 nr 1 [dok. elektr.] http://www.dlib.org/dlib/january00/01levy.html [odczyt: 30.11.2011].
13. McGovern, Nancy [i in.] (2004). *Virtual remote control: building a preservation risk management toolbox for web resources.* „D-Lib Magazine" vol. 10 nr 4 [dok. elektr.] http://www.dlib.org/dlib/april04/mcgovern/04mcgovern.html [odczyt: 30.11.2011].
14. Makri, Stephann [i in.] (2007). *A library or just another information resource? A case study of users' mental models of traditional and digital libraries.* „JASIS" *vol.* 58 nr 3, s. 433-445.
15. Melucci, Massimo (2004). *Making digital libraries effective: automatic generation of links for similarity search across hyper-textbooks.* „JASIS" vol. 55 nr 5, s. 414-430.
16. Nicholas, David [i in.] (2009). *Student digital information-seeking behavior in context.* „Journal of Documentation" vol. 65 nr 1, s. 106-132.
17. Pan, Bing [i in.] (2006). *One digital library, two undergraduate classes, and four learning modules: Uses of a digital library in classrooms.* „JASIS" vol.57 nr 10, s. 1315-1325.
18. Paterson, Lorraine; Low, Boon (2010). *Usability inspection of digital libraries.* „ARIADNE Magazine" nr 63 [dok. elektr.] http://www.ariadne.ac.uk/issue63/paterson-low/ [odczyt: 30.11.2011].
19. Pomerantz, Jefferey; Marchionini, Garry (2007). *The digital library as place.* „Journal of Documentation" vol. 63 nr 4, s. 505-533.
20. Powell, Andy; Heaney, Michael; Dempsey, Lorcan (2000). *RSLP Collection Description.* „D-Lib Magazine" vol. 6 nr 9 [dok. elektr.] http://www.dlib.org/dlib/september00/powell/09powell.html [odczyt: 30.11.2011].
21. Roes, Hans (2001). *Digital libraries and education: trends and opportunities.* „D-Lib Magazine", vol. 7 nr 7/8 [dok. elektr.] http://www.dlib.org/dlib/july01/roes/07roes.html [odczyt: 30.11.2011].
22. Song, Sangchul; JaJa, Joseph (2009). *Techniques to audit and certify the long-term integrity of digital archives.* „International Journal of Digital Libraries" vol. 10 nr 2-3, s. 123-131.
23. Weig, Eric; Terry Kopana, Kathryn Lybarger (2007). *Large scale digitization of oral history: a case study.* „D-Lib Magazine" vol. 13 nr 5/6 [dok. elektr.] http://www.dlib.org/dlib/may07/weig/05weig.html [odczyt: 30.11.2011].
24. Wilensky, Robert (2000). *Digital library resources as a basis for collaborative work.* „JASIS" vol. 51 nr 3, s. 228-245.
25. Zhang, Xiangmin [i in.] (2008). *Effects of interaction design in digital libraries on user interactions.* „Journal of Documentation" vol. 63 nr 3, s. 438-463.

Część 2

Ludzie związani z bibliotekami cyfrowymi

Użytkownicy polskich bibliotek cyfrowych

Jolanta Mazurek

W pracy przedstawiono problematykę użytkowników bibliotek cyfrowych z akcentem na użytkowników polskich bibliotek cyfrowych. Zasygnalizowano także zagadnienie 'Nie-Użytkowników' oraz omówiono kategorie odbiorców.

Wstęp

Biblioteka cyfrowa, funkcjonująca jako jedna z form usług biblioteki tradycyjnej (obok np. takich ofert jak OPAC, różnego typu bazy danych, repozytoria czy usługa „Zapytaj bibliotekarza"), wykorzystuje Internet jako medium w komunikacji z Użytkownikiem jej zasobów. Przestrzeń Internetu stanowi naturalne środowisko funkcjonowania tak zorganizowanej formy dostępu do informacji, w którym każdy potencjalny Użytkownik biblioteki cyfrowej jest przede wszystkim Internautą, wirtualnym odbiorcą globalnych zasobów informacyjnych, ale nie każdy Internauta jest Użytkownikiem biblioteki. Co więcej, nie każdy Internauta, który trafia na stronę biblioteki cyfrowej, jest jej Użytkownikiem.

W tak przyjętej koncepcji kluczowym zadaniem staje się wyznaczenie pewnych warunków, które stanowić będą niejako punkt wyjścia dla skategoryzowania odbiorców usług dostarczanych przez polskie biblioteki cyfrowe. Pamiętać jednakże należy, że propozycja usystematyzowania Użytkowników bibliotek cyfrowych, ze względu na dynamiczny rozwój bibliotek cyfrowych, może stanowić o jej labilności. Wydaje się, że już blisko 10 lat funkcjonowania w społeczeństwie takiej formy usług bibliotek instytucjonalnych, pozwala na formułowanie pewnych podsumowań i wyciąganie wniosków. Z drugiej jednakże strony silna zależność bibliotek cyfrowych od pewnych formalnych uwarunkowań[1], w jakich one funkcjonują, ma bezpośredni wpływ na ich

[1] Np. obowiązujących przepisów prawnych, czy też wysokości środków przeznaczanych na rozwój bibliotek cyfrowych.

rozwój i kształt, co pośrednio wpływa także na formę społeczności skupionej wokół tego typu serwisów.

Użytkownik a NIE-Użytkownik w bibliotece cyfrowej

Witryny bibliotek cyfrowych, podobnie jak i innego typu strony WWW, poprzez wykorzystanie dodatkowych narzędzi, mają możliwość monitorowania i gromadzenia danych odnośnie do wejść na stronę danej biblioteki cyfrowej. Cechą tego rodzaju danych gromadzonych przez system jest to, że dokładnie rejestrują one praktycznie każdą, nawet najkrótszą wizytę. Przykłady danych prezentujących wejścia na stronę Wielkopolskiej Biblioteki Cyfrowej [http://www.wbc.poznan.pl/] pokazują regularność takich zachowań, potwierdzoną bardzo krótkim czasem pobytu na tej witrynie. Na rys.1 zaprezentowano czas trwania wizyt na stronach WBC [dane za AWStats dla serwisu WBC] w czerwcu 2010 r., gdzie aż ponad 63% to wizyty trwające pół minuty lub krócej.

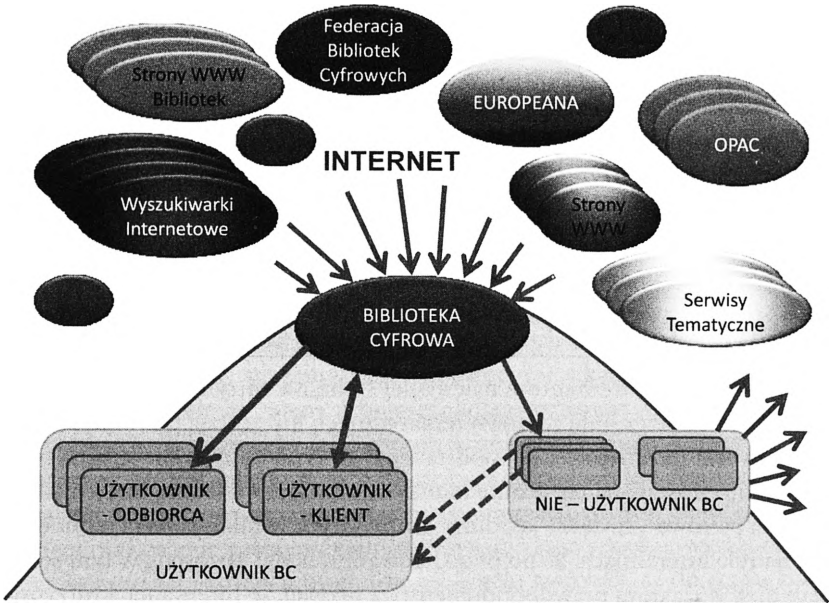
Visits duration		
Number of visits: 101152 - Average: 434 s	Number of visits	Percent
0s-30s	64000	63.2 %
30s-2mn	9188	9 %
2mn-5mn	6184	6.1 %
5mn-15mn	7124	7 %
15mn-30mn	4888	4.8 %
30mn-1h	5016	4.9 %
1h+	4594	4.5 %
Unknown	158	0.1 %

Rys. 1. Zestawienie czasu przebywania na stronach internautów odwiedzających WBC w czerwcu 2010 r.

Dla porównania, to samo źródło zarejestrowało, że w styczniu 2010 r. takich najkrótszych wizyt w WBC było ponad 55%, a w marcu 2010 r. ponad 58%.

Tak znacząca ilość bardzo krótkich wizyt na stronie biblioteki cyfrowej pozwala założyć, że na tego typu witryny trafiają także przypadkowi Internauci, którzy w żaden sposób nie są zainteresowani korzystaniem z usług biblioteki cyfrowej, a więc nie są jej Użytkownikami[2]. Sytuacja taka wymaga postawienia wyraźnej granicy pomiędzy właśnie tymi Internautami, którzy jedynie WIZYTUJĄ witrynę biblioteki a tymi, którzy KORZYSTAJĄ z usług oferowanych przez bibliotekę cyfrową, i wyróżnić Internautę wizytującego czyli NIE-Użytkownika biblioteki cyfrowej oraz Internautę korzystającego, czyli właściwego Użytkownika biblioteki cyfrowej. Przyjmując taką koncepcję, na rysunku 2 zobrazowano potencjalny model środowiska Internautów tworzącego społeczność biblioteki cyfrowej.

[2] Przyjmując założenie, że 30 sekund i krócej to czas zbyt krótki, aby skorzystać chociażby z jednej usługi oferowanej przez bibliotekę cyfrową.



Rys. 2 Model środowiska społeczności skupionej wokół biblioteki cyfrowej

Internauta wizytujący, NIE-Użytkownik biblioteki cyfrowej

Ten typ Internauty określić można jako Użytkownika globalnych zasobów Internetu, który odwiedził witrynę biblioteki cyfrowej i został zarejestrowany poprzez system monitorujący ruch na tej stronie, ale z subiektywnych, sobie znanych przyczyn, nie skorzystał z żadnej z usług oferowanych przez bibliotekę cyfrową i opuścił serwis. Okoliczności, dla których trafił i dla których nie pozostał na stronie, na podstawie danych monitorujących ruch na stronach WWW bibliotek cyfrowych, nie da się ustalić. Stwierdzić jedynie można, że najprawdopodobniej zarejestrowany krótki czas pobytu świadczy o przekierowaniu go na tę stronę, która w sposób jednoznaczny nie spełniła jego oczekiwań wyszukiwawczych. Z dużym prawdopodobieństwem przypuszczać można także, że trafił na tę stronę niezamierzenie i całkowicie przez przypadek. Zgromadzone dane [za Google Analytics dla serwisu WBC] potwierdzają (Tab. 1), że praktycznie jedynie około 40% trafiających na stronę bibliotek cyfrowych trafia tam bezpośrednio (Direct Traffic) – a wiec można założyć, że z pewnością są to celowi Użytkownicy bibliotek cyfrowych a nie tylko przypadkowi Internauci, wizytujący te strony.

Tabela 1. Dane zarejestrowane dla serwisu WBC kolejno w miesiącach marzec-czerwiec 2011 r. odnotowujące źródło dostępu do Wielkopolskiej Biblioteki Cyfrowej

Traffic Sources	Visits			
	03.2011	04.2011	05.2011	06.2011
Reffering Sites	33 140 (43,31%)	27 687 (41,20%)	24 659 (37,36%)	22 314 (36,10%)
Direct Traffic	28 363 (37,06%)	25 671 (38,20%)	26 506 (40,16%)	22 160 (35,85%)
Search Engines	14 999 (19,60%)	13 830 (20,58%)	14 827 (22,47%)	17 330 (28,04%)
Other	23 (0,03%)	8 (0,01%)	3 (0,00%)	7 (0,01%)

Pozostała większość Internautów najczęściej trafia na witryny biblioteki cyfrowej przede wszystkim poprzez linki z witryn tematycznych lub zostają tam skierowani poprzez wyszukiwarki internetowe. Na podstawie tego typu danych nie można określić jak wielu z nich pozostaje na tej stronie i korzysta chociażby z jednej funkcji jaką oferuje biblioteka cyfrowa[3], czyli dla jak dużej grupy Użytkowników oferta biblioteki cyfrowej jest na tyle interesująca, że po przekierowaniu, nadal pozostają w tym serwisie. Można jedynie z dużą dozą prawdopodobieństwa przyjąć, że Internauci, którzy zostali odesłani na stronę danej biblioteki cyfrowej z innej biblioteki cyfrowej (lub serwisów związanych z funkcjonowaniem bibliotek np. FBC[4], Europeana[5], WorldCat[6], NUKAT[7], KaRo[8], katalogi typu OPAC[9]) trafiają na tę stronę w wyniku świadomie podjętych przez nich czynności i mają zamiar skorzystać z usług tej biblioteki – a więc stanowią również grupę celowych jej Użytkowników.

Internauta korzystający, Użytkownik biblioteki cyfrowej

W odróżnieniu od powyżej wyróżnionego typu Internauty, który jedynie przypadkowo trafia na stronę biblioteki cyfrowej, **Użytkownikiem biblioteki cyfrowej** okre-

[3] Takich informacji dostarczyć by mogła analiza logów. Metoda ta ze względu na złożoność oraz koszty przedsięwzięcia, nigdy nie została przeprowadzona w odniesieniu do żadnej z funkcjonujących bibliotek cyfrowych w Polsce.

[4] Federacja Bibliotek Cyfrowych (FBC) – jest serwisem umożliwiającym przeszukanie zawartości wszystkich polskich regionalnych i instytucjonalnych bibliotek cyfrowych, oraz serwisów, w których zastosowane zostały otwarte protokoły wymiany danych. Serwis dostępny jest pod adresem: http://www.fbc.pionier.net.pl/.

[5] Europeana – Europejska Biblioteka Cyfrowa oferuje dostęp do zasobów europejskiego dziedzictwa kulturowego. Serwis dostępny jest pod adresem: http://www.europeana.eu/.

[6] WorldCat – największy na świecie katalog zasobów i usług bibliotecznych. Serwis dostępny jest pod adresem: http://www.worldcat.org/.

[7] NUKAT – centralny katalog polskich bibliotek akademickich i naukowych. Serwis dostępny jest pod adresem: http://www.nukat.edu.pl/.

[8] KaRo – wyszukiwarka zasobów bibliotecznych z katalogów kilkudziesięciu bibliotek naukowych w Polsce. Serwis dostępny jest pod adresem: http://karo.umk.pl/.

[9] OPAC – ang. Online public access katalog. – moduł katalogu bibliotecznego, dostępnego poprzez strony WWW.

ślić można Internautę, który najczęściej w świadomy (choć nie zawsze) sposób dotarł na witrynę biblioteki cyfrowej i skorzystał przynajmniej z jednej funkcji oferowanej przez tego typu serwis. W odróżnieniu od NIE-Użytkownika charakteryzuje się on określoną aktywnością w bibliotece cyfrowej, zorientowaną na zaspokojenie swoich potrzeb informacyjnych.

W tak przyjętym skategoryzowaniu Użytkownika biblioteki cyfrowej, należy zwrócić uwagę na wyraźną zależność pomiędzy biblioteką cyfrową traktowaną jako zaawansowaną platformę dostarczającą różnego typu usługi a interakcją podjętą ze strony odbiorcy tych usług. Produktem finalnym, dostarczanym przez bibliotekę jest obiekt cyfrowy o unikalnym identyfikatorze, mogący być samodzielną publikacją w obrębie biblioteki, lub dla wygodniejszej nawigacji po zasobach powiązanych formalnie lub tematycznie, prezentowany w postaci publikacji grupowych czy kolekcji tematycznych. Docelową zatem usługą biblioteki cyfrowej jest organizacja dostępu do tego obiektu. W oczywisty więc sposób Użytkownik, który korzysta z takiego zasobu jest tym, którego obecność w bibliotekach cyfrowych jest najbardziej pożądana. Przebywanie takiego typu Użytkownika w bibliotece cyfrowej charakteryzuje się zaistnieniem praktycznie pełnego cyklu użytkowania tego serwisu, wyrażonego zastosowaniem najbardziej istotnych modułów biblioteki cyfrowej, aż do etapu końcowego, tj. wyświetlenia obiektu cyfrowego.

Nie należy jednakże pomijać faktu, że Użytkownik biblioteki cyfrowej to także Internauta, którego obecność na witrynie biblioteki nie zawsze jest zakończona wyświetlaniem obiektu cyfrowego, a więc nie jest równoznaczna z faktem korzystania z zasobów tej biblioteki. Użytkownik ten w sposób świadomy może odwiedzać bibliotekę cyfrową także w celu skorzystania z innych usług, które ta biblioteka mu oferuje np. w celu monitorowania nowości pojawiających się w bibliotece cyfrowej, kompletowania danych bibliograficznych, czy pozyskania informacji na temat statusu dostępności zasobów planowanych do udostępnienia w bibliotece cyfrowej.

Do tej kategorii Użytkowników bibliotek cyfrowych trzeba zaliczyć także tych, którzy chcieli skorzystać z zasobów bibliotek cyfrowych ale z rozmaitych powodów nie byli w stanie tego uczynić. Użytkownicy biblioteki cyfrowej reprezentujący różne środowiska, w odmienny sposób korzystają z biblioteki cyfrowej, posiadają także różny stopień umiejętności w posługiwaniu się tego typu serwisami. Nie zawsze w związku z tym podjęte przez nich czynności prowadzą do wyświetlenia czy dotarcia do poszukiwanego obiektu. Nierzadko powodem takiego stanu rzeczy jest również brak pożądanego zasobu w bibliotece cyfrowej czy ograniczenia uniemożliwiające wyświetlenie zawartości publikacji elektronicznej[10]. Ten typ Użytkownika biblioteki cyfrowej można porównać przykładowo z czytelnikiem biblioteki tradycyjnej, który odwiedza

[10] Przykładem tego typu Użytkowników w WBC są najczęściej studenci spoza poznańskich uczelni, którzy poszukują dostępu do współczesnych skryptów. Lektury te, w poszanowaniu przepisów prawa autorskiego, są udostępniane jedynie studentom tych uczelni, którym macierzysta biblioteka wykorzystując możliwości biblioteki cyfrowej, stworzyła taki model dostępu do poszukiwanych treści. Obecność tego typu Użytkowników często jest akcentowana ich mailowym kontaktem z bibliotekarzami czy też administratorami WBC.

bibliotekę ale np. powodu braku dostępności wydawnictwa, nie dociera do pożądanej treści.

Modele funkcjonowania Użytkowników w bibliotece cyfrowej

Użytkownik-Odbiorca

Użytkownicy bibliotek cyfrowych funkcjonują w środowiskach tych serwisów w różnorodny sposób. W odróżnieniu od bibliotek tradycyjnych, nie mogąc liczyć na bezpośrednią i natychmiastową pomoc bibliotekarza w dotarciu do potrzebnych im informacji, są zobligowani do samodzielnej mniej lub bardziej zaawansowanej aktywności w tym kierunku. Niezależnie od rezultatów podjętych poczynań, część Użytkowników biblioteki cyfrowej pozostaje bierna i nie wyraża żadnej zwrotnej informacji odnoszącej się do rezultatów podjętych przez nich aktywności. O ich obecności w bibliotekach świadczą jedynie dane systemowe monitorujące ruch w serwisach. Dla sobie znanych pobudek pozostają anonimowi i wirtualni, nie nawiązują kontaktu z bibliotekarzami czy też administratorem biblioteki cyfrowej, co skutkuje tym, że nie jest możliwe również nawiązanie z nimi jakiegokolwiek kontaktu. Taki całkowity brak reakcji ze strony tych Użytkowników pozwala na określenie tej grupy mianem **Użytkowników-Odbiorców**, którzy poprzez mało aktywny sposób funkcjonowania w życiu bibliotek cyfrowych, sytuują się jedynie jako bierni odbiorcy zaproponowanych poprzez te biblioteki usług. Grupa ta nie przejawia jakiegokolwiek zainteresowania kształtem i rozwojem biblioteki cyfrowej i nie uczestniczy w żaden inny sposób w życiu tej biblioteki. Wśród tej grupy są stali Użytkownicy bibliotek cyfrowych, ale z dużą dozą prawdopodobieństwa zakładać można, że trzon tej grupy stanowią Użytkownicy okazjonalni i okresowi, którzy korzystają z biblioteki cyfrowej tylko w określonym celu (np. studenci w celu napisania pracy) i po zrealizowaniu zadania nie będą już wcale, lub tylko sporadycznie, korzystać z biblioteki cyfrowej.

Charakteryzując tę grupę Użytkowników bibliotek cyfrowych podkreślić należy, że z punktu widzenia jej twórców, to właśnie ta grupa stanowi najbardziej perspektywiczne grono odbiorców usług dostarczanych przez biblioteki cyfrowe. Plany związane z rozwojem bibliotek cyfrowych powinny zostać skierowane na zaktywizowanie Użytkowników-Odbiorców i być ukierunkowane na zatrzymanie tego typu Użytkowników w serwisie, także po wykonaniu przez nich określonego zadania, dla którego skorzystali z usług biblioteki cyfrowej. Nawiązanie z nimi interakcji i rozpoznanie a następnie zaspokojenie ich potrzeb informacyjnych, przyczyni się do rozbudowywania grona bardziej aktywnej społeczności skupionej wokół bibliotek cyfrowych.

Użytkownik-Klient

Zdecydowanie odmienny model funkcjonowania w środowisku bibliotek cyfrowych prezentuje grupa ich aktywnych Użytkowników. Użytkownicy ci są świadomi możliwości jakie dla ich potrzeb są w stanie oferować biblioteki cyfrowe, a ich postawa przypomina odbiorców usług komercyjnych, przez co najbardziej trafnym jest określenie ich jako **Użytkowników-Klientów.**

Ta grupa użytkowników bibliotek cyfrowych nie tylko korzysta z zasobów biblioteki cyfrowej, ale również przejawia w różnym stopniu potrzebę wpływu na kształt tej usługi, stając się tym samym aktywnym podmiotem tego środowiska. Użytkownicy-Klienci to najczęściej stali Użytkownicy tego typu usług, reprezentujący różny stopień umiejętności obcowania z bibliotekami cyfrowymi. Wielu z nich korzysta z bardziej zaawansowanych usług bibliotek cyfrowych np. z możliwości założenia konta oraz korzystania z dostępnych funkcjonalności jakie posiadanie takiego konta daje, czy też subskrybowanie informacji poprzez kanał RSS. W różnorodny sposób Użytkownicy-Klienci samodzielnie nawiązują kontakt z osobami zaangażowanymi w proces tworzenia bibliotek cyfrowych – co jest cechą zdecydowanie odróżniającą ich od Użytkowników-Odbiorców. Kontakt ten bywa nawiązany różnymi kanałami i z różną częstotliwością. Najczęściej Użytkownicy-Klienci komunikują się za pomocą poczty elektronicznej, dedykowanego do takich celów wątku na Forum Biblioteka 2.0 [http://forum.biblioteka20.pl/] lub nawet wybierają bardziej bezpośrednią formę komunikacji, tj. pomoc telefoniczną a nawet kontakt osobisty. Taki stan rzeczy powoduje, że niejednokrotnie przestają oni być anonimowymi Użytkownikami usług dostarczanych przez biblioteki cyfrowe, obalając tym samym mit wirtualnego – z założenia – odbiorcy tych usług.

Tak wyrażana postawa Użytkowników-Klientów świadczy przede wszystkim o pozytywnym odbiorze bibliotek cyfrowych w społeczeństwie, oraz wysokim stopniu zaufania wobec takiej formy usług biblioteki tradycyjnej w przeładowanym informacją Internecie. Twórcom bibliotek cyfrowych daje ona wspaniałą okazję na nawiązanie kontaktu z odbiorcami tych usług, zaproszenia ich do współtworzenia środowiska bibliotek cyfrowych, co w konsekwencji prowadzi do owocnej współpracy, przynoszącej profity każdej ze stron[11].

Skupienie uwagi twórców bibliotek właśnie na tej aktywnej grupie Użytkowników bibliotek cyfrowych, właściwe rozpoznanie i zaspokojenie ich potrzeb związanych z tą usługą, przejawia się wyrażanym przez tych Użytkowników stopniem ich zadowolenia z oferty usług świadczonych przez biblioteki cyfrowe. W konsekwencji stanowi to podstawę dla odpowiednio ukierunkowanego rozwoju tego typu serwisów zapewniając dalszy wzrost zadowolenia Użytkowników z usługi jaką ogólnie jest biblioteka cyfrowa. Wychodząc z tak przyjętego założenia, ten typ Użytkownika-Klienta jest

[11] Przykładem może być np. zaproszenie Użytkowników bibliotek cyfrowych do udziału w badaniach, których głównym celem jest pozyskanie od nich bezpośredniej opinii odnośnie funkcjonowania i wykorzystania przez nich bibliotek cyfrowych. W 2009r. dobrowolnie z takiego zaproszenia skorzystało 946 Użytkowników polskich bibliotek cyfrowych, a w roku 2011 – 530 przedstawicieli różnych grup Użytkowników bibliotek cyfrowych.

najbardziej pożądanym w bibliotece cyfrowej, a stopień jego zaangażowania i aktywność, ma wpływ na rzeczywisty obraz bibliotek cyfrowych oraz ich pozycję w społeczeństwie.

Aktywność Użytkownika-Klienta przejawia się na wielu polach funkcjonowania bibliotek cyfrowych. Najczęściej, jego aktywna rola w budowaniu społeczności skupionej wokół bibliotek cyfrowych, wyraża się poprzez:

- Samodzielną reakcję i komentarz (zarówno negatywny jak i pozytywny) na zaproponowane przez twórców bibliotek cyfrowych zmiany w korzystaniu z bibliotek cyfrowych. Dotyczą one najczęściej zmian wprowadzonych w kolejnej wersji oprogramowania, sposobie wyświetlania i prezentowania dostępnych publikacji, zmianie kształtu dostępnych funkcjonalności itd.,
- Zgłaszanie usterek i błędów np. odnoszących się do publikacji niekompletnych, zawierających pominięte podczas skanowania strony czy skany o złej jakości ale także błędne czy niekompletne opisy publikacji,
- Informowanie o kłopotach technicznych występujących podczas korzystania z biblioteki cyfrowej np. związanych z wydajnością systemów obsługujących biblioteki cyfrowe,
- Skierowanie zapytania informacyjnego (kwerendy) wyrażającego prośbę o pomoc w dotarciu do poszukiwanej informacji (zarówno w bibliotece cyfrowej jak i w zbiorach bibliotek tradycyjnych),
- Skierowanie prośby o umożliwienie dostępu do wyszukanego zasobu (dotyczy to jedynie sytuacji, w których w ramach bibliotek cyfrowych funkcjonują repozytoria uczelniane, które ze względu na poszanowanie przepisów obowiązującego prawa autorskiego, dają prawo korzystania z wybranych zasobów, jedynie studentom tych uczelni),
- Współtworzenie zasobów bibliotek cyfrowych tj. proponowanie tytułów do zamieszczenia w bibliotece, wartościowych z punktu widzenia Użytkowników,
- Rozpowszechnianie informacji o zasobach bibliotek cyfrowych np. poprzez umieszczanie informacji o ich istnieniu i działalności na prywatnych stronach WWW, linkowanie bezpośrednio do zasobów bibliotek cyfrowych z różnorodnych portali tematycznych,
- Wykorzystywanie bibliotek cyfrowych jako platformy mającej zastosowanie w zdalnym nauczaniu (e-learning),
- Wykorzystywanie bibliotek cyfrowych jako formy publikowania własnego dorobku naukowego (e-publishing).

W niektórych przypadkach formy aktywności Użytkowników-Klientów bibliotek cyfrowych w zdecydowany sposób wybiegają poza konwencjonalne ramy funkcjonowania tego typu usług. Przykładem może być postawa Użytkowników będących właścicielami unikatowych oryginałów, którzy wyrażają chęć publikowania cyfrowych wersji tych materiałów w bibliotekach. Zaangażowanie Użytkowników wyraża się również poprzez publikowanie zdjęć, albumów, fotografii, pamiętników i innych pamiątek rodzinnych za pośrednictwem bibliotek cyfrowych. Przykładem aktywnych Użytkowników bibliotek cyfrowych są także ci, którzy powodowani chęcią skomple-

towania lub też przyspieszenia procesu udostępnienia w bibliotece cyfrowej danego zasobu, przekazują, do publikacji na stronach biblioteki cyfrowej samodzielnie zdigitalizowany materiał, będący w swych początkowych założeniach kolekcją prywatną tworzoną jedynie na potrzeby jej twórcy. Niejednokrotnie tak pozyskane materiały stanowią bardzo cenny materiał źródłowy, który pozwala na stworzenie kompletnej kolekcji cyfrowej, niedostępnej w oryginale w żadnej z bibliotek w kraju, nawet tych z egzemplarzem obowiązkowym. Ogólnie, taka aktywna postawa Użytkowników bibliotek cyfrowych pozwala nie tylko na zwiększenie dostępności do materiałów trudnodostępnych i rzadkich ale stwarza jedyną możliwość dostępu do materiałów prywatnych, rozproszonych po całym kraju, mających w oryginale wartość jedynie dla posiadacza tego typu zasobów.

Wizytówka Użytkowników polskich bibliotek cyfrowych

Wyniki badania przeprowadzonego na przełomie 2008/2009r. pozwoliły twórcom bibliotek cyfrowych przybliżyć nieco społeczność Użytkowników tych serwisów [Mazurek 2009].

Pokazały one, że w większości z bibliotek cyfrowych korzystają częściej mężczyźni niż kobiety (67% badanych stanowili mężczyźni a 30% kobiety[12]). Użytkownicy tego typu usług bibliotek tradycyjnych, to praktycznie przedstawiciele każdej grupy wiekowej – większość (60%) to osoby w wieku pomiędzy 20-25 lat, 31-40 lat oraz 51-60 lat. Najmniej, bo tylko 1 % całej grupy badanych, miało powyżej 71 lat a 3% badanych miało mniej niż 20 lat.

Są to osoby w większości wykształcone. Więcej niż połowa posiada wykształcenie wyższe (56%), (w tym wyższe magisterskie – 44%). Częściej niż co 10 Użytkownik (11%) posiada stopień naukowy. Mniej niż co trzeci czytelnik (28%) legitymuje się wykształceniem średnim. Jedynie nieliczni (2%) posiadają wykształcenie podstawowe.

Częściej niż co 3 osoba odwiedzająca biblioteki cyfrowe jest studentem lub pracuje naukowo. (odpowiednio: 20% i 17% badanych).

Biorąc pod uwagę miejsce zamieszkania Użytkowników bibliotek cyfrowych, zauważyć można tendencję, że im mniej liczne miejscowości tym mniejszy jest tam procent Użytkowników korzystających z bibliotek cyfrowych, a im większa liczebność miejscowości tym procent ten wzrasta. Prawie połowa (49%) czytelników bibliotek cyfrowych mieszka w dużych miastach liczących powyżej 100 000 mieszkańców. 17% mieszka w miejscowościach powyżej 25 000 mieszkańców (25 000-100 000). Dla mieszkańców wsi i miast do 5 000 mieszkańców procent ten wynosi odpowiednio 9% i 8%.

Użytkownicy korzystający z polskich bibliotek cyfrowych w większości mieszkają w Polsce (89% badanych), choć serwisy te są odwiedzane także przez Internautów

[12] 3% respondentów nie udzieliło odpowiedzi na to pytanie.

zamieszkujących poza granicami naszego kraju. W badaniu udział wzięli Użytkownicy, którzy jako kraj zamieszkania podali następujące kraje: Niemcy (14 badanych), Ukraina (9), Francja (5), USA (5), Białoruś (5, Australia (2), Szwecja (2), Kanada (2), Czechy (2), Belgia (1),Bułgaria (1), Izrael (1), Norwegia (1), Rosja (1), Wielka Brytania (1).

Grupy kluczowych Użytkowników polskich bibliotek cyfrowych

W oparciu o ww. badanie a także analizę e-maili od Użytkowników skierowanych na skrzynkę Wielkopolskiej Biblioteki Cyfrowej, udało się wyodrębnić 3 podstawowe, najbardziej liczne grupy Użytkowników, którzy dla różnych celów odwiedzają biblioteki cyfrowe. Grupy te to Hobbyści, Studenci oraz Pracownicy naukowi[13].

Hobbyści

Hobbyści stanowią zdecydowanie najliczniejszą grupę Użytkowników, która w największym stopniu eksploruje zasoby polskich bibliotek cyfrowych. Głównym celem, dla którego odwiedzają oni ten typ serwisów internetowych jest ich hobby, które dzięki tego typu usługom bibliotek tradycyjnych, są w stanie rozwijać. W tej grupie Użytkowników przeważają ci, którzy interesują się historią (w szczególności regionalną) – 42% badanych podało, że „często" i „dość często" w tym celu odwiedza strony bibliotek cyfrowych, oraz genealogią – 31% badanych „często" i „dość często" poszukuje na własne potrzeby informacji odnośnie swoich przodków, wykorzystując w tym celu materiały udostępniane w bibliotekach cyfrowych.

Studenci

Studenci stanowią drugą co do wielkości grupę Użytkowników polskich bibliotek cyfrowych, która jednakże liczebnie jest dużo mniejsza niż grupa Hobbystów. Co piąty Użytkownik biblioteki cyfrowej jest Studentem (20% ogólnej liczby badanych) różnego typu (licencjackich, magisterskich, zaocznych, doktoranckich), kierunków, uczelni i lat studiów. W tej grupie badanych 6% odwiedza stronę biblioteki codziennie, ponad 13% kilka razy w tygodniu a kilka razy w miesiącu aż 33% badanych Studentów. W bibliotekach cyfrowych poszukują oni przede wszystkich materiałów potrzebnych im do studiów oraz skryptów uczelnianych. Zdecydowanie najwięcej badanych Użytkowników w tej grupie to Studenci nauk humanistycznych, takich

[13] Uzyskanie pełniejszego obrazu tych grup Użytkowników polskich bibliotek cyfrowych zostało uzyskane dzięki wynikom kolejnego badania przeprowadzonego w 2011 r., dedykowanego już bezpośrednio do przedstawicieli tych trzech grup. Bardziej wyczerpująca charakterystyka każdej z tych grup zaprezentowana zostanie w ramach niezależnych publikacji autorki w ogólnie dostępnych czasopismach branżowych w 2011 r.

kierunków jak: historia i – w mniejszym stopniu – informacja naukowa i bibliotekoznawstwo oraz prawo.

Pracownicy naukowi

Pracownicy naukowi stanowią kolejną największą pod względem liczebności grupę (17% ogólnej liczby badanych) Użytkowników polskich bibliotek cyfrowych. Wykorzystują oni biblioteki przede wszystkim do prowadzenia badań związanych z własną pracą naukową. W mniejszym stopniu zasoby bibliotek cyfrowych stanowią także wsparcie dla prowadzonych przez nich zajęć dydaktycznych na uczelniach. W tej grupie badanych, znacząca liczba Użytkowników to osoby mieszkające poza Polską. Podobnie jak w grupie Studentów, ponad 12% wśród badanych Naukowców odwiedza bibliotekę cyfrową codziennie, 33% kilka razy w tygodniu, a ponad 39% kilka razy w miesiącu.

Badani naukowcy to w przeważającej części przedstawiciele nauk humanistycznych, głównie historycy (o różnych specjalizacjach), ale także filolodzy, literaturoznawcy, językoznawcy, kulturoznawcy, archeolodzy, socjolodzy i in. Wśród badanych znaleźli się także przedstawiciele nauk ekonomicznych, biologicznych, medycznych, matematycznych, chemicznych, fizycznych, prawnych, rolniczych oraz nauki o Ziemi.

Podsumowanie

Polskie biblioteki cyfrowe, stanowiące jedną z prężniejszych ofert usług oferowanych przez biblioteki tradycyjne, z pewnością zadomowiły się już na stałe w świadomości pewnych grup Internautów. Rosnące z roku na rok coraz większe grono ich Użytkowników pozwala spekulować, że pomimo różnych ograniczeń w jakich przyszło je budować, są one ważnym i zauważalnym elementem w budowaniu społeczeństwa opartego na wiedzy, a ich popularność będzie wciąż rosła. Pozytywnym aspektem tworzenia bibliotek cyfrowych jest aktywność i zaangażowanie pewnej grupy (Użytkowników-Klientów) tych serwisów, którzy dobrowolnie angażują się w tworzenie bibliotek cyfrowych. Pozwala to na tworzenie przydatnych dla ich celów serwisów internetowych dystrybuujących wartościowe dla nich informacje, w odróżnieniu od ogromu mało wiarygodnych treści powszechnie dostępnych w sieci. Współcześnie funkcjonujące w Polsce biblioteki cyfrowe – z powodu różnego typu trudności – nie są w stanie zadowolić potrzeb informacyjnych wszystkich jej Użytkowników, stąd zauważalna jest również grupa tych Użytkowników, dla których oferta biblioteki cyfrowej jest atrakcyjna jedynie w ograniczonym zakresie. Zdecydowanie największą grupą Użytkowników, która obecnie w największym stopniu korzysta z bibliotek cyfrowych są Hobbyści. Wykorzystują oni biblioteki cyfrowe do rozwijania swojego hobby, którym najczęściej jest genealogia oraz ogólnie ujmując historia – z uwzględnieniem historii regionalnej. Kolejne dwie grupy kluczowych Użytkowników polskich bibliotek

cyfrowych stanowią Studenci (przede wszystkim nauk humanistycznych) oraz Pracownicy naukowi, dla których biblioteki cyfrowe stanowią ważny element w ich warsztatach pracy.

Biblioteki cyfrowe, stale powiększające swoje zasoby, a także rozwijające różnego rodzaju inne funkcjonalności, mają szansę na oferowanie nowych możliwości oraz intensyfikację już oferowanych usług. W przyszłości pozwoli to na rozbudowanie społeczności skupionych wokół bibliotek cyfrowych, które staną się również atrakcyjne dla innych, nowych grup odbiorców. W takim ujęciu najbardziej perspektywicznymi grupami odbiorców usług bibliotek cyfrowych, są właśnie ci Użytkownicy, których nie trzeba przekonywać do samej idei bibliotek cyfrowych, a których obecne potrzeby informacyjne, serwisy te zaspakajają jedynie w umiarkowanym lub minimalnym zakresie.

Bibliografia

Mazurek, Jolanta (2009). *Drogi czytelniku, jak korzystasz z biblioteki cyfrowej? Próba odpowiedzi* [dok. lektr.] „EBIB – Elektroniczny Biuletyn Bibliotekarzy" nr 105 http://www.ebib.info/2009/105/a.php?mazurek [odczyt: 15.06.2011].

Modele mentalne użytkowników bibliotek cyfrowych

Monika Krakowska

Wstęp

Do zagadnień pragmatyki tworzenia oraz implementowania rozbudowanych systemów informacyjno-wyszukiwawczych, w tym także bibliotek cyfrowych, coraz częściej włączane są kwestie dotyczące projektowania zorientowanego na użytkownika (*user-oriented design*) [Xie 2008, s. 1346-1351; Comeaux 2008; Frias-Martinez i in. 2007, s. 305-309]. Problematyka komponowania środowiska cyfrowego odzwierciedlającego paradygmat kognitywny wraz z indywidualnymi cechami użytkownika, jego wiedzę i doświadczenia, motywacje, stany emocjonalne, zachowania informacyjne, staje się niezwykle istotna w organizacji repozytoriów cyfrowych. Interpretowane jako systemy, biblioteki te, stanowią także rodzaj specyficznego narzędzia, które pozwala na umysłową aktywność użytkownika, bez barier w dostępie do informacji i logicznych, koncepcyjnych, fizycznych lub czasowych granic. Traktowane jako unikalne centra działalności intelektualnej, umożliwiają one także realizowanie różnorodnych procesów, zachodzących w systemie, którego komponentami są kolekcje, infrastruktura sprzętowo-programowa, społeczność użytkowników, zarówno twórców i odbiorców oraz projekty [Fuhr i in., dok. elektr.]. Biblioteki cyfrowe można także rozpatrywać jako przestrzeń, infrastrukturę skupioną wokół użytkownika oraz miejsce interakcji, jakie podejmuje odbiorca w tym środowisku informacyjnym [Chowdhury 2010, s. 208]. W bibliotece cyfrowej pojmowanej jako obiekt, którego komponentami są system komputerowy, kolekcja oraz wielorako rozumiani użytkownicy (zarówno administratorzy, bibliotekarze twórcy, naukowcy, informatycy, odbiorcy), istotnym elementem analizy i ewaluacji jakości oraz czynnikiem całościowego rozpatrywania będzie percypowany użytkownik. Kompleksowość pojmowania tego konceptu, powoduje rozważania z różnych perspektyw, uzależnione od informacyjnej, socjologicznej i psychologicznej interpretacji

biblioteki cyfrowej. Zarówno wykładnia, tudzież ewaluacja jakościowa służy ustaleniu poziomu jej funkcjonowania, czyli efektywności i skuteczności, jak i samej wartości. Skoncentrowanie się w badaniach na użytkowniku oraz konsytuacyjności, w jakiej się znajduje, ma także wpływ na ocenę tych systemów informacyjnych w czterech wymiarach: społecznościowym, serwisu, technologii oraz zawartości. Różnorodne ujęcia biblioteki cyfrowej, umożliwiają jej ewaluację w perspektywie oddziaływania (Impact-ED, *Impact Evaluation for Digital Libraries*), uwzględniającym modele mentalne użytkowników, a także czynniki kognitywne wpływające na interakcje użytkownika z systemem informacyjno-wyszukiwawczym. Te z kolei wpływają na podejmowane przez nich procesy oraz interakcje z systemem [Madle, Kostkova, Roudsari 2008, s. 100-101]. W prowadzonych analizach o orientacji psychologii poznawczej, istotny jest użytkownik, indywidualne struktury wiedzy, sposoby przetwarzania informacji, interakcje, elementy subiektywne oraz afektywne.

W filozoficznym ujęciu Georga Lakoffa i Philipa Johnsona-Lairda [Piłat, dok. elektr.] umysł odzwierciedla podłoże, kontekst rzeczywistości, zaś pojęcia urzeczywistniane są w ludzkim mózgu, który cechuje ewolucyjność i emocjonalne zaangażowanie[1]. Podłoże wyznacza również ogólne ramy, podejścia badawcze oraz probierze ewaluacji bibliotek cyfrowych. Analiza implementowania, funkcjonowania, interakcji bibliotek cyfrowych prowadzona w Sieci Doskonałości dla Bibliotek Cyfrowych DELOS (*Network of Excellence on Digital Libraries*) oraz badaczy takich jak Tefko Saracevic, Hong Iris Xie, Ying Zhang [Saracevic, dok. elektr.; Zhang, dok. elektr.; Xie 2006] spowodowała wypracowanie pewnych wskaźników charakteryzujących paradygmaty ich eksploracji. Są one zorientowane przede wszystkim na użytkownika (zawierają komponenty psychologiczne, antropologiczne, społecznościowe), w tym również zachowania informacyjne jakie podejmuje on w interakcji z systemem oraz wpływające na użyteczność i funkcjonalność tego swoistego organizmu [Saracevic, dok. elektr.]. Doskonalenie bibliotek cyfrowych opiera się choćby na ocenie jakości i funkcjonowania systemu, wpływie komponentów oraz zachodzących między nimi relacjach na zmiany w zachowaniach informacyjnych, reakcje poznawcze. Staje się zatem konieczna podczas wdrażania oraz inicjowania bibliotek cyfrowych wiedza o tym w jaki sposób użytkownik tworzy w umyśle reprezentację, umysłowy model tego obiektu[2]. Elementy

[1] Potwierdzają to również badania w psychologii ewolucyjnej, gdzie na uwagę zasługuje koncepcja MacLeana, a także teoria Johna Toobyiego i Ledy Cosmides. Istotne jest tu uwypuklenie adaptacyjnej roli mózgu oraz czynnika afektywnego (emocji, motywacji). Za: Lewis, M., Haviland-Jones, J. M. (2005). Psychologia emocji. GWP, 904 s.

[2] Należy zaznaczyć, że w informatologii oraz badaniach nad interakcjami człowiek-komputer (prowadzonymi także w informatyce, psychologii, kognitywistyce) reprezentacje powstające w ludzkim umyśle określa się jako:

a) **mapy koncepcyjne** (Kenneth Craik, 1943; Dania Bilal, Peiling Wang (2005). Children's Conceptual Structures of Science Categories and the Design of Web Directories. *Journal of the American Society for Information Science and Technology*, 56(12): 1303-1313; Peiling Wang, Stephen Bales, Jason Rieger, Yan Zhang (2000) Survey of Learners' Knowledge Structures: Rationales, Methods and Instruments. dok. elektr.: http://ils.unc.edu/~yanz/cogmapping.pdf [odczyt: 12.12.2011]);

b) **mapy umysłu** (*mind map*) (Tony Buzan, Barry Buzan (2005). The mind map book. BBC Active, pp. 75-82);

konstrukcyjne i ewaluacyjne odnoszą się zatem do modeli mentalnych wszystkich użytkowników tworzących system biblioteki cyfrowej, znajdujący się w określonej konsytuacji społecznej, antropologicznej, technologicznej, afektywnej[3]. Reprezentowanie konkretnej sytuacji, obiektu lub systemu umożliwione jest również poprzez językowe i semantyczne frazeologizmy oraz utożsamianie w kontekście sytuacyjnym, co pozwala na obserwowanie zachowań, odwzorowywania i rozumienia motywacji wpływających na procesy informacyjne [Westbrook 2006, s. 564-566]. W przypadku implementowania modeli mentalnych do konstruowania systemów ważne jest skoncentrowanie się na efektywnych, emocjonalno-poznawczych interakcjach, opierających się na doświadczeniach, zapamiętanych stanach afektywnych, a także na poznawczych, behawioralnych oraz społecznych umiejętnościach użytkownika [Izard 2009, s. 2-4].

Interdyscyplinarność badań nad modelami mentalnymi

Teoria modeli mentalnych mieści się w nurcie badań psychologicznych, ale włączana jest także do paradygmatów innych dyscyplin jak informatologia, informatyka, socjologia, antropologia i wykorzystywana do badań nad procesami poznawczymi jednostki. Wiedza dotycząca tworzonych indywidualnie i wykorzystywanych w procesach informacyjnych, specyficznych reprezentacjach umysłowych może być podstawą tworzenia, zarządzania i stosowania różnorodnych systemów, serwisów, stron WWW, obiektów. W informacji naukowej kontynuuje się komplementarne analizy nad tymi strukturami wewnętrznymi, przede wszystkim w charakterystyce systemów informacyjno-wyszukiwawczych, interakcji człowiek-komputer. Szczególną uwagę poświęcono atrybutom, cechom i właściwościom modeli mentalnych (np. Papastergiou 2005; Zhang 2009), komponentom modeli mentalnych użytkowników systemów informacyjno-wyszukiwawczych (np. Makri i in. 2007; Westbrook 2006), a także wpływowi modeli mentalnych na zachowania informacyjne i procesy wyszukiwania informacji (np. Yan Zhang 2008; Xiangmin Zhang, Mark Chignell 2001; Jack Hsu i in. 2011).

W informatologicznych eksploracjach rozważano nad modelami mentalnymi w kontekście: 1) katalogów bibliotecznych online (Borgman 1989, Kerr 1990, Slone 2002), 2) baz danych (Katzeff 1990; Tallman, Henderson 2011), 3) serwisów WWW oraz wyszukiwarek internetowych (He i in. 2008, Efthimiadis, Hendry 2005; Zhang 2008), 4) eksperymentalnych systemów informacyjno-wyszukiwawczych i modelu za-

c) **mapy kognitywne** (*cognitive maps*) (Patrick Péruch, Loïc Belingard, Catherine Thinus-Blanc (2000). Transfer of Spatial Knowledge from Virtual to Real Environments. *Spatial Cognition II Lecture Notes in Computer Science*, 2000, Volume 1849/2000, pp. 253-264);

d) **modele mentalne** (*mental models*).

[3] Modele te, o charakterze prototypów, schematów psychicznych określa się poprzez pryzmat indywidualnych i uwarunkowanych wewnętrznie zachowań oraz wiedzy, uwzględniających stany emocjonalne użytkowników. Przez to mają istotny wpływ na rozumienie, przewidywanie procesów, zjawisk i podejmowane decyzje i reagowanie na bodźce (Zhang 2007, p. 1330-1331).

chowań informacyjnych *Anomalous State of Knowledge* (nietypowych stanów wiedzy) (Belkin 2005, Darabi i in. 2009, Zhang 2012), 5) wizualizacji interfesju (Chalmers 2003, Roth i in. 2010), 6) map kognitywnych tworzonych dla środowiska wirtualnego (Billinghurst, dok. elektr., Morganti, Carassa, Geminiani 2007), 7) Internetu, stron WWW (Zhang 2008, Holman 2011), 8) ewaluacji portali i serwisów internetowych (Roth et. al. 2010) oraz 9) bibliotek cyfrowych (Veldof, Beavers 2001, Makri i in. 2007).

Najczęściej wykorzystywanymi metodami pomiaru, odwzorowania lub przedstawienia modeli mentalnych użytkowników są: 1) wywiady lub protokoły głośnego myślenia (*think-aloud protocols*), pozwalające na semantyczne, werbalne wytłumaczenie umysłowej reprezentacji przez użytkownika. Analiza wypowiedzi słownych nie jest jednak pozbawiona nieścisłości. Wpływ na jakościowe wyniki ma indywidualna zdolność użytkowników do artykułowania swoistych rozważań. Istnieje także dyspozycja do uzasadnienia lub racjonalizowania indywidualnych działań użytkowników w protokołach werbalnych; 2) wizualizacja poprzez metodę rysowania (*drawings*). W przypadku obrazowych reprezentacji modeli mentalnych trudności następują podczas interpretowania rysunków, mogą także przedstawiać wyłącznie wybrane aspekty funkcjonowania określonego systemu, obiektu i nieumiejętnie stosować się do wiedzy użytkownika o strukturze np. biblioteki cyfrowej; 3) narzucanie różnorodnych warunków, kontekstu związanego z określonym obiektem, systemem, zjawiskiem podczas badania i założenie, że wynikiem będzie powstanie wielorakich, nierelewantnych modeli mentalnych systemu. Możliwe jest także, że użytkownicy, którzy posługują się instrukcjami schematycznymi, dotyczącymi architektury np. biblioteki cyfrowej (*model-based instructions*) wytworzą bardziej kompleksowe, rozbudowane reprezentacje umysłowe, niż odbiorcy stosujący się do wytycznych proceduralnych (*procedure-based instructions).* Możliwe jest również wytworzenie modeli mentalnych bez wcześniejszego instruowania użytkowników w zakresie posługiwania się systemem [Zhang 2012, s. 108-109].

Charakterystyka modeli mentalnych

Ukierunkowanie i właściwości procesów tworzenia wiedzy i przekształcania już nabytej, a także wynikających z tych sekwencji relewantnych zachowań informacyjnych, opierają się na reprezentacjach wewnętrznych wspomagających rozumienie systemu, obiektu, zjawiska, bodźca, działania, etc. We wspomnianym ujęciu koncepcji psychologii kognitywnej każdy podmiot wizualizuje eksploracyjnie wzorce porównawcze otaczającego go świata. Są to wyobrażenia, często niekompletne, związane z jednostkowym doświadczeniem, wiedzą, trudne do zaobserwowania, imitacje elementów i struktury realnych obiektów [Zhang 2007, s. 1330]. Modele mentalne cechuje abstrakcyjność, co stanowi o specyfice tej *nietrwałej reprezentacji słowno-obrazowej, o dość dużym poziomie złożoności, powstająca w toku procesu myślenia lub rozumowania, jako sposób na przedstawienie sobie w umyśle istoty problemu i składających się nań zależności* [Nęcka, Orzechowski, Szymura 2008, s. 640].

Po raz pierwszy modele mentalne zostały opisane przez Kennetha Craika w książce *The Nature of Explanation* opublikowanej w 1943 r. [Craik 1943], w której autor określił je jako imitacje, struktury o mniejszej skali (*small-scale models*) świata fizycznego, konstruowane w psychice człowieka, stanowiące podłoże ludzkich działań oraz antycypacji zdarzeń, uwypuklania objaśnienia zewnętrznego uniwersum. Teoria Kennetha Craika została eksplorowana czterdzieści lat później, m.in. przez Philipa Johnsona-Lairda oraz Ruth Byrne, którzy scharakteryzowali modele mentalne jako psychiczne reprezentacje rzeczywistości, hipotetycznej lub wyimaginowanej sytuacji, zjawiska. Spostrzeżenia Kennetha Craika zostały rozwinięte o nowe poglądy, sugerując, iż reprezentacje te są kombinacją analogowych i logicznych schematów, relacyjnych w stosunku do obiektów, jednak różnicujących je w strukturze umysłu [Makri 2004, s. 13]. Rozbieżność wyobrażeń rzeczywistości w umyśle użytkownika dotyczy dysproporcji w niezmienności, trwałości modeli mentalnych, optymalizujących kreatywne podejmowanie procesów poznawczych, informacyjnych. Ułatwiają także rozwiązywanie komplikacji poprzez coraz bardziej odpowiednie, zastępujące się wzajemnie reprezentacje umysłowe. Antynomiczni użytkownicy posiadają mniej lub bardziej skomplikowane modele mentalne, uzależnione od poziomu wiedzy i indywidualnych potrzeb informacyjnych. Charakterystyczne jest również to, że reprezentacje mają formę nietrwałą, subiektywną, tworzone zaś są w konkretnej sytuacji. Tworzą strukturę wiedzy, która jest podstawą rozumienia świata, nadawania mu sensu oraz wzajemnego oddziaływania z człowieka z uniwersum [Johnson-Laird 2001, s. 434-435]. Philip Johnson-Laird zauważył również semantyczne oraz strukturalne podobieństwo modeli mentalnych do rzeczywistości. Model mentalny kwantyfikuje kognitywny mechanizm do reprezentowania lub wnioskowania odnoszącego się do systemu, zjawiska, obiektu, bądź problemu, który użytkownik tworzy podczas interakcji oraz uczenia się tego systemu, zjawiska, obiektu [Borgman 1986, s. 48]. Według Jeroena Van Marriënboera, Norberta Seela i Paula Kirshnera model taki jest niezwykle złożoną konstrukcją faktów, konceptów, planów, zasad, relewantnych wzajemnie (gdzie relacje nie są przypadkowe). Postrzegany jest on jako rozbudowana reprezentacja umysłowa organizacji zakresu zadania (*task domain*) [Marriënboer, Seel, Kirshner 2002]. Procesy rozumowania, takie jak uczenie się, pojmowania, rozwiązywania problemów są uzależnione od tworzonych mentalnie mechanizmów modelu, rzeczywistości samej w sobie (*reality of its own*), ujednoliconej strukturalnie i autonomicznej, choć nakładającej na obiekt, zjawisko, system pewne ograniczenia[4]. Reprezentacja jest symbolem, ukazaniem zjawiska, zdarzenia, odzwierciedleniem intensywności zmysłów rozpoznających i percypujących bodziec. Forma może odnosić się do języka (jako element, słowo), idei, zbioru sądów, twierdzeń (*prepositions*), lub do obrazu, wizualizacji schematu i analogowego wyobrażenia [Nęcka, Orzechowski, Szymura 2008, s. 60-63].

[4] Ograniczenia te wynikają z samej psychiki jednostki, kreowanego w umyśle modelu rzeczywistości, nigdy zaś z oryginalnej, faktycznej realności. Modele budowane są z istniejącej wiedzy, doświadczenia i nakładane na obraz rzeczywisty.

Kathleen Carley i Michael Palmquist zauważają, że modele mentalne zbudowane są ze zorganizowanych struktur wiedzy – konceptów, zasad, połączeń. Stanowią one reprezentacje wewnętrzne człowieka, mogą być wyrażone poprzez system koncepcji, zaś nowe znaczenie koncepcji nadawane przez indywidualnego użytkownika jest osadzone na relacjach z innymi ideami, poglądami, intencjami, które już znajdują się w modelu mentalnym [Carley, Palmquist 1992, s. 603].

Obrazowi mentalnemu nadawany jest charakter określonych i prawdziwych cech materialnego obiektu, zjawiska, zdarzenia. Proces ten dotyczy wizualizowania w umyśle mentalnego modelu abstrakcyjnych pojęć, choć te, pomimo ścisłego konceptu mogą nie mieć sensorycznego odpowiednika w realiach [Paivo, dok. elektr.]. W odniesieniu do systemów informacyjno-wyszukiwawczych i interakcji człowiek-komputer, wiedza użytkownika zawarta w umysłowej reprezentacji dotyczy komponentów, organizacji i funkcjonowania tego układu. Im bardziej kompletny jest model, tym mniejsza liczba popełnianych błędów w procesie informacyjnym i określone, udoskonalone wzorce zachowań użytkownika [Dimitroff 1992, s. 142; Young 1983, s. 40, 45, 51].

John H. Holland, Keith J. Holyoak, Richard E. Nisbett i Paul R. Thagard charakteryzują model mentalny jako zbiór synchronicznych oraz diachronicznych, następujących po sobie reguł, zasad zorganizowanych w niejednoznaczne hierarchie oraz pogrupowanych w kategorie [Holland i in. 1989, s. 22]. Kategorie te obejmują trzy rodzaje wiedzy użytkownika: deklaratywną, strukturalną oraz proceduralną [Kahler, dok.elektr.]. Wiedza ta jest istotna dla rozumienia samego modelu mentalnego, jego zastosowania i udoskonalania. Deklaratywna orientacja oznacza wiedzę o czymś (*knowing what*), bez konieczności logicznego jej zaimplementowania do zastosowania obiektu, systemu, zjawiska oraz podjęcia interakcji. Strukturalna wiedza reprezentuje powiązania, związki pomiędzy poznaniem deklaratywnym, a przez to pozwala na tworzenie schematów oraz modeli mentalnych[5] konkretnego przedmiotu, obiektu, zjawiska. Wiedza proceduralna oznacza kompetencje do podjęcia sugestywnego działania (*knowing how to do*), włączając także umiejętności zdobyte poprzez doświadczenie. Zastosowanie trzech struktur wiedzy pozwala na formułowanie logicznego, przyczynowego oraz operacyjnego (proceduralnego) modelu, umożliwiającego umysłowe działania oraz interakcje z systemem, obiektem, etc.[6].

Reprezentacje umysłowe, pomimo tego, że są nieuporządkowane, niedokładne, niekompletne, stanowią podstawę rozwiązywania problemów podczas interakcji oraz akceptacji nowych sytuacji, bodźców, właśnie poprzez możliwość przewidywania reakcji na podstawie wcześniejszego modelu i doświadczenia [Norman 2002, s. 36-37]. Donald Norman stwierdził ogólnie, że: 1) modele mentalne są niekompletne, 2) umiejętność zastosowania modelu umysłowego przez użytkownika może być częścio-

[5] Wiedza strukturalna stanowi element kluczowy do kreowania modelu mentalnego rzeczywistości (systemu, obiektu, zjawiska, etc.), a przez to rozumowania, poznawania i zastosowania systemu, zasobu poprzez dostarczenie podstawowych zasad, reguł i powiązań (semantycznych, logicznych). [David Jonassen, Katherine Beissner, Michael Yacci, 1993, p. viii, 9.

[6] Model mentalny oparty jest na wiedzy np. o komponentach systemu, co pozwala na uświadomienie sobie w jaki sposób działa ten system, a także przewidywanie czynności, reakcji i rezultatów. Norman, 2002.

wo ograniczona, 3) modele mentalne są niestabilne, zaś użytkownicy mogą zapomnieć o szczególnych cechach charakterystycznych dla konkretnego systemu, obiektu, który wykorzystują, szczególnie jeśli nastąpiła dość duża przerwa w jego zastosowaniu, 4) modele mentalne nie posiadają sztywnych odgraniczeń, ram (wiedzę o działaniu systemu lub strukturze obiektu można przenieść na zupełnie nowe formacje), 5) modele mentalne są nienaukowe (użytkownika cechują zachowania informacyjne ignorujące wiedzę o bezcelowości podejmowanych działań, niepotrzebnych interakcjach, ze względu na to, że nie wymagają zbyt wiele wysiłku i utrzymują sprawdzony już model), 6) modele mentalne są oszczędne – zwykle użytkownik podejmuje się dodatkowych działań niż nadobowiązkowego projektowania, które pozwoliłoby na ograniczenie operacji i złożoności kognitywnych procesów [Norman 2002, s. 8].

Typologia modeli mentalnych

Zrozumienie funkcjonowania i działalności tradycyjnej biblioteki odwzorowywane jest w modelu mentalnym użytkownika, a następnie przekładane często na bibliotekę cyfrową, jej interpretację, zachowania podejmowane wewnątrz tego systemu informacyjno-wyszukiwawczego oraz procesy informacyjne [Makri i in. 2007, s. 433, s. 443]. Analizy tego modelu zasadzone są na przede wszystkim interpretowaniu, zdefiniowaniu biblioteki cyfrowej, określeniu czego metaforą staje się w momencie zastosowania przez odbiorcę oraz w jaki sposób jest rozumiana: system informacyjno-wyszukiwawczy, obiekt, serwis, wyszukiwarka internetowa, analogia tradycyjnej biblioteki, etc. [Makri, dok. elektr.] Reprezentacje umysłu wykorzystywane są tym samym do tworzenia struktur systemu informacyjno-wyszukiwawczego, serwisów, obiektów, z którymi człowiek podejmuje jakiekolwiek interakcje. Kluczowym substratem implementowania tychże systemów lub serwisów, stają się zatem wewnętrzne wizualizacje tworzone przez użytkowników, którzy z nich korzystają, inicjując procesy i zachowania informacyjne. Istotna jest również tendencja do mapowania nowych doświadczeń lub wiedzy i nakładanie na istniejące strukturalne relacje istniejące w modelu. Mapowanie podobnych elementów w reprezentacji, rozumowanie poprzez metaforę wydaje się być racjonalne dla zachowania aktualnego stanu wiedzy oraz wzmocnienia modelu mentalnego i zrozumienia nowego konceptu. Konieczne jest także uwzględnienie stanów afektywnych pełniących rolę aktywatorów, umysłowych stanów gotowości do podejmowania działań behawioralnych, optymalizujących indywidualne nastawienie do zastanej rzeczywistości oraz sposób reakcji [Baudry, Pinsonneault 2010, s. 690]. Konsolidację emocji oraz wpływających na modele mentalne stanów afektywnych, a także uwzględniające te probierze zasady tworzenia systemów informacyjno-wyszukiwawczych, w tym bibliotek cyfrowych, określa się jako „emocjonalne projektowanie" (*emotional design*) lub „afektywne przetwarzanie danych" (*affective computing*). Uwzględnienie czynników kognitywnych w konstruowaniu systemów komputerowych stanowi próbę ich formowania, zbliżonego do naturalnych reprezentacji powstających w umyśle użytkownika. Zaimplementowanie indywidualnych procesów kognitywnych, modeli mentalnych i emocji pozwala na

skonstruowanie efektywnych systemów, serwisów, obiektów, które są przyjazne, zrozumiałe, łatwe do zastosowania, rozpoznające i odpowiadające na powstające reakcje użytkownika (afektywne, pozytywne lub negatywne) w trakcie procesu informacyjnego [Lim, Blevis, Stolterman 2007, s. 106-115].

Model mentalny użytkownika, poprzez zastosowanie subiektywnych kryteriów, wypracowanych na podstawie wiedzy i doświadczenia odbiorcy, kreuje adekwatne zachowania informacyjne oraz decyduje o najczęściej pojawiających się interakcjach człowieka z komputerem. Przykładem jest reprezentacja mentalna interfejsu użytkownika, który skonstruowany jest z określonych wyobrażeń o relacjach jakie powstają podczas działań użytkownika a odpowiedzią systemu na te poczynania. Model mentalny jest zatem manipulatorem procesu rozumowania, nadawania sensu obiektowi, zbiorem możliwych reprezentacji informacji [Johson-Laird 2001, s. 434-436]. Użytkownicy posługują się nieidentycznymi, antytecznymi reprezentacjami umysłowymi. Można je podzielić na modele strukturalne[7] oznaczające reprezentowanie wewnętrznych mechanizmów narzędzia (np. biblioteki cyfrowej, Internetu, komputera, etc.), systemu oraz modele funkcjonalne, które umożliwiają odbiorcy zrozumieć w jaki sposób zastosować narzędzie [Zhang 2009, s. 14]. William M. Newman, Michael G. Lamming zasugerowali cztery rodzaje modeli mentalnych, opartych na stanach kognitywnych i poziomie wiedzy odbiorców. Zaliczyli do nich: 1) stany przejściowe (*state transition*), w których wizualizacja w umyśle użytkownika odwołuje się do zmian ogólnej wydajności i działania systemu, 2) odwzorowywanie (*mapping*), polegające na realizacji zamierzeń i przekształcaniu ich w sekwencję czynności, działań, reakcji z systemem, 3) struktury analogiczne (*analogical*), które opierają się na wcześniejszych doświadczeniach użytkownika zakodowanych podczas interakcji z systemem oraz 4) modele zbieżne (*similar*), dzięki którym odbiorca zapamiętuje komponenty systemu, a następnie kontroluje późniejsze oddziaływanie człowiek-komputer [Zhang 2009, s. 14].

Modele mentalne użytkowników systemów informacyjno-wyszukiwawczych

Modele mentalne jako reprezentacje wiedzy, są kluczowym elementem struktur kognitywnych, które charakteryzują procesy informacyjne oraz cykl uczenia się. Użytkownik konstruuje niewerbalne, obrazowe konfiguracje[8] w celu wyjaśnienia bądź imitacji specyficznych fenomenów obiektów, zdarzeń w przypadku braku odpowiednich odzwierciedleń mentalnych. Wyodrębniono także modele mentalne nowicjuszy i ekspertów [Zhang 2008, s. 1332], które różnią się złożonością, abstrakcyjnością. Początkujący użytkownicy charakteryzują się ograniczonym zasobem wiedzy i mniej-

[7] Wiedza o tym jak funkcjonuje narzędzie, jak działa (*how-it-works knowledge*) [Makri, dok. elektr.].

[8] Szczegółowo opisujące rzeczywistość poprzez nadawanie tym obrazom odpowiednich znaczeń słownych [Makri, dok. elektr.].

szą liczbą strategii pozwalających na rozwiązywanie problemów podczas interakcji z systemem, obiektem, zjawiskiem. Biegli użytkownicy, profesjonalni, wyróżniają się pewnością siebie, zorganizowana wiedzą i odpowiednim stosowaniem procesów informacyjnych (m.in. poprzez zastosowanie podczas strategii wyszukiwania metody „prób i błędów"). Wpływają one na podejmowane zachowania informacyjne, wzbogacone, rozbudowane w zależności od kompleksowości oraz interakcji z komputerem. Donald Norman zdefiniował w ujęciu kognitywistycznym modele mentalne w kontekście innych konfiguracji, schematów, które pojawiają się i są wykorzystywane w procesach oddziaływania człowiek-komputer (*human-computer interaction)*. Są to:

1) konceptualny model systemu (*system conceptual model*), który dotyczy kompleksowej wizualizacji, współmiernego i adekwatnego obrazowania systemu informacyjno-wyszukiwawczego, ale także może odnosić się do obiektu lub serwisu. W przypadku twórców, struktura konceptualna może zawierać: a) metafory i analogie (sposób organizacji informacji, np. w katalogu), b) koncepty (wraz z atrybutami – mogą być to metadane; procesami i serwisami bibliotecznymi – np. moduł wypożyczania, rezerwacji; zasoby cyfrowe; c) relacje – przykładowo numery, woluminy czasopism fachowych online, typ zasobów – książki, fotografie, etc.; d) odwzorowania (przykładowo każdy element systemu – biblioteki cyfrowej, odpowiada rzeczywistym obiektom, komponentom biblioteki tradycyjnej) [Makri, dok. elektr.];
2) wizualizacja systemu (*system image*), która obejmuje wygląd oraz wrażenie związane z zastosowaniem, użyciem systemu; obrazowanie mentalne systemu dotyczy odzwierciedlenia interfejsu, stron systemu, dokumentacji oraz informacji pojawiających się podczas pracy z komputerem. Empiria użytkownika związana z systemem poprzez jego mentalne reprezentowanie wyrażane jest także jako:
3) interfejs użytkownika mieszczący się w modelu systemu twórcy (*user interface designer's model of the system*). Model interfejsu tworzony przez twórców odnosi się równocześnie do konceptualnego modelu użytkowników końcowych;
4) model systemu użytkownika (*user's model*) tworzony jest w celu reprezentowania w jaki sposób odbiorcy percypują oraz interpretują system;
5) konceptualny obraz naukowców modelu mentalnego użytkowników (*Scientists' conceptualizations of the users' mental models*) dotyczy postrzegania i eksplikowania reprezentacji umysłowych odbiorców systemu tworzonych przez teoretyków;
6) systemowy model użytkownika (*system's model of the user*) tworzony jest wewnątrz architektury obiektu, np. biblioteki cyfrowej i działa poprzez różnorodne elementy, takie jak profile odbiorców, ustawienia, logowania, a także błędy popełniane w trakcie interakcji [Zhang 2009, s. 16-17].

Modele mentalne stanowią zatem swoiste mechanizmy pozwalające użytkownikom zrozumieć, akceptować oraz podejmować procesy informacyjne i wzajemne interakcje z systemem. Wizualizacje umysłowe kreowane przez interesariuszy bibliotek cyfrowych są przede wszystkim skomplikowane, uzależnione od stopnia zaawansowania zrozumienia i zastosowania narzędzi i form technologiczno-informacyjnych

odbiorcy, jak również poziomu wiedzy koniecznej do podejmowania procesów informacyjnych [Carroll, Olson 1988, s. 5].

Tego typu konstrukcje odwzorowane zostały między innymi w referencyjnym modelu DELOS, który jako abstrakcyjny schemat komponentów koniecznych do odwzorowania struktury biblioteki cyfrowej stosuje się do modeli mentalnych interesariuszy (użytkowników: architektów systemu, twórców, bibliotekarzy, naukowców, odbiorców końcowych), systemu oraz koncepcyjnego modelu twórców (projektodawców, administratorów, zarządzających). Podstawowy koncept opiera się na modułach, ogniwach współtworzących przestrzeń biblioteki cyfrowej wraz z architekturą, zawartością, użytkownikami, funkcjonalnością, jakością biblioteki cyfrowej, określoną polityką. Każdy z tych relacyjnych komponentów wzajemnie oddziałuje i funkcjonuje, opierając się na mapach koncepcyjnych, czyli mentalnych reprezentacjach, kreowanych przez poszczególne osoby tworzące specyfikę tego systemu informacyjno-wyszukiwawczego, określając odpowiedni poziom abstrakcji dla każdego modułu [Casarosa, dok. elektr.]. Model referencyjny biblioteki cyfrowej, jej mentalna reprezentacja tworzona przez wszystkich użytkowników, stanowi swoiste ramy pojęciowe, których zadaniem jest zrozumienie oraz przełożenie na prototypowy schemat różnorodnych komponentów tego systemu informacyjno-wyszukiwawczego. Celem jest eksplikacja zachowań informacyjnych użytkowników oraz ich relacji ze skonkretyzowaną strukturą biblioteki cyfrowej. Model wzorcowy zbudowany jest między innymi ze zbioru aksjomatów, wzajemnych relacji, ujednoliconych pojęć w danym obszarze zagadnieniowym. Stanowi prototyp niezależny od określonych standardów, technologii, a także pojedynczych wdrożeń bibliotek cyfrowych. Abstrakcyjny schemat referencyjny dotyczy kategoryzacji systemu na trzy poziomy konceptualizacji: 1) rozpoznawalną (namacalną) bibliotekę cyfrową, czyli zasób, z którym użytkownik podejmuje różnorodne interakcje, 2) system biblioteki cyfrowej, stanowiący wielorakie aplikacje, zainstalowane oraz indywidualnie skonfigurowane oraz 3) system zarządzania biblioteką cyfrową, który odwzorowywany jest przez odpowiednie oprogramowanie, konieczne do stworzenia, zaimplementowania oraz zarządzania biblioteką [Candela i in., dok. elektr.].

Modele mentalne tworzone są przez różnych użytkowników (każda jednostka ma własne reprezentacje mentalne) zaangażowanych w procesy projektowania, zarządzania i użytkowania bibliotek cyfrowych. Twórcy kreują modele pojęciowe (*conceptual model*) systemu oraz modele użytkownika (*user model*). Modele mentalne systemów bibliotek cyfrowych powstają z kolei w umysłach użytkowników końcowych. Model użytkownika stanowi stereotypowy sposób myślenia o adresacie (konsumencie) systemu, stworzony przez twórcę, zaś pojęciowe wyobrażenie systemu, wizja powstająca w umyśle użytkownika stanowi unikalną, specyficzną metaforę, wyrażającą działanie systemu oraz jego strukturę. Istotna w rozumieniu modeli mentalnych, a przez to zachowań informacyjnych użytkowników jest świadomość tego, iż osoby projektujące oraz administrujące systemem, jak również odbiorcy kreują własne wyobrażenia o sobie nawzajem oraz obiektach i narzędziach, z którymi się stykają. Obrazy te są oparte na wiedzy o obiektach, spełnianych przez nie funkcjach, zachodzących interakcjach, zdarzeniach występujących podczas pracy z systemem. Podczas konsytuacji modelu projektu (*design model*),

odpowiadającego wizji twórcy, uwzględniać należy informację o wiedzy i potrzebach użytkownika końcowego, scalająca reprezentacje umysłowe systemu i użytkowników [Zhang dok. elektr.; Makri, dok. elektr.].

Donald Norman wyróżnił także trzy funkcjonalne czynniki, które pojawiają się zarówno w modelu użytkowników, jak i twórców lub badaczy. Zaliczył do nich 1) system przekonań, oparty na przeświadczeniu i poglądowi związanemu z budową i działaniem systemu poprzez obserwowanie, uczenie się i wnioskowanie; 2) obserwowalność poprzez kompatybilność i zazębianie się z parametrami i stanami dostępnych modeli mentalnych użytkownika oraz wyglądem i formą rzeczywistego systemu, który można konstatować; 3) system przewidywań, który oznacza mechanizm antycypowania przez użytkownika zachowań systemu poprzez zastosowanie odpowiednich reguł konkludowania lub procesu twórczego (derywacyjnego) [Johson-Laird 2010, s. 438].

Stephen Payne analizując koncepcje modeli mentalnych Donalda Normana, stwierdził istnienie podstawowych komponentów, formujących te obrazowe reprezentacje. Według teorii Payne'a: 1) przekonania wpływają na prognozowanie zachowań – jednakże tylko spontanicznie stworzone wyobrażenie konstruujące model mentalny może zostać wykorzystane do interakcji i wpłynąć na zachowanie użytkownika podczas korzystania np. z systemu; 2) interferencje mogą nastąpić poprzez mentalną symulację, integrację wiedzy proceduralnej, nadaje również znaczenie obrazowym trajektoriom tworzonym w umyśle użytkownika podczas nakładania się modeli mentalnych oraz dostosowywaniu różnorodnej wiedzy; 3) reprezentacje mentalne mogą być analogiczne[9]: a) model mentalny jest symptomatycznym przykładem wizualizacji, spośród zbioru możliwych struktur umysłowych gwarantujących afektywną, satysfakcjonującą interakcję użytkownika z obiektem, systemem, b) nowa informacja niezgodna z aktualnym modelem mentalnym pobudza wyszukiwanie nowego modelu, wykorzystując równocześnie zapamiętany, sugestywny fundament interakcji, c) zawartość modelu mentalnego częściowo zdeterminowana jest poprzez funkcje, jakie pełni; 4) modele mentalne tworzone są na podstawie subiektywnego (fizycznego) doświadczenia [Makri, dok. elektr.].

John M. Carroll i Judith Reitman Olson wymieniają trzy aspekty, komponenty systemu, które podczas analizowania interaktywnych struktur zostały wyselekcjonowane z modelu mentalnego. Moduły, które można także odnieść do bibliotek cyfrowych, są to: 1) wiedza o zadaniu (*task-knowing*), która dotyczy znajomości celu oraz poszczególnych zadań, wytycznych koniecznych do osiągnięcia wyznaczonego zamierzenia (celu); 2) interfejs systemu (*system inteface*) oraz wiedza o tym, w jaki sposób osiągnąć kolejno podzadania (*sub-task*), pozwalająca również na wizualizację danych oraz interakcje semantyczne (językowe) z systemem; 3) architektura systemu (*system architecture*), oznaczającą wiedzę o strukturze danych w systemie, przechowywaniu zasobów, a także wszelkich zewnętrznych procesach wymaganych podczas interakcji oraz ogólnym funkcjonowaniu systemu [Carroll, Olsen 1988, s. 45-48].

[9] W ujęciu Johnsona-Lairda także pojawia się teza o analogicznej konfiguracji modeli mentalnych (Johnson-Laird, P.N. (1989). *Mental models.* W: Michael I. Posner (Ed.), *Foundations of cognitive science* (pp. 469-499). Cambridge, MA: MIT Press.

Kategorie modeli mentalnych

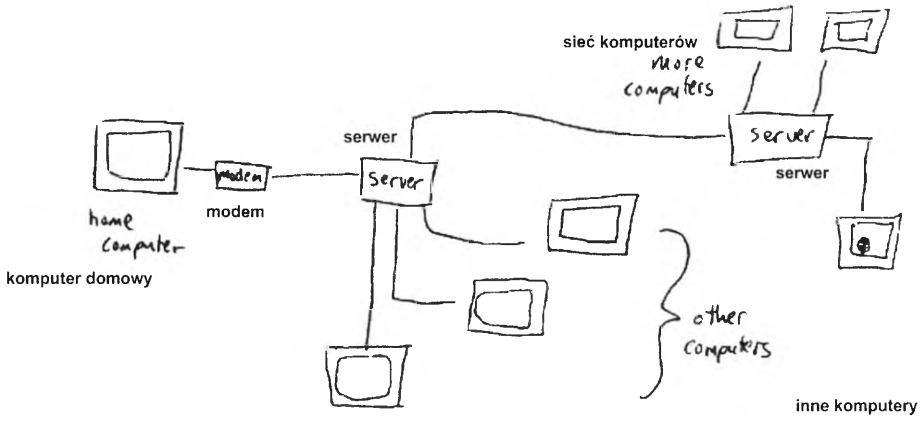
Richard Young wyróżnił osiem prototypowych wyobrażeń oraz kryteriów modelu mentalnego dla obiektów, interaktywnych systemów, które stanowią: 1) silną, wyraźną analogię (*strong analogy*), 2) substytut (*surrogate*), imitację systemu, obiektu, serwisu (określający częściowo mechanizmy działania) 3) odwzorowanie, zadaniowo-funkcjonalne mapowanie (*mapping, task action mapping*), 4) koherencja, spójność (*coherence*), 5) leksyka (*vocabulary*), 6) przestrzeń zadaniową (*problem space*), 7) gramatykę generatywną[10] (*psychological grammar*) oraz 8) typowość (*commonality*) [Zhang 2008, s. 18]. Optyka ta dotyczy różnorodnych procesów kognitywnych pojawiających się podczas asymilacji (np. poprzez wyraźny ekwiwalent i relacje w stosunku do innych systemów) i adaptacji (przyswojenia i komplementarnego zrozumienia systemu) rzeczywistych zjawisk, obiektów, systemów.

Andrea diSessa, podobnie jak Philip Johnson-Laird, rozróżnił strukturalne (uniwersalne, niezależne, charakteryzujące wewnętrzną strukturę systemu) i funkcjonalne (stanowiące zbiór informacji o działaniu systemu, odnoszący się do konkretnego zadania, modele mentalne oraz zasugerował koncepcję dualnych, rozproszonych modeli (*distributed models*) charakteryzujących się istnieniem różnorodnych danych poprzez które, na podstawie doświadczenia oraz indywidualnej wiedzy, użytkownik określa w umyśle obiekt, system, serwis [Zhang 2008, s.19-20].

Reprezentacje umysłowe systemów, bibliotek cyfrowych można zhierarchizować według istotnych kategorii na: a) modele zastępcze (*surrogate models*), b) metafory (*metaphors*), c) modele tzw. szklanej skrzynki (*glass box*) będącej przenośnią koncepcji transparentnego programowania, przejrzystości struktury, architektury informacji w systemie, zrozumienia funkcjonowania oraz jego racjonalnego zastosowania [Zhang 2008, s. 20; Subiekta 1999, s. 26, s. 195].

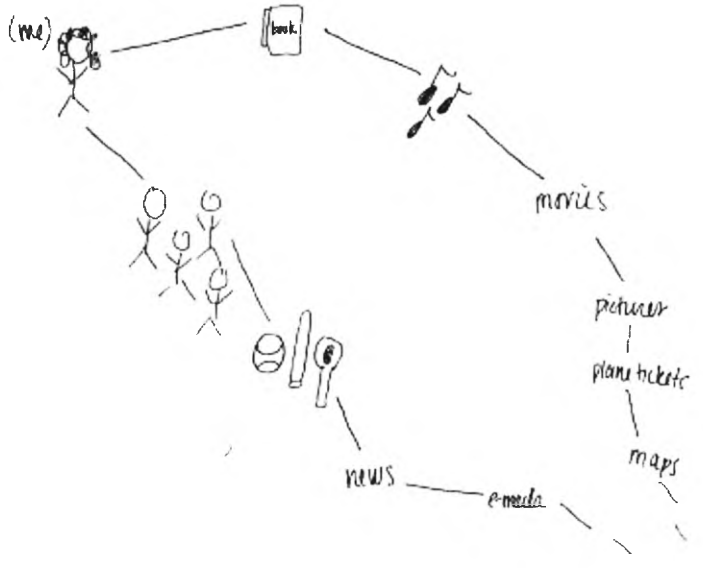
Modele mentalne mogą być postrzegane z różnych perspektyw, a przez to uszeregowane w poszczególne kategorie. Badania nad obrazowaniem systemów informacyjnych lub sieci pozwoliły m.in. Yanowi Zhang na ocenę reprezentacji umysłowych oraz skategoryzowanie ich w: 1) perspektywie technicznej (*technical view*) – w której najważniejsze elementy to system komputerów, serwerów, narzędzia komunikacji i informacji oraz użytkownicy;

[10] Psychologia gramatyki odnosząca się do procesów i mechanizmów tworzących lub/i przetwarzających struktury językowe; będąca reakcją kognitywnych struktur na bodźce językowe (stanowiąca mechanizm – w ujęciu Noama Chomskiego [Paweł Mecner (2007). Elementy gramatyki umysłu. Kraków, Wyd. Universitas, 272 p.]



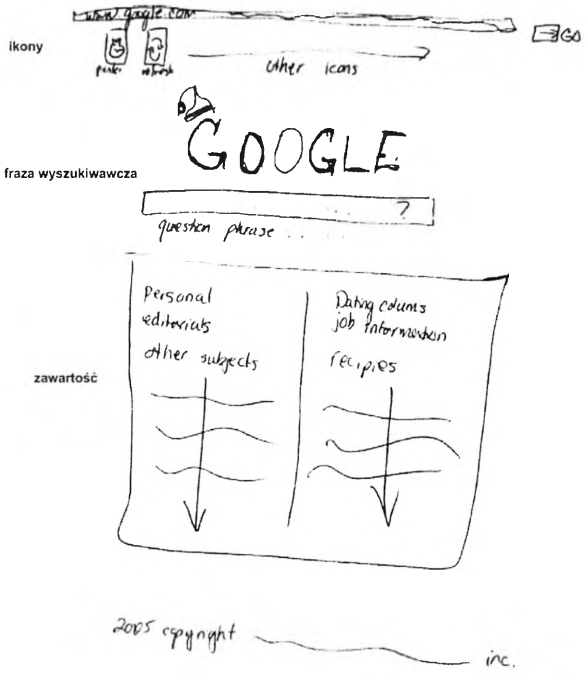
Rys. 1. Pryzmat techniczny modelu mentalnego [Za Y.Zhang, 2008, p. 1335]

2) perspektywie funkcjonalnej (*functional view*), reprezentującej różnego rodzaju usługi, a także proces wyszukiwania informacji;



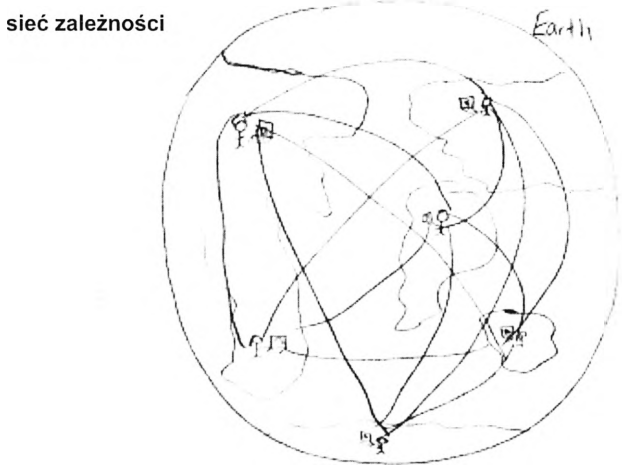
Rys. 2. Funkcjonalny model mentalny [Za Y.Zhang, 2008, p. 1335]

3) perspektywie przekształceń (*processed view*), gdzie system lub serwis traktowane jest jako centrum wszelkich procesów, interakcji, do którego przynależą różnorodne elementy np. biblioteki cyfrowej;



Rys. 3. Aktywizujący model mentalny – perspektywa procesu zachodzącego w systemie, obiekcie, sieci. [Za Y.Zhang, 2008, p. 1335]

4) perspektywie zależności (*connection view*), która uwzględnia specyfikę powiązań pomiędzy poszczególnymi elementami, z których wynikają procesy, interakcje oraz podejmowane przez użytkownika działania [Zhang 2008, s. 1334].

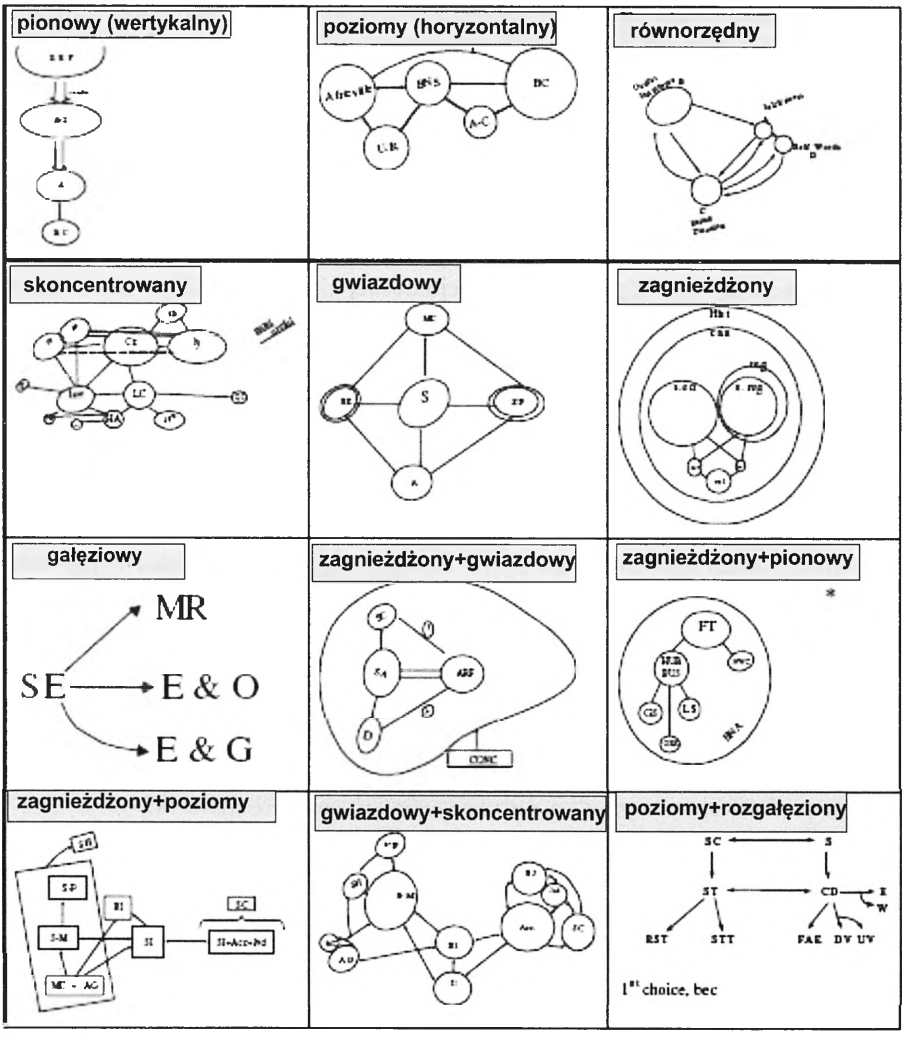


Rys. 4. Pryzmat zależności modelu mentalnego [Za Y.Zhang, 2008, p. 1335]

Charles Cole wraz z Yang Lin, Johnem Leide, Andrew Largem, and Jamshidem Beheshti skategoryzowali poszczególne obrazowania, warianty modeli mentalnych w klasy i określili je jako: 1) wertykalne, poziome – gdzie istnieje więcej warstw, powiązań semantycznych i metafor poziomych niż pionowych, 2) horyzontalne, pionowe – w tym przypadku przewaga dotyczy warstw poziomych nad wertykalnymi, 3) paralelne, równorzędne – wówczas model mentalny zostaje przedstawiony w równorzędnych warstwach poziomych i pionowych, 4) radialne, skoncentrowane – w tej kategorii model mentalny skupiony zostaje wokół centralnie wizualizowanego konceptu, który rozgałęzia się na mniej istotne, aż do przypadkowych skojarzeń, przekonań, idei, 5) skoncentrowane, systemowe (klasterowe) – wówczas pojawia się kilka stref, obszarów koncepcyjnych, które mogą być pogrupowane w przestrzenie, rzadko kiedy połączone ze sobą, 6) zagnieżdżone – kategoria ta została przedstawiona poprzez umiejscowienie konceptów jeden pośród drugich, 7) rozgałęzione (drzewo) – reprezentacja obrazowa może być pionowa lub pozioma, rozpoczynająca się od jednego, szerszego konceptu, rozgałęziając się wertykalnie lub horyzontalnie w mniejsze podkategorie, pośrednie idee, przekonania, koncepcje[Cole i in. 2007, s. 2097]. Przykłady obrazowania modeli mentalnych, zaproponowane przez Colea i jego zespół, które charakteryzują procesy wyszukiwania informacji w sieci, zostały przedstawione na Rys. 1. Należy zwrócić uwagę, że użytkownicy[11] przedstawiali subiektywne modele mentalne rzeczywistych obiektów poprzez większe okręgi reprezentujące istotne koncepcje, metafory, a także mniejsze, określające mało znaczące idee. Uczestnicy grupowali relatywne koncepcje oraz szeregowali w pewnym oddaleniu mało istotne przekonania, zamysły. Tylko niektóre reprezentacje określały zależności pomiędzy metaforami, konceptami poprzez zastosowanie wskaźników (strzałek).

W oparciu o typologię Yan Zhang, Lucy Holman analizując modele mentalne procesu wyszukiwania użytkowników (należących do Generacji Y, Milenium, aktywnie korzystających z nowych form komunikacji i informacji), posiadających doświadczenie w zakresie zastosowania zasobów cyfrowych, wyznaczyła przewagę trzech typów reprezentacji umysłowych. Zaliczyła do nich: 1) modele w ujęciu proceduralnym (process view) – postrzegające rzeczywiste obiekty, system, serwis jako całościowy proces składający się z diagramu zadań, strategii wyszukiwawczych i wszelkich procesów informacyjnych; 2) modele w ujęciu hierarchicznym (hierarchical view), gdzie wyszczególniony zostaje główna problematyka, kategoria, temat, zasób oraz podkategorie, zagadnienia mieszczące się w obszarze konkretnego obiektu, systemu, metafory; 3) model w ujęciu schematycznym, systemu powiązań (network view), który obrazowany jest przez strukturę diagramu relatywnych określeń, metafor, idei, konceptów [Holman 2011, s. 23].

[11] Badania zostały przeprowadzone przez Charlsa Colea, Yang Lin, Johna Leidę, Andrew Largea i Jamshida Behesthi wśród 80 studentów studiów I stopnia (w 3 seriach – pozwoliło to na zgromadzenie 240 modeli mentalnych), reprezentujących 5 fakultetów (nauki społeczne) na Uniwersytecie McGill w Montrealu, Kanadzie [Cole i in. 2007, 2095-2096].



Rys. 5. Kategoryzacja modeli mentalnych według Charlesa Colea i in.
[Za: Ch. Cole i in. 2007, p. 2098]

Podsumowanie

W artykule przedstawiono najistotniejsze problemy mieszczące się w obszarze interdyscyplinarnych badań nad modelami mentalnymi, prowadzonych w informatologii. Zostały one przedstawione w perspektywie kognitywistycznej, ze szczególnym uwzględnieniem ujęcia stosowanego w dziedzinie interakcji człowiek-komputer, procesów informacyjnych oraz konieczności zastosowania reprezentacji umysłowych

w tworzeniu, implementowaniu i doskonaleniu bibliotek cyfrowych. Systemy te, stanowiąc pewnego rodzaju narzędzie wspomagają intelektualne interakcje, działania mające na celu realizację przez użytkownika zamierzonych celów i zadań poprzez wielorakie procesy. Stanowiąc niejako ośrodek rozumowych, konceptualnych aktywności, odzwierciedlana jest w różnorodnych modelach mentalnych, obrazowaniach umysłowych wszystkich interesariuszy – administratorów, bibliotekarzy, programistów, teoretyków, wreszcie odbiorców, użytkowników końcowych. Modele te umożliwiają zrozumienie funkcji i zadań bibliotek cyfrowych, udoskonalają je oraz pozwalają na implementacje najlepszych referencyjnych struktur wraz z cyfrowymi zasobami, narzędziami, serwisami. Bardzo ważne staje się analizowanie specyfiki wiedzy, którą dysponuje użytkownik, reprezentując rzeczywistość w umyśle, podejmując odpowiednie zachowania informacyjne poprzez zastosowanie reprezentacji zjawisk, obiektów, systemów, serwisów, wytworzonych wewnętrznie. Perspektywa użytkownika staje się niezwykle istotnym komponentem wdrażania, realizowania polityki digitalizacji repozytoriów, a także ewaluowania bibliotek cyfrowych. W komplementarnych oraz interdyscyplinarnych badaniach podejmowanych w informacji naukowej nad tymi systemami informacyjno-wyszuiwawczymi, analiza modeli mentalnych oraz rozwój poznawania tych struktur stanowi preludium do holistycznej eksploracji zasobów cyfrowych, jak również do wzmocnienia orientacji i wiedzy o samych użytkownikach, potrzebach informacyjnych, interakcjach i zachowaniach podejmowanych przez nich.

Bibliografia

1. Bales, Stephen; Rieger, Jason; Zhang, Yan (2004). *Survey of Learners' Knowledge Structures: Rationales, Methods and Instruments* [dok. elektr.] http://ils.unc.edu/~yanz/cogmapping.pdf [odczyt: 12.12.2011].
2. Belkin, Nicholas J. (2005). *Anomalous State of Knowledge.* [W:] Karen E. Fisher, Sanda Erdelez, Lynne (E.F.) McKechnie red. *Theories of Information Behavior.* New Jersey: ASIST Monograph Series, Information Today, s. 44-47.
3. Bilal, Dania; Wang, Peiling (2005). *Children's Conceptual Structures of Science Categories and the Design of Web Directories.* "Journal of the American Society for Information Science and Technology" 56(12), s. 1303-1313.
4. Billinghurst, Mark; Weghorst, Suzanne (1995). *The Use of Sketch Maps to Measure Cognitive Maps of Virtual Environments.* [dok. elektr.] ftp://www.hitl.washington.edu/pub/publications/p-94-1/p-94-1.htm [odczyt: 9.01.2012].
5. Borgman, Christine L. (1985). *The users mental model of an information retrieval system.* [dok. elektr.] http://web.simmons.edu/~benoit/lis466/p268-borgman.pdf [odczyt: 12.12.2011].
6. Borgman, Christine L. (1986). *The user's mental model of an information retrieval system: An experiment on a prototype online catalog.* "International Journal of Man-Machine Studies" 24(1), s. 47-64.
7. Brandt, Scott D. (2001). *Reference, Mental Models and Teaching Technology.* "The Reference Librarian" Vol. 74, s. 37-48.

8. Carroll, John M.; Olson, Judith R. (1988). *Mental models in human-computer interaction: Research Issues about What the User of Software Knows.* [W:] Helander, Martin (red.). *The Handbook of Human-Computer Interaction.* Amsterdam: North Holland Press, s. 45-65.
9. Candela, Leonardo; Nardi, Alessandro red. (2011). *The Digital Library Reference Model* [dok. elektr.] http://www.dlorg.eu/index.php/outcomes/reference-model [odczyt: 2.12.2011].
10. Candela, Leonardo [i in.] (2007). *Setting the Foundations of Digital Libraries. The DELOS Manifesto.* "D-Lib Magazine" Vol. 13 nr 3/4 [dok. elektr.] http://www.dlib.org/dlib/march07/castelli/03castelli.html [odczyt: 09.01.2012].
11. Carley, Kathleen; Palmquist, Michael (1992). *Extracting, reprezenting and analyzing mental models.* "Social Forces" 70 (3), s. 601-636 [dok. elektr.] http://www.casos.cs.cmu.edu/publications/papers/carley_1992_extractingrepresenting.PDF [odczyt: 20.12.2011].
12. Casarosa, Vittore (2007). *DELOS Reference Model for Digital Libraries* [dok. elektr.] http://elag2007.upf.edu/papers/casarosa.pdf [odczyt: 2.01.2012].
13. Chalmers, Patricia A. (2003). *The role of cognitive theory in human-computer interface.* "Computers in Human Behavior" Vol. 19 nr 5, s. 593-607.
14. Chowdhury, Gobinda (2010). *From digital libraries to digital preservation research: the importance of users and context.* "Journal of Documentation" Vol.66 nr 2, s. 207-223.
15. Cole, Charles Lin; Yang, Leide; John, Large; Andrew; Beheshti, Jamshid (2007). *A classification of mental models of undergraduates seeking information for a course essay in history and psychology: preliminary investigations into aligning their mental models with online thesauri.* "Journal of The American Society for Information Science and Technology" 58(13), s. 2092-2104.
16. Comeaux, David J. (2008). *Usability Studies and User-Centered Design in Digital Libraries.* "Journal of Web Librarianship" vol. 2 nr 2-3, s. 457-475.
17. Darabi, Aubteen A.; Nelson, David W.; Seel, Norbert M. (2009). *Progression of mental models throughout the phases of a computer-based instructional simulation: Supportive information, practice, and performance.* "Computers in Human Behavior" vol. 25 nr 3, s. 723-730.
18. Dimitroff, Alexandra (1992) *Mental models theory and search outcome in a bibliographic retrieval system.* "Library and Information Science Research" 14(2), s. 141-156.
19. Efthimiadis, Efthimis N.; Hendry, David G. (2005). *Search engines and how students think they work.* [W:] *Proceedings of SIGIR'05.* New York: ACM, s. 595-596.
20. Frias-Martinez, Enrique; Chen, Sherry Y.; Macredie, Robert D.; Liu, Xiaohui (2007). *The role of human factors in stereotyping behavior and perception of digital library users: a robust clustering approach.* "User Model User-Adapted Interaction" 17, s. 305-337.
21. Fuhr, Norbert [i in.] (2007). *Evaluation of digital libraries* [dok. elektr.] "International Journal of Digital Libraries". http://www.scribd.com/doc/185523/Evaluation-of-Digital-Libraries-Fulltext [odczyt: 29.12.2011].
22. He, Wu; Erdelez, Sanda; Wang, Feng-Kwei; Shyu, Chi-Ren (2008). *The effects of conceptual description and search practice on users' mental models and information seeking in a case-based reasoning retrieval system.* "Information Processing & Management" Vol. 44 nr 1, s. 294-309.
23. Holland, John [i in.] (1989). *Induction: processes of inference, learning, and discovery.* Massachusetts: MIT Press, 398 s.
24. Holman, Lucy (2011). *Millennial Students' Mental Models of Search: Implications for Academic Librarians and Database Developers.* "The Journal of Academic Librarianship" vol. 37 nr 1, s. 19-27.

25. Hsu, Jack S.C.; Chang, Jamie Y.T; Klein, Gary; Jiang, James J. (2011). *Exploring the impact of team mental models on information utilization and project performance in system development.* "International Journal of Project Management" 29, s. 1-12.
26. Izard, Carroll E. (2009). *Emotion Theory and Research: Highlights, Unanswered Questions, and Emerging Issues* [dok. elektr.] http://www.scribd.com/doc/51919582/Izard-2009--Emotion-Theory [odczyt: 29.12.2011].
27. Johnson-Laird, Philip N. (1989). *Mental models.* [W:] Michael I. Posner (red.). *Foundations of cognitive science.* Cambridge, MA: MIT Press, s. 469-499.
28. Johnson-Laird, Philip N. (2001). *Mental models and deductive reasoning.* "Trends in Cognitive Science" 5 (8), s. 434-442.
29. Jonassen, David H.; Beissner, Katherine; Yacci, Michael (1993). *Structural knowledge: techniques for representing, conveying, and acquiring structural knowledge.* New Jersey, Hillsdale: Lawrence Erlbaum Associates Inc., 265 s.
30. Kahler, Susan Elizabeth (2002). *A comparison of knowledge acquisition methods for the elicitation of procedural mental models* [dok. elektr.] http://repository.lib.ncsu.edu/ir/bitstream/1840.16/4900/1/etd.pdf [odczyt: 20.12.2011].
31. Katzeff, Cecilia (1990). *System demands on mental models for a fulltext database.* "International Journal of Man-Machine Studies" 32, s. 483-509.
32. Kerr, Stephen T. (1990). *Wayfinding in an electronic database: the relative importance of navigational cues VS mental models.* "Information Processing & Management" 26(4), s. 511-523.
33. Lim, Youn-Kyung; Blevis, Eli; Stolterman, Erik (2007). *Grand Challenges in Design Research for Human-Centered Design Informatics.* [W:] Douglas Schuler (red.): *Online Communities and Social Computing*, HCII 2007, LNCS "Lecture Notes in Computer Science" Vol. 4564, s. 106-115.
34. Madle, Gemma; Kostkova, Patty; Roudsari, Abdul, (2008). *Impact-ED: The development of a new Digital Library Evaluation Model.* [W:] Birte Christensen-Dalsgaard, Donatella Castelli, Bolette Ammitzbøll Jurik, Joan Lippincott (red). *Research and Advanced Technology for Digital Libraries.* "Lecture Notes in Computer Science" Vol. 5173, s.100-105.
35. Makri, Stephann (2004). *Investigating users' mental models of traditional and digital libraries. Master of Science (Human-Computer Interaction with Ergonomics) thesis in the Faculty of Life Sciences, University College London*, 2004, s. 142 [dok. elektr.] http://discovery.ucl.ac.uk/231312/ [odczyt: 12.12.2011].
36. Makri, Stephan [i in.] (2007). *A library Or Just another information resource? A case study of users' mental models of traditional and Digital libraries.* "Journal of the American Association for Information Science and Technology" 58 (3), s. 433-445.
37. Morganti, Francesca; Carassa, Antonella; Geminiani, Giuliano (2007). *Planning optimal paths: A simple assessment of survey spatial knowledge in virtual environments.* "Computers in Human Behavior" Vol. 23 nr 4, s. 1982-1996.
38. Muramatsu, Jack; Pratt, Wanda (2001). Transparent queries: investigating users' mental models of search engines. [W:] *Proceedings of the twenty-fourth international ACM conference on research and development in information retrieval SIGIR* 2001. New York: ACM, s. 217-224.
39. Nęcka, Edward; Orzechowski, Jarosław; Szymura, Błażej (2008). *Psychologia poznawcza.* Warszawa: Academica Wydaw. SWPS, Wydaw. Naukowe PWN, 740 s.
40. Norman, Donald A. (2002). *Emotion & design: attractive things work better.* "Interactions" 9 (4), s. 36-42.

41. Pavio, Allan. *Dual Coding Theory and Education* [dok. elektr.] http://www.umich.edu/~rdytolrn/pathwaysconference/presentations/paivio.pdf [odczyt: 12.12.2011].
42. Payne, Stephen J. (2003) *Users' Mental Models: The Very Idea.* [W:] John M. Carroll (red.). *HCI Models, Theories, and Frameworks.* San Francisco: Morgan Kaufmann, s.135-156.
43. Péruch, Patrick; Belingard, Loïc; Thinus-Blanc, Catherine (2000). *Transfer of Spatial Knowledge from Virtual to Real Environments. Spatial Cognition* II "Lecture Notes in Computer Science" Vol. 1849/2000, s. 253-264.
44. Piłat, Robert (2000). *Umysł jako model świata* [dok. elektr.] http://www.robertpilat.republika.pl/umysl.pdf [odczyt: 12.12.2011].
45. Roth, Sandra P.; Schmutz, Peter; Pauwels, Stefan L.; Bargas-Avila, Javier A.; Opwis, Klaus (2010). *Mental models for web objects: where do users expect to find the most frequent objects in online shops, news portals, and company web pages.* "Interacting with Computers" 22, s. 140-152.
46. Saracevic Tefko (2004). *Evaluation of digital libraries: an overview*, paper presented at *DELOS Workshop on the Evaluation of Digital Libraries* [dok. elektr.] http://dlib.ionio.gr/wp7/WS2004_Saracevic.pdf [odczyt: 12.12.2011].
47. Slone, Debra J. (2002). *The influence of mental models and goals on search patterns during Web interaction.* "Journal of the American Society for Information Science and Technology" 53(13), s.1152-1169.
48. Subieta, Kazimierz (1999). *Słownik terminów z zakresu obiektowości.* Warszawa: Akad. Oficyna Wydawnicza PLJ, 290 s.
49. Tallman, Julie I.; Henderson, Lyn (1999). *Constructing Mental Model Paradigms for Teaching Electronic Resources* [dok. elektr.] http://www.ala.org/aasl/aaslpubsandjournals/slmrb/slmrcontents/volume21999/vol2tallman [odczyt: 12.12.2011].
50. Westbrook, Lynn (2006) *Mental models: a theoretical overview and preliminary study.* "Journal of Information Science" 32 (6), s. 563-579.
51. Van Merriënboer, Jeroen J.G.; Seel, Norbert M.; Kirschner, Paul A. (2002). *Mental models as a new foundation for instructional design.* "Educational Technology" Vol. 42 nr 2, s. 60-66.
52. Veldof, Jerilyn; Beavers Karen (2001). *Going mental: Tackling mental models for the online library tutorial.* "Research Strategies" Vol. 18 nr 1, s. 3-20.
53. Xie, Hong I. (2006). *Evaluation of digital libraries: Criteria and problems from users' Perspectives.* "Library and Information Science Research" 28, s. 433-452.
54. Xie, Hong Iris (2008). *Users' evaluation of digitallibraries (DLs): Their uses, their criteria, and their assessment.* "Information Processing & Management" Vol. 44 nr 3, s. 1346-1373.
55. Young, Richard M. (1983). *Surrogates and mappings: Two kinds of conceptual models for interactive devices.* [W:] Dedre Gentner, Albert L. Stevens (red). *Mental models.* New Jersey, Hillsdale: Lawrence Erlbaum Associates, s. 43-52.
56. Zhang, Xiangmin; Chignell, Mark (2001). *Assessment of the effects of user characteristics on mental models of information retrieval systems.* "Journal of the American Society for Information Science and technology" 52(6), s. 445-459.
57. Zhang, Yan (2008). *The influence of mental models on undergraduate students' searching behavior on the Web.* "Information Processing and Management" 44, s. 1330-1345.

58. Zhang, Yan (2010). *Dimensions and elements of people's mental models of an information-rich web space.* "Journal of the American Society for Information Science and Technology" 61 (11), s. 2206-2218.
59. Zhang, Yan (2012). *The impact of task complexity on people's mental models of MedlinePlus.* "Information Processing & Management" Vol. 48 nr 1, s. 107-119.
60. Zhang, Ying. *Developing a Holistic Model for Digital Library Evaluation* [dok. elektr.] http://onlinelibrary.wiley.com/doi/10.1002/asi.21220/pdf [odczyt: 12.12.2011].

Model kompetencji zawodowych bibliotekarza cyfrowego

Łukasz Mesek

W tekście przedstawiono model kompetencji bibliotekarzy opracowany na podstawie praktyki tj. w trakcie przyjęć pracowników do Biblioteki Jagiellońskiej na stanowiska bibliotekarzy cyfrowych. Stanowiska te zostały utworzone aby zapewnić obsługę w otwartej w lipcu 2010 r. Jagiellońskiej Bibliotece Cyfrowej.

Stworzenie sprawnie działającej biblioteki cyfrowej wymaga, oprócz zapewniania infrastruktury komputerowej, powołania do życia kompetentnego zespołu fachowców. Zintegrowany system zarządzania kompetencjami pozwala na skuteczne powiązanie wszystkich obszarów zarządzania personelem w jeden spójny system, którego fundamentem są kompetencje zebrane w model. Stworzenie takiego modelu jest podstawą w doborze specjalistów, co w praktyce przekłada się na efektywność działania zespołu.

W nowoczesnym środowisku pracy jakim jest biblioteka cyfrowa kompetencje bibliotekarza musiały ulec znacznemu poszerzeniu w stosunku do kompetencji bibliotekarza zarządzającego zbiorem dokumentów w wersji drukowanej. Udostępnianie treści w postaci cyfrowej, wymaga przede wszystkim o wiele większych umiejętności związanych z wiedzą informatyczną, grafiką komputerową i społeczeństwem wirtualnym. Dlatego przedstawiony poniżej model kompetencji w praktyce generuje nowe stanowisko pracy, wraz z modelem tym tworzące się w bibliotece – stanowisko bibliotekarza cyfrowego.

W przypadku samego pojęcia kompetencji to jest ono rozumiane w sposób jak najbardziej tradycyjny [Oleksyn 2006] i definiuje sumę wiadomości, umiejętności oraz postaw i zachowań, które pozwalają wykonywać zadania na określonym poziomie zawodowości. W ramach modelowania kompetencji stworzony został opis stanowisk pracy, określono kwalifikacje zawodowe, którymi powinni wykazywać się wszyscy pracownicy. Można jednak wśród nich wyróżnić kompetencje o kluczowym znaczeniu dla organizacji biblioteki cyfrowej. Kompetencje kluczowe zostały oznaczo-

ne skrótem KL. Wśród nich zdefiniowano takie umiejętności i zachowania, których posiadanie lub przejawianie sprzyja realizacji celów strategicznych i wymagane jest od wszystkich pracowników, niezależnie od pełnionych przez nich obowiązków czy ich miejsca w strukturze organizacyjnej.

Zastosowanie modelu kompetencji stanowi podstawę systemu zarządzania kwalifikacjami zawodowymi obejmującego m.in. rekrutację i selekcję, systemy ocen, premiowania oraz ścieżki karier i programy rozwoju pracowników zatrudnionych w bibliotece cyfrowej.

Sylwetka i kompetencje bibliotekarza cyfrowego

Przedstawiony model kompetencji pracownika zatrudnionego jako bibliotekarz cyfrowy wynika z uniwersyteckiego charakteru studiów w zakresie informacji naukowej i bibliotekoznawstwa, a w szczególności specjalizacji oferowanej w ramach tych studiów jaką są zasoby cyfrowe. Dzięki podbudowie teoretycznej z zakresu bibliotekoznawstwa jaką otrzymuje w pierwszych latach studiów, absolwent jest w stanie bez większych problemów zaadaptować się do warunków pracy panujących w bibliotece naukowej. Ważna jest tu zarówno znajomość charakteru zbiorów jak i systemów bibliotecznych, organizacja procesów bibliotecznych, jak i przepisów katalogowania. W toku studiów z jednej strony została ukształtowana u przyszłego bibliotekarza wrażliwość historyczna pozwalająca na rozpoznawanie materiałów o wartości wieczystej, a z drugiej umiejętności techniczne pozwalające na cyfrowe zachowanie, obróbkę i powszechne ich udostępnienie jako dobra kultury narodowej.

Model ten został opracowany na podstawie modelu kompetencji zawodowych archiwistów i zarządców dokumentacji [Chorążyczewski, 2008] oraz praktycznych obserwacji prowadzonych w trakcie uruchamiania biblioteki cyfrowej i zatrudniania do jej obsługi nowych pracowników, wśród których znakomitą większość stanowią absolwenci informacji naukowej i bibliotekoznawstwa.

Należy zwrócić uwagę na fakt iż poszczególne kompetencje nie muszą być właściwe tylko dla absolwenta informacji naukowej i bibliotekoznawstwa, a dotyczą raczej pracownika zatrudnionego na stanowisku bibliotekarza cyfrowego. Dotyczy to zwłaszcza kwalifikacji informatycznych czy graficznych. Dla bibliotekarza cyfrowego są one konieczne. Dopiero pełny zestaw kompetencji konstytuuje zawód bibliotekarza cyfrowego, a nie którakolwiek z kwalifikacji zawodowych z osobna.

Bibliotekarz cyfrowy powinien posiadać ogólne wykształcenie humanistyczne i być przygotowanym zarówno do pracy w bibliotece jak i poza nią, do zarządzania różnego rodzaju dokumentami tak w bibliotece cyfrowej jak i na stronie portalu internetowego. Powinien mieć podstawowe przygotowanie do pracy na stanowisku bibliotekarza, niezależnie od typu i wielkości biblioteki, w zakresie gromadzenia, przechowywania, opracowywania i udostępniania materiałów bibliotecznych, zarówno w ich wersji drukowanej jak i cyfrowej, a więc do właściwej i systematycznej kontroli przyjmowania, obiegu, przechowywania i posługiwania się dokumentami.

W tym momencie uzyskanie poziomu kompetencji kwalifikujących pracownika na stanowisko bibliotekarza cyfrowego jest realizowane na studiach prowadzonych na Uniwersytecie Jagiellońskim, II stopnia (magisterskich) z zakresu informacji naukowej i bibliotekoznawstwa – ze specjalizacją zasoby cyfrowe, które stanowią uzupełnienie studiów I stopnia (licencjackich). Podane poniżej zakresy kompetencji w poszczególnych dziedzinach stanowią wypadkową całego ogromu wiedzy uzyskanej przez absolwenta w ciągu pięciu lat studiów.

Kompetencje w zakresie warsztatu badawczego dla bibliotekarza cyfrowego:

- zna i umie poprawnie stosować terminologię informacyjną i biblioteczną (KL),
- zna i umie poprawnie stosować zasady teoretyczne informacji naukowej i bibliotekoznawstwa (KL),
- zna i umie stosować zespół metod badawczych stosowanych w informacji naukowej i bibliotekoznawstwie,
- zna i stosuje technikę pracy naukowej, potrafi przygotować tekst wyposażony w aparat naukowy,
- zna język angielski w stopniu pozwalającym na korzystanie z literatury fachowej oraz bibliotek cyfrowych na świecie.

Kompetencje w zakresie bibliotekoznawstwa i historii bibliotek:

- zna podstawowe zjawiska i procesy zachodzące w bibliotece (KL),
- zna różne formy dokumentów z różnych okresów historycznych i sposoby ich powstawania,
- zna dzieje bibliotek, rozumie ich wpływ na współczesne rozmieszczenie zasobów bibliotecznych i potrafi wykorzystać tę wiedzę w przypisaniu danego dokumentu do właściwej kolekcji.

Kompetencje w zakresie obiegu dokumentów w bibliotece:

- zna i umie stosować zasady „records management" (KL),
- potrafi obserwować i opisać obieg dokumentu wewnątrz i na zewnątrz biblioteki (KL),
- umie zorganizować i prowadzić księgozbiór bieżący,
- zna i potrafi obsługiwać systemy informatyczne automatyzujące obieg dokumentów w bibliotece,
- zna i potrafi stosować zasady i metody logistyki dla potrzeb biblioteki,
- zna problematykę tworzenia i przechowywania dokumentów elektronicznych; potrafi uczestniczyć w procesach tworzenia i zarządzania dokumentami elektronicznymi na wszystkich etapach ich życia.

Kompetencje w zakresie gromadzenia i kształtowania zasobu bibliotecznego:

- zna i umie stosować zasady, kryteria, metody i narzędzia selekcji dokumentów (KL),
- zna i umie stosować formy i narzędzia kształtowania narastającego zasobu bibliotecznego,
- zna i umie stosować sposoby gromadzenia (pozyskiwania) zasobu bibliotecznego,

- zna zasady rozmieszczenia zasobu bibliotecznego i umie rozwiązywać praktyczne problemy wynikające z ich stosowania.

Kompetencje w zakresie przechowywania dokumentów:

- zna ogólne zasady projektowania budynków bibliotecznych i wyposażenia magazynów,
- zna możliwości i metody profilaktyki i konserwacji dokumentów; umie rozpoznawać rodzaj zagrożeń dla materiałów bibliotecznych,
- zna i umie stworzyć odpowiednie warunki przechowywania wszystkich rodzajów materiałów bibliotecznych,
- zna i umie stosować scenariusze działań w sytuacjach kryzysowych.

Kompetencje w zakresie opracowania bibliograficznego dokumentu:

- zna i potrafi stosować standardy metadanych dla dokumentów drukowanych (KL),
- zna i potrafi stosować standardy metadanych dla dokumentów elektronicznych (KL),
- zna i potrafi stosować standardy regulujące postępowanie z dokumentami na każdym etapie ich życia (KL),
- zna i umie stosować zasady, metody, narzędzia i etapy opracowania dokumentu (KL),
- jest przygotowany do samodzielnego opracowania dokumentów współczesnych w formacie MARC 21.

Kompetencje w zakresie działalności informacyjnej i udostępniania dokumentów:

- zna i poprawnie stosuje zasady, metody i narzędzia nowoczesnego udostępniania dokumentów (KL),
- zna i rozumie zasady powstawania i funkcjonowania systemów informacyjnych (KL),
- zna i umie wykorzystywać nowoczesne technologie informacyjne i komunikacyjne, w tym systemy informatyczne stosowane w zarządzaniu informacją i dokumentami (KL),
- zna zasady konstruowania strony internetowej i potrafi ją skonstruować dla potrzeb bibliotecznych (KL),
- potrafi wyszukiwać informacje w sieciach informatycznych (KL),
- zna i potrafi stosować zasadę swobodnego i równego dostępu do zasobu bibliotecznego przy zachowaniu prawa do prywatności (KL),
- przestrzega zasady równości wszystkich użytkowników (KL),

Kompetencje w zakresie prawa autorskiego, prasowego i wydawniczego.

- zna podstawowe pojęcia z zakresu prawa autorskiego, prasowego i wydawniczego (KL),
- zna normy prawne regulujące postępowanie z dokumentami (KL),
- zna i umie wykorzystywać przepisy prawne obowiązujące obecnie, jak i w przeszłości zarówno w Polsce jak i na świecie,

- zna zasady na jakich mogą być udostępniane dokumenty w sieci rozległej i ograniczenia prawne regulujące udostępnianie tylko na poziomie biblioteki.

Kompetencje w zakresie procesu obróbki i publikacji kopii cyfrowych dokumentów:

- zna i potrafi stosować zasady i techniki digitalizacji dokumentów (KL),
- potrafi kontrolować jakość i kompletność kopii cyfrowych dokumentów uzyskanych w procesie digitalizacji (KL),
- zna zasady obróbki graficznej kopii cyfrowych dokumentów (KL),
- zna zasady konwersji kopii cyfrowych dokumentów z formatu master do formatu publikowalnego (KL),
- zna zasady redagowania opisu bibliograficznego w formacie Dublin Core (KL),
- zna zasady tworzenia i zarządzania publikacjami na platformie biblioteki cyfrowej (KL),
- zna zasady przygotowania kopii cyfrowych dokumentów do zabezpieczenia w postaci mikrofilmów,
- potrafi obsłużyć pracowników w zakresie udostępniania kopii cyfrowych zabezpieczonych dokumentów.

Kompetencje w zakresie ogólnych postaw i zachowań (KL):

- zna, akceptuje i potrafi stosować postawy i zachowania pożądane w zawodzie bibliotekarza,
- nie rozpowszechnia informacji uzyskanych w trakcie pracy z dokumentami, przestrzegając tajemnicy służbowej swojego pracodawcy,
- nie wykorzystuje swojego uprzywilejowanego dostępu do dokumentów dla osiągnięcia korzyści osobistych jakiegokolwiek rodzaju,
- odznacza się nastawieniem na ustawiczne podnoszenie swoich kwalifikacji
- potrafi pracować w grupie,
- wykazuje się solidnością i dbałością o szczegóły oraz odpowiedzialnością za powierzony do pracy materiał,
- dba o terminowe i prawidłowe wykonanie powierzonych zadań służbowych i poleceń przełożonych zgodnie z obowiązującymi przepisami,
- przestrzega przepisów o ochronie danych osobowych,
- przestrzega przepisów o ochronie tajemnicy służbowej i państwowej,
- przestrzega porządku i dyscypliny pracy,
- przestrzeganie przepisów oraz zasad BHP i przeciw pożarowych,
- należycie dba o powierzone mienie.

Zastosowanie modelu kompetencji zawodowych bibliotekarza cyfrowego w praktyce.

Powyższy model kompetencji został zastosowany dwukrotnie podczas rekrutacji pracowników przy tworzeniu zespołu dla nowo powstałej Jagiellońskiej Biblioteki Cyfrowej.

Pierwszym etapem rekrutacji było wypełnienie przez petenta formularza rekrutacyjnego, którego wzór został przedstawiony na rys. 1.

Formularz rekrutacyjny

Data wypełnienia .....

Staram się o stanowisko:

Imię | Nazwisko | Data urodzenia

Adres do korespondencji

Adres e-mail | Telefon:

WYKSZTAŁCENIE

NAZWA SZKOŁY	**Specjalność**	**Okres nauki**

SZKOLENIA, KURSY	**Nazwa firmy szkolącej**	**Rok ukończenia**

ZNAJOMOŚĆ PROGRAMÓW KOMPUTEROWYCH	**Nazwa programu/pakietu**	**Stopień zaawansowania**

ZNAJOMOŚĆ JĘZYKÓW OBCYCH	**Stopień opanowania języka** (biegły, zaawansowany, komunikatywny, podstawowy)

PRZEBIEG ZATRUDNIENIA (ostatnie 5 lat)

Nazwa i siedziba pracodawcy	Okres zatrudnienia	Stanowisko	Główne obowiązki

Oczekiwania finansowe (wynagrodzenie zasadnicze, kwota brutto): .....

O cyfryzacji zbiorów BJ i tworzeniu Jagiellońskiej Biblioteki Cyfrowej dowiedziałem/am się z:

....

.....

Dodatkowe informacje (np. rzadkie umiejętności, kwalifikacje i zainteresowania):

OŚWIADCZENIE

Wyrażam zgodę na przetwarzanie moich danych osobowych zawartych w ankiecie dla potrzeb procesu zatrudnienia zgodnie z Ustawą z dnia 29.08.1997 roku o Ochronie Danych Osobowych (Dz. Ust. Nr 133, poz. 883).

.....

Data i podpis wypełniającego ankietę

UWAGI OSOBY ZAJMUJĄCEJ SIĘ REKRUTACJĄ:

Rys. 1. Formularz rekrutacyjny na stanowisko bibliotekarza cyfrowego w Bibliotece Jagiellońskiej

Na podstawie wypełnionych formularzy rekrutacyjnych została przeprowadzona wstępna selekcja i na rozmowę kwalifikacyjną zostały zaproszone osoby, które osiągnęły największą liczbę punktów. Na potrzeby wstępnej selekcji i następującej po niej rozmowy kwalifikacyjnej został opracowana tabela 1:

Tabela 1. Tabela oceny pracownika w trakcie rozmowy kwalifikacyjnej

imię	nazwisko	rok urodzenia	telefon	rok rozpoczęcia pracy	wykształcenie	doświadczenie zawodowe	predyspozycje interpersonalne		punkty razem	oczekiwane wynagrodzenie brutto	uwagi
							oceniający 1	oceniający 2			

która wraz z przedstawioną poniżej punktacją pozwoliła na wybranie najlepszych kandydatów.

		wykształcenie		**predyspozycje interpersonalne**	
zaproszenie na rozmowę	od 3 pkt.	kierunkowe	2 pkt.	maksymalne	3 pkt.
		inne/zbyt wysokie	0-1 pkt.	umiarkowane	2 pkt.
przyjęcie pod warunkiem uzyskania MGR		**doświadczenie zawodowe**		minimalne	1 pkt.
		właściwe dla stanowiska	3 pkt.		

Do etapu rozmowy kwalifikacyjnej zostali dopuszczeni tylko kandydaci, którzy wykazali się posiadaniem kompetencji kluczowych. Wśród nich największą grupę stanowili absolwenci lub studenci piątego roku kierunku informacja naukowa i bibliotekoznawstwo, głównie wybierający jako kierunek specjalizacji zasoby cyfrowe, broker informacji lub elektroniczne publikowanie. Drugą pod względem liczebności była grupa absolwentów historii o specjalizacji archiwistyka i zarządzanie dokumentacją. Bardzo ważnym czynnikiem jaki miał wpływ na wybór kandydatów, oprócz ich odpowiedzi na pytania dotyczące kompetencji w zakresie ogólnych postaw i zachowań, były prezentowane w czasie rozmowy kwalifikacyjnej portfolia. Należy nadmienić, że większość prezentacji przedstawiana była tylko w postaci elektronicznej, co również dawało oceniającym wgląd w kwalifikacje zawodowe przyszłych pracowników. Sposób prezentacji i umiejętność omówienia przedstawianych przez kandydatów projektów był kolejnym kryterium, które oceniającym dawało pewność, że praca została wykonana samodzielnie.

Podsumowując, w wyniku naboru pracowników mających tworzyć główny zrąb zespołu tworzącego Jagiellońską Bibliotekę Cyfrową wybrano:

- 8 kandydatów będących absolwentami kierunki Informacja Naukowa i Bibliotekoznawstwo, którzy ukończyli specjalizację zasoby cyfrowe
- 1 kandydat studiów polonistycznych, który ukończył specjalizację edytorstwo
- 1 kandydat studiów historycznych, który ukończył specjalizację archiwistyka i zarządzanie dokumentacją.

Na przestrzeni czasu jaki minął od momentu pierwszej rekrutacji, okazuje się iż pracownicy kształceni zgodnie z programem studiów, a w szczególności programem specjalizacji zasoby cyfrowe oferowanym przez Instytut Informacji Naukowej i Bibliotekoznawstwa Uniwersytetu Jagiellońskiego w pełni wywiązują się z zadań stawianych przed dynamicznie rozwijającą się biblioteką cyfrową, a po wdrożeniu w specyfikę zbiorów i systemu ich organizacji w Bibliotece Jagiellońskiej w pełni zasługują na stanowisko bibliotekarzy cyfrowych.

W artykule przedstawiona została próba stworzenia modelu kompetencji na podstawie pożądanych u bibliotekarza cyfrowego kwalifikacji zawodowych. Model o tyle ciekawy, że bardzo szybko możliwy do przetestowania w praktyce, dzięki dotacjom unijnym na stworzenie Jagiellońskiej Biblioteki Cyfrowej. Zaś przybliżenie poszczególnych etapów procesu rekrutacji, być może pozwoli absolwentom wchodzącym na rynek pracy oswoić pierwszy stres i dobrze przygotować się do miłej rozmowy kwalifikacyjnej.

Bibliografia

1. Oleksyn, Tadeusz (2006). *Zarządzanie kompetencjami: teoria i praktyka*. Kraków: Oficyna Ekonomiczna, 300 s.
2. Oleksyn, Tadeusz (2011). *Zarządzanie kompetencjami: teoria i praktyka przedsiębiorstw z „listy 500"* [W:] Bylok Felicjan, Słocińska Anna red. (2011). *Współczesne oblicza kapitału ludzkiego i intelektualnego.* Częstochowa: Wydawnictwo Politechniki Częstochowskiej, s. 97-123.
3. Chorążyczewski, Waldemar; Kwiatkowska, Wiesława oprac. (2008). *Model kompetencji zawodowych archiwistów i zarządców dokumentacji: rekomendacja Zarządu Głównego Stowarzyszenia Archiwistów Polskich uchwalona 4 listopada 2010 roku.* Warszawa: Stowarzyszenie Archiwistów Polskich [dok. elektr.] http://www.sap.waw.pl/node/355 [odczyt 9.08.2011].

Część 3

Zasoby i procesy

Tworzenie i rozwój bibliotek cyfrowych

Cezary Mazurek, Tomasz Parkoła, Marcin Werla

W pracy przedstawiono modele i organizację funkcjonowania biblioteki cyfrowej oraz przede wszystkim etapy tworzenia i rozwoju biblioteki cyfrowej od strony jej projektowania. Skupiono się na dwóch modelach tworzenia biblioteki cyfrowej: instytucjonalnym i konsorcjalnym. Na tym tle omówiono też najważniejsze obszary organizacji pracy i samego funkcjonowania biblioteki cyfrowej.

Wstęp

Pierwsze biblioteki cyfrowe pojawiły się w Polsce na początku XX wieku. Inicjatywą, która zapoczątkowała ich dynamiczny rozwój była zbudowana z wykorzystaniem oprogramowania dLibra (http://dlibra.psnc.pl/), Wielkopolska Biblioteka Cyfrowa (http://dlibra.psnc.pl/) [Górny 2003]. Powstała ona dzięki wspólnej inicjatywie Poznańskiej Fundacji Bibliotek Naukowych (PFBN) i Poznańskiego Centrum Superkomputerowo-Sieciowego (PCSS), przy współpracy wielkopolskich bibliotekarzy.

Oprogramowanie dLibra, rozwijane w PCSS już od 1996 r., początkowo przeznaczone było dla wydawnictw elektronicznych. Od 1999 r. rozpoczęto dostosowywanie jego funkcji do potrzeb instytucji kultury i nauki, mając na uwadze przede wszystkim udostępnianie zdigitalizowanych zbiorów dziedzictwa kulturowego. Celem tych działań było stworzenie narzędzia, które pozwoliłoby na zwiększenie możliwości dostępu do dziedzictwa kulturowego online. Analogiczne projekty rozwijające się w tym samym czasie za granicą były nastawione przede wszystkim na publikacje naukowe (systemy DSpace oraz ePrints) lub też były realizowane w starszych technologiach i z ograniczonym zakresem funkcjonalnym, ze względu na specyficzne docelowe grupy użytkowników (rozwój systemu Greenstone finansowany m.in. przez UNESCO z myślą o budowie bibliotek cyfrowych na potrzeby edukacji w krajach rozwijających się). Rozwój systemu dLibra prowadzony był od początku w ścisłej współpracy

z przedstawicielami środowiska bibliotek zainteresowanych digitalizacją zbiorów, co pozwoliło na dostosowanie tego oprogramowania do specyfiki funkcjonowania tych instytucji. Intensywny rozwój bibliotek cyfrowych w Polsce możliwy był dzięki realizacji od 2001 r. programu *PIONIER: Polski Internet Optyczny – Zaawansowane Aplikacje, Usługi i Technologie dla Społeczeństwa Informacyjnego* [Rychlewski 2000], w którym jednym z obszarów zaawansowanych usług sieciowych był sieciowy system bibliotek cyfrowych [Mazurek 2001].

Wykorzystanie opracowanego przez PCSS oprogramowania umożliwiło zaangażowanie do współuczestnictwa w WBC wielu instytucji kulturalnych i naukowych z regionu. W kolejnych latach powstawały kolejne biblioteki cyfrowe, które w zależności od charakteru instytucji je tworzących oraz ich możliwości finansowo-organizacyjnych, tworzone były zazwyczaj w modelu instytucjonalnym lub konsorcjalnym.

1. Modele bibliotek cyfrowych w Polsce

Biblioteki cyfrowe, budowane w modelu instytucjonalnym lub konsorcjalnym, różnią się przede wszystkim liczbą instytucji, które je współtworzą. Instytucjonalne biblioteki cyfrowe tworzone są przez jedną instytucję, natomiast konsorcjalne przez więcej, z jedną lub kilkoma jednostkami wiodącymi [Mazurek 2007]. Poza tym bardzo wymiernym rozróżnieniem, istnieje z reguły różnica związana z zasobami cyfrowymi, które są gromadzone i udostępniane w ramach dwóch wymienionych wyżej typów bibliotek cyfrowych. W przypadku instytucjonalnych bibliotek cyfrowych, zasoby te związane są przede wszystkim z charakterem i działalnością instytucji je tworzących czy zbiorami będącymi w posiadaniu tych instytucji. Przykładami tego typu bibliotek są:

- Biblioteka Cyfrowa Ośrodka Rozwoju Edukacji (http://www.bc.ore.edu.pl/dlibra) – niewielka biblioteka cyfrowa, działająca od 2007 r. i udostępniająca raporty, książki i materiały edukacyjne, szczególnie przydatne dydaktykom.
- e-biblioteka Uniwersytetu Warszawskiego (http://ebuw.uw.edu.pl/dlibra) – duża biblioteka cyfrowa, działająca od 2007 r. i udostępniająca online zarówno zabytkowe jak i współczesne zbiory Biblioteki Uniwersyteckiej w Warszawie, w tym również publikacje naukowe pracowników UW.
- Publiczne Archiwum Cyfrowe Agnieszki Osieckiej (http://www.archiwumagnieszkiosieckiej.pl/dlibra) – biblioteka cyfrowa prowadzona przez Fundację im. Agnieszki Osieckiej „Okularnicy" od 2009 r., zawierająca wiele różnego rodzaju materiałów związanych z życiem i twórczością Agnieszki Osieckiej.

Na tych przytoczonych przykładach widać, jak różne mogą być instytucjonalne biblioteki cyfrowe. Jest to oczywisty efekt specyfiki konkretnych instytucji, które chcą popularyzować informacje o swojej działalności czy zgromadzonych zbiorach. Zaletą takiego rozwiązania jest jego bezkompromisowość, tzn. reguły działania biblioteki cyfrowej, struktura podziału zasobów, tempo rozwoju czy charakter zasobów ustalane są wyłącznie pod kątem potrzeb jednej instytucji.

W przypadku bibliotek cyfrowych tworzonych przez instytucje zrzeszone w różnego rodzaju konsorcjach, charakter zasobów, ich organizacja oraz zasady tworzenia leżą już w gestii wszystkich współpracujących instytucji (choć często zasady działania ustalane są przez instytucję wiodącą). Przykładami bibliotek cyfrowych tworzonych przez konsorcja są:

- Wielkopolska Biblioteka Cyfrowa – regionalna biblioteka cyfrowa, uruchomiona w 2002 r. z inicjatywy poznańskiego środowiska akademickiego. Uczestnikami projektu są biblioteki naukowe i publiczne Wielkopolski. Zbiory udostępniane w ten sposób poprzez Internet przez te instytucje należą do czterech typów zasobów: edukacyjne, dziedzictwo kulturowe, regionalne i muzyczne. Techniczną opiekę nad WBC sprawuje PCSS.
- Repozytorium Cyfrowe Instytutów Naukowych (http://rcin.org.pl/dlibra) – za tę inicjatywę funkcjonującą od 2011 r. odpowiada ponadregionalne porozumienie 16 polskich instytutów naukowych oraz ich bibliotek, tworzących Konsorcjum Repozytorium Cyfrowego Instytutów Naukowych. Cechą wspólną członków tego typu konsorcjum jest naukowy charakter ich działalności, który przekłada się również na charakter zbiorów.
- Księgozbiór Wirtualny Federacji Bibliotek Kościelnych FIDES (http://digital.fides.org.pl/dlibra) – inicjatywa Federacji Bibliotek Kościelnych FIDES działająca od 2006 r.. Biblioteka cyfrowa gromadzi i prezentuje zbiory bibliotek rozproszonych po całej Polsce i zrzeszonych w Federacji a także utwory przekazane do publikacji bezpośrednio przez twórców i wydawców.
- Morska Biblioteka Cyfrowa (http://mbc.fundacjamorska.org/dlibra) –uruchomiona w 2009 r. przez Polską Fundację Morską tematyczna biblioteka cyfrowa, której celem jest udostępnianie cyfrowych zasobów piśmiennictwa polskiej kultury morskiej i żeglarskiej. Projekt ma charakter otwarty, przez co uczestnikami mogą być i są różne instytucje nie tylko z kraju, ale również z zagranicy. W ramach Morskiej Biblioteki Cyfrowej współpracują ze sobą organizacje z Ameryki Północnej, Ameryki Południowej oraz Europy.

Powyższe przykłady pokazują, jakie motywy mogą kierować instytucjami, chcącymi wspólnie budować biblioteki cyfrowe. W przypadku konsorcjów wszelkie reguły współpracy muszą być zaakceptowane przez wszystkich uczestników. Zazwyczaj jedna, wybrana instytucja wiodąca opracowuje podstawowe założenia oraz administruje biblioteką cyfrową, a instytucje współpracujące dostosowują swoje procedury do tych założeń i dostarczają treść do biblioteki. Sprawy najistotniejsze dyskutowane i ustalane są najczęściej na forum kluczowych partnerów konsorcjum.

Ostatnim z najważniejszych aspektów rozróżniającym model instytucjonalny od konsorcjalnego jest zaplecze sprzętowo-programowe biblioteki cyfrowej. W przypadku bibliotek cyfrowych instytucjonalnych cały ciężar utrzymania konfiguracji sprzętowo-programowej spada na jedną instytucję. Zalicza się do tego m.in. infrastruktura i personel techniczny, bieżąca obsługa czy usuwanie awarii. Z tego względu na profesjonalną instytucjonalną bibliotekę cyfrową mogą sobie pozwolić zazwyczaj tylko duże ośrodki, które mają odpowiednie zaplecze lub też mogą skorzystać z zaple-

cza istniejącego w nadrzędnej jednostce (np. biblioteki uczelniane). Inaczej wygląda to w przypadku biblioteki cyfrowej tworzonej przez konsorcjum. Tutaj najczęściej partnerzy partycypują w kosztach, a instytucja wiodąca koordynuje i odpowiada za działania w zakresie utrzymania i rozwoju infrastruktury sprzętowo-programowej biblioteki cyfrowej. W przypadku wspomnianych regionalnych bibliotek cyfrowych, częstą praktyką jest współpraca bibliotek z regionu z jednostką wiodącą naukowej sieci miejskiej (tzw. jednostki wiodące MAN) lub centrum obliczeniowym (tzw. centra KDM)[1]. Zapewnia to odpowiednią wartość merytoryczną zasobów oraz bardzo dobre zaplecze techniczne, które oferowane jest przez centrum MAN/KDM z danego regionu. Tego typu współpraca istnieje w Polsce, np. w Wielkopolsce (Wielkopolska Biblioteka Cyfrowa działająca przy wsparciu PCSS), na Podlasiu Podlaska Biblioteka Cyfrowa działająca przy wsparciu BiaMAN), czy na Dolnym Śląsku (Dolnośląska Biblioteka Cyfrowa działająca przy wsparciu WCSS).

Zarówno regionalne jak i instytucjonalne biblioteki cyfrowe działające w Polsce mogą na poziomie technicznym włączyć się w istniejącą sieć bibliotek cyfrowych. W ramach takiej sieci biblioteki cyfrowe m.in. automatycznie wymieniają informacje o udostępnianych zasobach i o planach digitalizacji. Jest to podstawą do udostępniania użytkownikom tych bibliotek nowych, zaawansowanych usług sieciowych. Portalem udostępniającym takie usługi w Polsce jest Federacja Bibliotek Cyfrowych sieci PIONIER (http://fbc.pionier.net.pl/) [Mazurek 2008].

2. Organizacja funkcjonowania biblioteki cyfrowej

Poza przedstawionymi w poprzednim rozdziale modelami, wpływ na docelową postać biblioteki cyfrowej mają również rozwiązania przyjęte w zakresie organizacji funkcjonowania biblioteki cyfrowej. Przy tworzeniu biblioteki cyfrowej konieczne jest uwzględnienie szeregu obszarów jej funkcjonowania. Do najistotniejszych należą:

- zasady i proces doboru zbiorów do digitalizacji i udostępniania online,
- struktura udostępniania online,
- schemat opisu obiektów cyfrowych,
- digitalizacja,
- przygotowanie finalnej postaci obiektu,
- archiwizacja i długoterminowe przechowywanie,
- udostępnianie w Internecie.

Zasady i proces doboru zbiorów do digitalizacji i udostępniania online obejmują wiele kryteriów takich jak np. zakres tematyczny poszczególnych kolekcji biblioteki cyfrowej, fizyczny stan zbiorów, potrzeby użytkowników oraz odpowiedni status zbiorów, np. w kontekście praw autorskich.

Struktura udostępniania online jest niezwykle ważna z punktu widzenia użytkowników końcowych i decydujące o łatwości przeglądania zawartości biblioteki cyfro-

[1] Obecnie w Polsce funkcjonuje 21 jednostek MAN oraz 5 centrów KDM. Są one zrzeszone w konsorcjum PIONIER, a ich lista dostępna jest na stronie http://www.pionier.net.pl/online/pl/jednostki/.

wej. Warto poświęcić tej kwestii odpowiednio dużo uwagi, aby w efekcie podział zasobów udostępnianych przez strony internetowe był czytelny i łatwy w odbiorze dla użytkownika. W Polsce mamy z reguły do czynienia z tematycznym podziałem zawartości bibliotek cyfrowych. Obiekty cyfrowe przypisane są do tzw. kolekcji, które zorganizowane są najczęściej w hierarchie. Oznacza to, że obiekty, biblioteki cyfrowej przypisane do określonej kolekcji tematycznej mogą być podzielone dodatkowo na kolekcje podrzędne, zawężające ich nadrzędną kolekcję tematyczną. Podział ten może być realizowany wg jednego kryterium (np. kolekcje tematyczne w Wielkopolskiej Bibliotece Cyfrowej) lub też wg kilku równolegle stosowanych kryteriów (np. Śląska Biblioteka Cyfrowa – http://www.sbc.org.pl/dlibra proponuje podział wg subregionów Śląska, podział tematyczny analogiczny do WBC, oraz podział wg instytucji dostarczających obiekty). Jest to możliwe dzięki temu, że obiekty mogą należeć równocześnie do wielu kolekcji z równoległych hierarchii.

Schemat opisu obiektów cyfrowych (tzw. schemat metadanych) powinien być dopasowany do charakteru konkretnej biblioteki cyfrowej. Warto w tym kontekście rozważyć:

- typy obiektów cyfrowych, które będą udostępniane (i opisywane) – w tym również zróżnicowanie tych typów (np. pocztówki, starodruki, mapy, nuty, podręczniki akademickie),
- stopień zaawansowania użytkowników biblioteki cyfrowej – w tym również ich potrzeby eksplorowania zawartości biblioteki cyfrowej,
- organizacja pracy redaktorów biblioteki cyfrowej – w tym również czas jaki może być poświęcony na opracowanie poszczególnych obiektów oraz możliwości importu metadanych z zewnętrznych systemów.

Analiza powyższych elementów powinna dostarczyć informacji o tym, jakiego rodzaju schemat metadanych będzie w danej bibliotece cyfrowej optymalny. Ogólnie znanym schematem metadanych, który wykorzystywany jest na szeroką skalę na świecie jest schemat Dublin Core oraz jego rozszerzenie [DCMI Metadata Terms 2010]. Minusem tego schematu jest fakt, że kreuje on dość ogólne ramy opisu, przez co interpretacja poszczególnych pól może się znacznie różnić między instytucjami [Potęga 2009]. To z kolei tworzy ryzyko niejednoznaczności w opisach obiektów cyfrowych, skutkujące ograniczonymi możliwościami ich automatycznego przetwarzania (np. w kontekście agregacji metadanych) i utrudnieniami dla czytelników, którzy równocześnie korzystają z kilku bibliotek cyfrowych. Fakt, że schemat Dublin Core jest schematem uniwersalnym, a jednocześnie z powodzeniem wspomaga wyszukanie obiektów biblioteki cyfrowej w Internecie spowodował, że jest on wykorzystywany również w Polsce, przy czym często modyfikowany jest według potrzeb konkretnych bibliotek cyfrowych. Modyfikacja polega z reguły na dodaniu nowych pól, których z punktu widzenia danej instytucji lub konsorcjum brakuje. Pomocny w tej sytuacji jest ujednolicony schemat metadanych, który może z jednej strony uformować dostatecznie precyzyjne ramy opisu dla analizy automatycznej, a z drugiej strony jest na tyle uniwersalny i rozbudowany, że pozwala na opis różnych typów dokumentów i zarazem jest zrozumiały dla użytkownika. W 2011 r. Poznańskie Centrum Superkom-

puterowo-Sieciowe opracowało taki schemat i opublikowało go pod nazwą PLMET (Dokumentacja schematu metadanych PLMET, 2011) i docelowo ma on zastąpić schemat Dublin Core stosowany obecnie w Federacji Bibliotek Cyfrowych.

Digitalizacja to element kluczowy w procesie tworzenia każdej biblioteki cyfrowej udostępniającej obiekty dziedzictwa kulturowego. Podstawowymi elementami do analizy w tym kontekście są: jakość wynikowych plików digitalizacji (tzw. „pliki master" lub „pliki matki") oraz zakres ich późniejszego przetwarzania (np. OCR, obróbka graficzna). Od podjętych decyzji w tym zakresie zależy rozmiar potrzebnego archiwum danych oraz nakład pracy potrzebny do przygotowania wersji dokumentów udostępnianych online (tzw. wersje prezentacyjne, por. punkt 4.4.6).

Przygotowanie finalnej postaci obiektu, czyli opracowanie wynikowych plików skanowania wiąże się przede wszystkim z przygotowaniem ich wersji prezentacyjnych oraz przygotowaniem opisu obiektu (metadanych). Wersje prezentacyjne mogą być przygotowane w różnych formatach (np. DjVu, PDF, HTML, JPG) w zależności od potrzeb konkretnej instytucji. Poszczególne formaty mają swoje wady oraz zalety, dlatego warto przeanalizować ich właściwości i dokonać odpowiedniego wyboru w kontekście różnych typów dokumentów. Opis obiektu przygotowywany jest przez specjalizowane kadry biblioteki cyfrowej, często na podstawie danych importowanych z innych systemów informatycznych, takich jak systemy katalogowe. Obszerność opisu, czy jego wielojęzyczność (równoległe opracowanie w języku innym niż polski) należy dostosować do możliwości konkretnej instytucji. Jednym z elementów tego etapu pracy może być również wzbogacenie metadanych o miniaturę obiektu cyfrowego, która może być wykorzystana do uatrakcyjnienia wyglądu stron WWW biblioteki cyfrowej.

Archiwizacja i długoterminowe przechowywanie są niezmiernie ważne w pracach biblioteki cyfrowej. Z biegiem czasu sprzęt, oprogramowanie czy formaty plików ulegają zmianie, dlatego istotne jest, aby, pliki cyfrowe powstałe w wyniku digitalizacji (często bardzo kosztownej) były dostępne i możliwe do odczytania w wieloletniej perspektywie (tj. przy użyciu nowych konfiguracji sprzętowo-programowych jakie pojawią się w przyszłości). Choć problematyka długoterminowego przechowywania jest wciąż nowym zagadnieniem, należy ją mieć na uwadze i formować prace przy bibliotece cyfrowej w taki sposób, aby obejmować również ten aspekt. Warto sobie uświadomić, że z reguły podstawowym celem prac digitalizacyjnych jest archiwizacja i długoterminowe przechowywanie obiektów, a dopiero w drugiej kolejności znajduje się tworzenie ich wersji prezentacyjnych i udostępnianie użytkownikom końcowym. W tym kontekście warto nadmienić, iż w pewnych sytuacjach zachodzi potrzeba dostępu do obiektów cyfrowych wysokiej jakości – tzw. plików matek (ang. *Master Copy*). Może to być związane, np. z prowadzeniem badań naukowych w kontekście humanistyki czy lingwistyki.

Udostępnianie w Internecie odbywa się z reguły poprzez tzw. wersje prezentacyjne obiektów. Rozróżnienie na wersje prezentacyjne i archiwalne wynika ze względów praktycznych – rozmiar plików matek jest na tyle duży, że korzystanie z nich na dużą skalę za pośrednictwem Internetu jest kosztowne i niewygodne zarówno po stronie

czytelnika jak i biblioteki cyfrowej. W tym obszarze istotne są ponadto takie aspekty jak funkcjonalność, ergonomia i architektura informacji dostępnych na stronie WWW biblioteki cyfrowej, dobranie możliwości i wymagań formatu udostępniania dla czytelników, czy jasne określenie licencji udostępniania i praw dalszego wykorzystania obiektu cyfrowego [Fundacja Europeana 2010].

3. Etapy tworzenia i rozwoju biblioteki cyfrowej – podejście projektowe

Mając pogląd na obszary związane z organizacją funkcjonowania biblioteki cyfrowej możliwe jest przejście do poszczególnych etapów jej tworzenia. Zasadniczymi krokami przy tworzeniu i rozwoju biblioteki cyfrowej są:

- definicja docelowej wizji biblioteki cyfrowej,
- decyzja co do charakteru organizacyjnego biblioteki cyfrowej – własna inicjatywa, bądź uczestnictwo w konsorcjum czy powołanie konsorcjum,
- przygotowanie planu prac, uruchomienie i monitorowanie biblioteki cyfrowej.

Kroki te omówiono w kolejnych podpunktach niniejszego rozdziału.

3.1. Wizja biblioteki cyfrowej

Pierwszym etapem przy realizacji dowolnego projektu, w tym projektu tworzenia biblioteki cyfrowej, jest określenie jego wizji. Należy przeanalizować powody, dla których biblioteka cyfrowa ma powstać i określić, jaka postać biblioteki cyfrowej jest pożądaną postacią końcową. Istotne jest, aby decyzja została podjęta z uwzględnieniem długiej perspektywy czasu, w której to biblioteka cyfrowa ma funkcjonować. Częstym rozwiązaniem jest aplikowanie o finansowanie projektu, np. z budżetu Państwa, czy programu UE. Jest to na pewno dobry sposób na zainicjowanie biblioteki cyfrowej, natomiast nie powinien on być jedynym powodem jej uruchomienia, chociażby dlatego że biblioteka cyfrowa powinna działać również po zakończeniu takiego projektu. Stąd też, instytucja decydująca się na utworzenie biblioteki cyfrowej musi mieć jasną wizję co do modelu na jej dalsze utrzymanie i rozwój (w tym modelu finansowo-organizacyjnego).

Istotnym aspektem wizji biblioteki cyfrowej jest świadomość, że biblioteka cyfrowa to nie tylko zasoby dostępne online, ale również duże ilości danych składowane w archiwum cyfrowym. Wpływa to istotnie na koszty i nakład pracy związane z uruchomieniem i utrzymaniem biblioteki cyfrowej.

W ramach określania wizji biblioteki cyfrowej należy również zwrócić uwagę na identyfikację docelowych użytkowników, którzy będą korzystać z udostępnianych zasobów. W związku z tym, iż biblioteka cyfrowa sama w sobie dedykowana jest przede wszystkim tzw. użytkownikom końcowym, czyli czytelnikom, musi odzwierciedlać ich preferencje zarówno jeżeli chodzi o udostępniane materiały jak i oferowane usługi. Niekiedy warto przeanalizować profil użytkowników danej biblioteki (np. poprzez

przegląd statystyk wypożyczeń, czy przeprowadzenie ankiety), aby trafniej odpowiedzieć na ich potrzeby podczas określania zakresu planowanych do udostępnienia materiałów cyfrowych.

W samej wizji biblioteki cyfrowej nie może również zabraknąć specyfikacji celów i oczekiwanych wyników – czy ma to być określona liczba obiektów udostępnionych użytkownikom, liczba wyświetleń tych obiektów, większa rozpoznawalność samej biblioteki, większe zainteresowanie materiałami w bibliotece tradycyjnej (np. zwiększenie liczby wypożyczeni etc.)? Bardzo dobrze, jeżeli ta specyfikacja celów i wyników może zostać zdefiniowana w sposób precyzyjny i łatwy do weryfikacji – np. poprzez podanie numerycznych wskaźników.

3.2. Model organizacyjny biblioteki cyfrowej

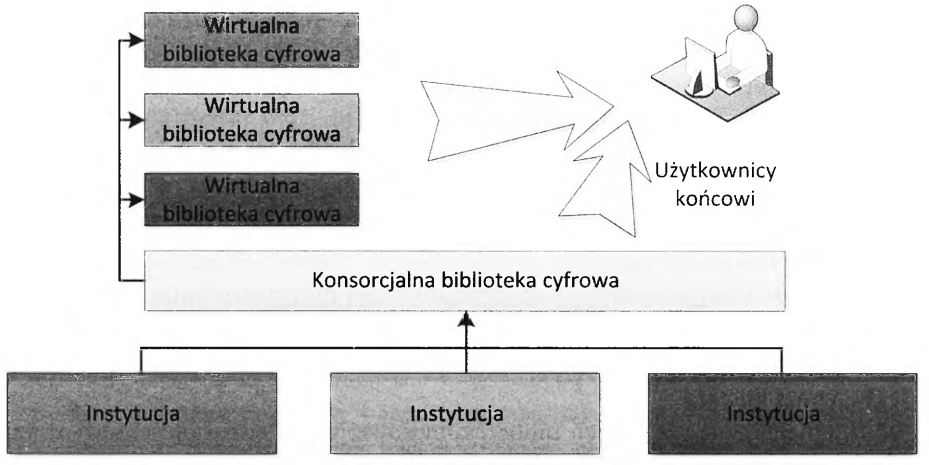
Po sprecyzowaniu wizji biblioteki cyfrowej, jej charakteru i celów oraz docelowych użytkowników następuje etap decyzji w odniesieniu do sposobu tworzenia biblioteki cyfrowej czy też modelu organizacyjnego. Chodzi mianowicie o decyzję o utworzeniu własnej, nowej biblioteki cyfrowej (instytucjonalnej lub konsorcjalnej), bądź też o przyłączeniu się do istniejącego konsorcjum.

Samodzielne uruchomienie biblioteki cyfrowej wiąże się z większym wysiłkiem instytucji, co przekłada się przede wszystkim na zwiększone nakłady finansowe oraz konieczność posiadania wyspecjalizowanego zaplecza merytoryczno-technicznego. Przy samodzielnym budowaniu biblioteki cyfrowej, elementy które będą wymagały doprecyzowania to m.in.: schemat metadanych i reguły jego stosowania w kontekście różnych typów obiektów, szkolenia personelu, rozwój/modyfikacja funkcji oprogramowania biblioteki cyfrowej, kontrola jakości, czy polityka udostępniania zasobów. Wybór ten może dawać z drugiej strony większy prestiż, poczucie indywidualności oraz większą elastyczność jeśli chodzi o zmiany organizacji zasobów, formy ich udostępniania, funkcjonalność portalu biblioteki cyfrowej itp. W przypadku projektu o dużej skali działania, może okazać się zasadne powołanie konsorcjum instytucji zainteresowanych utworzeniem nowej biblioteki cyfrowej.

Przyłączenie się do istniejącej już biblioteki cyfrowej jest mniej wymagające i wiąże się przede wszystkim z realizacją prac digitalizacyjnych oraz stosowaniem reguł wypracowanych przez konsorcjum, do którego nastąpi przyłączenie. W porównaniu z tworzeniem nowej biblioteki cyfrowej, odciążenie następuje tu częściowo w kwestiach specjalizowanych kadr (administracja, monitorowanie, opieka techniczna), rozwoju oprogramowania biblioteki cyfrowej oraz definicji schematu metadanych i reguł jego stosowania. Nie oznacza to definitywnego braku wpływu na kształt biblioteki cyfrowej, czy schemat metadanych, ale praca wymagana na wejściu jest znacznie mniejsza, a wszelkie zmiany następują z reguły w formie wypracowania decyzji optymalnych dla wszystkich członków konsorcjum.

W praktyce wybór jest często podyktowany możliwościami finansowymi i zapleczem technicznym instytucji. W przypadku małych instytucji kultury częściej nastąpi przyłączenie do istniejącego konsorcjum (np. regionalnej biblioteki cyfrowej).

W przypadku instytucji dużych będzie to albo własna inicjatywa (instytucjonalna biblioteka cyfrowa) lub rola koordynatora w konsorcjum. Warto również mieć świadomość, że współczesne oprogramowanie do budowy bibliotek cyfrowych pozwala na tworzenie czy też wydzielanie wirtualnych bibliotek cyfrowych funkcjonujących na podstawie bibliotek konsorcjalnych (por. Rysunek 1). Takie podejście umożliwia osiągnięcie kompromisu cechującego się zaletami udziału w konsorcjum (choćby finansowymi) oraz osiągnięcie stopnia niezależności zbliżonego do instytucjonalnej biblioteki cyfrowej.



Rys. 1. Wirtualne biblioteki cyfrowe tworzone na podstawie biblioteki konsorcjalnej.

3.3. Utworzenie i uruchomienie biblioteki cyfrowej jako projekt

Po dokonaniu wyboru co do formy organizacyjnej dla biblioteki cyfrowej, następuje etap najważniejszy i najbardziej intensywny – przygotowanie planu prac, uruchomienie i monitorowanie biblioteki cyfrowej. Etap ten warto rozdzielić na podetapy, które określają konkretne obszary działania:

- przygotowanie planu prac oraz określenie sposobu monitorowania i oceny postępów prac a także zarządzania ryzykiem przy tworzeniu biblioteki cyfrowej,
- definicja ról w zespole,
- opracowanie planu przepływu informacji i przepływu pracy w bibliotece cyfrowej,
- wybór oprogramowania dla biblioteki cyfrowej,
- promocja biblioteki cyfrowej,
- monitorowanie wykorzystania zasobów biblioteki cyfrowej.

Zasadniczym elementem tworzenia biblioteki cyfrowej jest przygotowanie samego planu jej utworzenia, oraz monitorowania postępów i zarządzania ryzykiem. W zależności od wielkości projektowanej biblioteki cyfrowej, możliwe jest skorzystanie

z wielu różnych podejść do zarządzania projektami. W przypadku bardzo dużych projektów niezbędne może okazać się posiłkowanie się formalnymi metodami takimi jak na przykład PRINCE2 (Projects In a Controlled Environment). Jest to sformalizowana metoda zarządzania projektem, ogólnego przeznaczenia, wykorzystywana do bardzo dużych projektów, realizowanych nawet na poziomie krajowym lub międzynarodowym. Niezależnie od wybranej metody, warto też zapoznać się z kompendiami wiedzy na temat zarządzania projektami. Tutaj przykładem może być PMBOK Guide (*A Guide to the Project Management Body of Knowledge*) – zestaw zaleceń ogólnego przeznaczenia do stosowania przy planowaniu i zarządzaniu projektem. Są też bardzo przydatne narzędzia menedżera projektu, takie jak na przykład Struktura Podziału Prac (ang. *Work Breakdown Structure*, WBS), która pozwala na przygotowanie hierarchicznej struktury zadań w projekcie, oraz dowolne doprecyzowywanie zakresu działań w odniesieniu do poszczególnych zadań[2].

W bieżącym zarządzaniu projektem kluczową rolę odgrywa też zarządzanie ryzykiem, a brak zdefiniowanych zasad zarządzania ryzykiem jest jedną z ważniejszych przyczyn niepowodzeń w realizacji projektów. Warto pamiętać o tym przy planowaniu zadań w ramach projektu tworzenia biblioteki cyfrowej. Niektórymi z potencjalnych zagrożeń związanych z takimi pracami mogą być: awarie sprzętu, problemy natury prawnej, brak wystarczającego czasu na poprawki w przypadku błędów, niedostateczna kontrola jakości, niepowodzenie planu na utrzymanie biblioteki cyfrowej po jej wstępnej inicjacji.

Ważnym elementem realizacji biblioteki cyfrowej w podejściu projektowym jest monitorowanie postępu prac oraz ich ocena. W tym celu należy przygotować zestaw procesów i zadań, jakie stać będą przed zespołem realizującym projekt oraz określić harmonogram ich realizacji. Należy również jasno zidentyfikować role, które będą przydzielone do odpowiedzialnych za konkretne zadania osób. Dzięki temu monitorowanie postępu prac będzie możliwe, a poszczególne osoby przypisane do ról w projekcie, będą miały jasno wyznaczone zadania oraz cele do realizacji. Na tej podstawie, w zestawieniu z harmonogramem prac, możliwe będzie monitorowanie postępu oraz ocena jakości bieżących działań. Wnioski płynące z monitorowania mogą prowadzić do modyfikacji pierwotnie opracowanego procesu tworzenia biblioteki cyfrowej. Należy mieć świadomość, że w przypadku projektowego charakteru pracy, szczególnie w przypadku projektów obarczonych dużym ryzykiem, znacznie ważniejsze od trzymania się z góry określonego planu jest bieżące monitorowanie prac i adaptowanie planu do zaistniałych sytuacji.

3.4. Przepływ pracy a role pracowników w zespole biblioteki cyfrowej

Kluczowym elementem w realizacji projektu utworzenia biblioteki cyfrowej jest personel. W ramach prac przygotowawczych niezbędna jest identyfikacja ról, jakie przybiorą osoby pracujące na bieżąco przy bibliotece cyfrowej. Warto podkreślić, że

[2] Zestawienie narzędzi informatycznych, które mogą być pomocne dla menadżera projektu dostępne jest m.in. na stronie http://en.wikipedia.org/wiki/Comparison_of_project_management_software

z jednej strony jedna rola może być wykonywana przez kilka osób, z drugiej strony jedna osoba może mieć kilka ról w projekcie. To która tendencja będzie przeważająca w konkretnym przypadku, zależy głownie od skali projektu i liczby zaangażowanych w nim osób.

Role jakie mogą przyjmować członkowie zespołu biblioteki cyfrowej są ściśle powiązane z docelowym bieżącym przepływem pracy w procesie tworzenia obiektów cyfrowych. Zasadniczymi etapami tego procesu są:

- selekcja dokumentów, które będą poddane digitalizacji,
- opracowanie dokumentu (przygotowanie metadanych),
- przygotowanie dokumentu do digitializacji,
- digitalizacja dokumentu,
- obróbka wynikowych plików digitalizacji,
- przygotowanie wersji prezentacyjnej obiektu cyfrowego,
- weryfikacja wykonanych prac,
- upublicznienie wersji prezentacyjnej,
- archiwizacja „plików matek" dokumentu cyfrowego.

Etapy te omówiono dokładniej w następnych podpunktach.

3.4.1. Selekcja dokumentów do digitalizacji

Selekcja dokumentów do digitalizacji opiera się z reguły na kilku kryteriach, z których najważniejszymi są:

- Fizyczny stanu oryginału – jeśli dany materiał oryginalny może wkrótce ulec nieodwracalnej degradacji, być może warto przeprowadzić jego digitalizację w pierwszej kolejności. Decyzja powinna zostać oczywiście podjęta z uwzględnieniem wartości intelektualnej, historycznej, czy kulturowej danej pozycji.
- Potrzeby użytkowników – ukierunkowanie się na oczekiwania docelowej grupy odbiorców biblioteki cyfrowej. Należy jednak tutaj pamiętać o tym, że w ogólnym przypadku selekcja nie powinna dotyczyć tylko tych pozycji, które należy upublicznić ze względu na ich popularność, ale również tych, które mają być zabezpieczone przed ich fizyczną degradacją.
- Stan praw autorskich – uzależnienie digitalizacji od możliwości szerokiego udostępnienia zbiorów w kontekście praw autorskich. Częstą praktyką jest digitalizacja przede wszystkim tych zbiorów, które należą do tzw. domeny publicznej. Warto pamiętać, aby racjonalnie podchodzić do kwestii digitalizacji dokumentów będących w domenie publicznej. Niekiedy może okazać się, iż korzystniej będzie wykonać digitalizację współczesnych pozycji, do których będzie ograniczony dostęp, ale które mimo to będą cieszyły się dużą popularnością (np. skrypty akademickie dostępne wyłącznie w sieci uczelnianej), niż szereg dokumentów z domeny publicznej, które będą wykorzystywane sporadycznie.

Warto podkreślić, że w celu odpowiedniej selekcji dokumentów do digitalizacji, należy wyważyć priorytety, które nakreślone są przez konkretną instytucję, w odniesieniu do poszczególnych kryteriów. Choć z perspektywy zachowania wszystkich do-

kumentów, może okazać się, że fizyczny stan obiektu jest najistotniejszy, to w praktyce digitalizacja wszystkich zagrożonych degradacją dokumentów może być niemożliwa, choćby ze względów finansowych. W takich sytuacjach należy zastosować dodatkowe kryteria (np. wartość intelektualna, wartość historyczna), które pozwolą wybrać pozycje przeznaczone do digitalizacji w pierwszej kolejności.

3.4.2. Opracowanie

Opracowanie dokumentu wiąże się zazwyczaj z przygotowaniem zestawu metadanych. Tutaj w odniesieniu do konkretnej pozycji, należy rozważyć schemat oraz języki, w których opis będzie przygotowany. Jeśli dokument został już wcześniej opisany, należy skorzystać z takiego opisu uzupełniając ewentualne braki. Warto podkreślić, że przy opracowywaniu opisu dokumentu należy mieć na uwadze dwa przypadki wykorzystania opisu. Pierwszy to opis dedykowany do archiwizacji, drugi to opis dedykowany dla użytkowników biblioteki cyfrowej. W przypadku pierwszym opis powinien być możliwe najpełniejszy, przez co może zawierać informacje istotne z punktu widzenia profesjonalistów bibliotekarzy. W drugim przypadku należy ukierunkować się na bardziej ogólne informacje i być może również użyć sformułowań bardziej zbliżonych do języka wyszukiwawczego stosowanego przez użytkowników portalu biblioteki cyfrowej, tak by użytkownicy ci mogli w łatwy sposób skorzystać z udostępnianych informacji i je zrozumieć. W pewnych sytuacja, w celu ujednolicenia sposobu opisu obiektów cyfrowych, warto zastosować słowniki kontrolowane dla określonych pól. Tzn. takie, które wymuszają wykorzystanie jednej ze zdefiniowanych wcześniej wartości dla danego pola. Przykładem może być pole typ zasobu, które w polskich bibliotekach cyfrowych przyjmuje z reguły wiele postaci do określenia jednego typu, np. czasopismo, czasopisma, Czasopismo. Przy zastosowaniu słownika kontrolowanego w prosty sposób uzyskujemy spójność na poziomie opisu różnych obiektów cyfrowych.

3.4.3. Przygotowanie do digitalizacji

Przed samą digitalizacją dokumentu konieczne jest odpowiednie przygotowanie pozycji. Może sprowadzać się to tylko do odnalezienia jej w magazynie, ale może również znaczyć renowację/naprawy w związku z powstałymi uszkodzeniami. Połączenie prac digitalizacyjnych i renowacji/napraw jest sensowne, choćby ze względu na minimalizację pracy związanej z wydawaniem pozycji z magazynu (może to być szczególnie istotne w przypadku wartościowych pozycji). Decyzja co do przekazania pozycji do naprawy/renowacji powinna być dyktowana dostępnymi możliwościami finansowymi oraz poziomem uszkodzeń konkretnego dokumentu. Na tym etapie istotne jest monitorowanie lokalizacji oryginału, w celu minimalizacji ryzyka zaginięcia lub uszkodzenia oryginału. Ma to szczególne znaczenie w przypadku digitalizacji przez podwykonawcę.

3.4.4. Digitalizacja

Digitalizacja dokumentu jest etapem, w którym z postaci fizycznej powstaje cyfrowa wersja dokumentu. Parametry digitalizacji, takie jak rozdzielczość, głębia kolorów, czy format zapisu plików muszą być ustalone z uwzględnieniem charakteru doku-

mentu. Inne będą parametry dla dokumentów tekstowych, inne dla dźwiękowych, jeszcze inne dla audiowizualnych. W ramach poszczególnych typów dokumentów także będą różnice – inaczej będzie digitalizowany zdobiony rękopis, inaczej współcześnie drukowana książka. Zestawy zaleceń i dobrych praktyk w odniesieniu do parametrów digitalizacji, można odnaleźć na stronach wiodących światowych instytucji, np. JISC Digital Media, Biblioteka Kongresu.

W związku z tym, że sprzęt do digitalizacji dokumentów jest stosunkowo kosztowny, a dodatkowo różne typy dokumentów powinny być digitalizowane przy wykorzystaniu różnych maszyn digitalizujących, dobrym rozwiązaniem może okazać się zlecenie digitalizacji podwykonawcy (np. firmie komercyjnej). Dzięki temu nie jest konieczny zakup i utrzymanie kosztownego sprzętu, szczególnie jeśli dokumenty, które mają być digitalizowane są różnorodne (np. rękopisy, starodruki, nagrania audio, nagrania wideo). Jeśli dokumenty do digitalizacji mają podobny charakter i typ, to zakup specjalizowanego sprzętu do digitalizacji będzie prawdopodobnie korzystniejszy finansowo. Również jeżeli digitalizacja jest planowana w perspektywie długoterminowej, to korzystne może być zbudowanie własnej pracowni digitalizacji.

3.4.5. Obróbka wynikowych plików digitalizacji

Efektem digitalizacji są tzw. pliki matki (pliki master), które w pewnych przypadkach wymagają dodatkowych zabiegów korekcyjnych. Chodzi o przygotowanie plików w taki sposób, aby były wygodne w przeglądaniu/czytaniu. Przykładem potrzeby korekty może być zbyt szeroko ustawiony obszar skanowania, przez co każdy skan zawiera margines z widoczną częścią skanera. W takim przypadku korekta sprowadza się do automatycznego lub półautomatycznego usunięcia zbędnych obszarów. Korekta może dotyczyć również poprawienia kontrastu, czy odkształceń powstałych przy skanowaniu. W przypadku dokumentów audio może to być np. eliminacja szumów. Bardzo ważne jest to, aby nie usuwać/modyfikować plików matek podczas korekty. To ze względu na archiwizację, gdyż archiwizacji poddawane są przede wszystkim pliki matki. Warto podkreślić, że krok w którym korygowane są pliki matki, jest krokiem opcjonalnym. Gdy wynikowe pliki skanowania nadają się bezpośrednio do upublicznienia w postaci prezentacyjnej, nie jest konieczna ich korekta. Poza korektą w przypadku współczesnych dokumentów drukowanych możliwe jest wykonanie tzw. OCR (Optical Character Recognition). Dzięki temu uzyskiwana jest warstwa tekstowa dokumentu, która ułatwia użytkownikom wyszukiwanie informacji w obiekcie cyfrowym oraz ułatwia wyszukiwarkom internetowym eksponowanie informacji o nim.

3.4.6. Przygotowanie wersji prezentacyjnej obiektu cyfrowego

Bazując na efektach prac korekty plików matek (lub na oryginalnych plikach matkach gdy korekta nie była potrzebna), należy przygotować wersję prezentacyjną obiektu cyfrowego. Zazwyczaj polega to na zmniejszeniu rozmiaru plików kosztem ich jakości. Dostępne algorytmy kompresji, pozwalają na uzyskanie małych rozmiarów plików, przy zachowaniu stosunkowo dobrej jakości danych, choć pewna część danych w wyniku takiej operacji jest zawsze tracona (jest to tzw. kompresja strat-

na). Przeważająca większość tekstowych dokumentów cyfrowych w Polsce ma postać prezentacyjną w formacie DjVu. Format DjVu charakteryzuje się wysoką kompresją danych, przez co pozwala uzyskać bardzo małe rozmiary plików prezentacyjnych. Dodatkowo w celu wyświetlenia danej strony książki, komputer użytkownika musi pobrać dane dotyczące tylko tej strony, a nie całego dokumentu. Minusem tego formatu jest mała popularność i konieczność instalacji dodatkowego oprogramowania (wtyczki) wyświetlającego pliki DjVu oraz brak wsparcia dla tego formatu ze strony wyszukiwarki Google. Innym formatem wykorzystywanym do prezentacji dokumentów tekstowych jest PDF, który może być szczególnie atrakcyjny do prezentowania dokumentów posiadających warstwę tekstową. Jest to wręcz format zalecany np. dla publikowanych online artykułów naukowych. Dzięki funkcje pobierania progresywnego plików w formacie PDF możliwe jest też szybsze rozpoczęcie przeglądania publikacji w przypadku dużych obiektów. Wśród innych formatów prezentacyjnych znajdują się HTML, JPG, MP3 (audio), czy FLV (wideo). Warto podkreślić, że mając pliki matki, lub pliki powstałe w efekcie korekty plików matek, możemy w stosunkowo prosty sposób utworzyć wersję prezentacyjną. Bardzo często może się to odbyć całkowicie automatycznie. Dzięki temu, potencjalne zmiany formatu plików prezentacyjnych są znacznie ułatwione.

3.4.7. Weryfikacja wykonanych prac i upublicznienie wersji prezentacyjnej

Przed udostępnieniem efektów digitalizacji szerszemu gronu odbiorców warto przeprowadzić weryfikacje plików. Sprowadza się to najczęściej w praktyce do przeglądu wersji prezentacyjnej obiektu cyfrowego i weryfikacji plików matek. Jeśli wersja prezentacyjna wygląda poprawnie i nie ma uchybień, to materiały źródłowe (pliki matki, pliki po korekcie) powinny też być poprawnie przygotowane. Jeżeli nad całością prac sprawuje opiekę dedykowane oprogramowanie, to może ono weryfikować poprawność plików do pewnego stopnia automatycznie.

Udostępnienie obiektu cyfrowego następuje najczęściej po pozytywnej weryfikacji. Obiekt w postaci prezentacyjnej wraz z opisem umieszczany jest na stronach biblioteki cyfrowej, przez co automatycznie dostępny jest dla docelowej grupy użytkowników.

3.4.8. Archiwizacja „plików matek" obiektu cyfrowego

Ostatnim elementem przepływu pracy w ramach procesu digitalizacji jest archiwizacja plików matek oraz opcjonalnie plików wynikowych z obróbki plików matek. Celem archiwizacji jest przechowywanie obiektu w postaci cyfrowej, tak aby utrwalić jego wizerunek i uchronić przed całkowitą jego utratą, w razie degradacji nośnika oryginalnego (np. papieru). W kontekście archiwizacji nie należy zapominać o aspekcie długoterminowego przechowywania. Z biegiem czasu konfiguracje sprzętowe i programowe oraz formaty plików wychodzą z użycia, zużywają się również elektroniczne nośniki na których zapisano dane. Z tych względów należy regularnie weryfikować dostępność i możliwość odczytu oraz wyświetlenia konkretnych pozycji. Dlatego też

niezbędne staną się operacje takie jak migracja formatu pliku, migracja sprzętowa, itp. Opcja przechowywania w archiwum plików będących efektem obróbki plików master jest atrakcyjna ponieważ na ich podstawie możliwe jest automatyczne wygenerowanie wersji prezentacyjnej pliku. Dzięki temu z biegiem czasu można w łatwy sposób zmieniać formaty plików prezentacyjnych danego obiektu.

3.4.9. Role w zespole biblioteki cyfrowej

Mając pogląd na cały przepływ pracy związany z digitalizacją obiektów cyfrowych, możliwa jest identyfikacja ról dla poszczególnych etapów tego przepływu. Role biorące bezpośredni udział w procesie digitalizacji obiektu cyfrowego to:

- Redaktor – wykonuje selekcję, opracowuje dokument, przekazuje dokument do digitalizacji, a po zakończeniu procesu weryfikacji wprowadza wersję prezentacyjną obiektu cyfrowego do biblioteki cyfrowej.
- Skanujący – osoba odpowiedzialna za digitalizację.
- Korektor – osoba odpowiedzialna za obróbkę graficzną plików matek.
- Konwertujący – osoba odpowiedzialna za konwersję plików matek lub plików wynikowych obróbki do wersji prezentacyjnej.
- Weryfikujący – osoba przeprowadzająca weryfikację plików wynikowych procesu digitalizacji, obróbki oraz plików prezentacyjnych.
- Archiwizujący – osoba odpowiedzialna za archiwizację plików matek oraz opcjonalnie plików po obróbce.

Wyżej wymienione role mogą być ze sobą łączone, np. role skanujący, korektor oraz konwertujący mogą być scalone w jedną rolę. Niektóre role (np. konwertującego) mogą wziąć na siebie odpowiednio przygotowane automatyczne narzędzia. Inne role mogą być dodatkowo podzielone – np. z roli Redaktor można wyłączyć rolę osoby wykonującej selekcję i tej, która opracowuje dokument i przekazuje go do digitalizacji. Ważne jest, aby zestaw ról był adekwatny do realnych warunków w danej instytucji i odpowiadał faktycznym działaniom poszczególnych osób do nich przypisanych.

3.5. Wybór oprogramowania dla biblioteki cyfrowej

Przygotowany przepływ pracy oraz zdefiniowane w nim role przypisane do poszczególnych osób definiują podstawowy kręgosłup organizacji biblioteki cyfrowej. Na tej podstawie możliwe jest zdefiniowanie wymagań funkcjonalnych oprogramowania, które służyć będzie do wspomagania prac digitalizacyjnych oraz archiwizowania i udostępniania efektów tych prac. Na rynku istnieje wiele rozwiązań, które mogą być do tych celów wykorzystane i warto przeanalizować ich możliwości, aby podejmując ostateczną decyzję dokonać optymalnego wyboru. Oprogramowanie, które może być wykorzystywane przy budowaniu biblioteki cyfrowej to szeroka gama produktów, poczynając od ogólnego zastosowania systemów operacyjnych, przeglądarek internetowych i pakietów biurowych, poprzez oprogramowanie dedykowane do obsługi konkretnego sprzętu digitalizacyjnego, aż do specjalistycznego systemu

biblioteki cyfrowej czy oprogramowania wspomagającego przepływ prac digitalizacyjnych. Właściwy wybór zwłaszcza tych dwóch ostatnich systemów ma kluczowy wpływ na powodzenie projektu utworzenia biblioteki cyfrowej i efektywność jej dalszego operacyjnego działania.

Oprogramowanie służące do budowania bibliotek cyfrowych umożliwia gromadzenie i udostępnianie wersji prezentacyjnych obiektów cyfrowych, a także zarządzanie takimi obiektami. Warto wyróżnić następujące funkcje (grupy funkcji) biblioteki cyfrowej:

- unikalna identyfikacja obiektów cyfrowych (w tym również łatwość tworzenia referencji do obiektów i gromadzenie statystyk wyświetleń na poziomie poszczególnych obiektów),
- przeglądanie zbiorów biblioteki cyfrowej oraz przeszukiwanie tekstu i metadanych – zarówno w postaci podstawowej jak i przeznaczonej dla użytkownika zaawansowanego,
- stopień złożoności i możliwości modyfikacji/konfiguracji architektury informacji w bibliotece cyfrowej (struktura kolekcji, obiekty cyfrowe proste i złożone etc.),
- sposób prezentacji różnych typów obiektów cyfrowych i wsparcie dla udostępniania obiektów interaktywnych i multimedialnych,
- szerokie możliwości autoryzacji dostępu do obiektów cyfrowych (w tym również możliwość wykorzystania zewnętrznych baz użytkowników oraz autoryzacja poprzez IP),
- elastyczna obsługa metadanych obiektów cyfrowych (w szczególności możliwość dowolnej konfiguracji schematu metadanych, wsparcie dla otwartych i zamkniętych słowników, wielojęzycznych metadanych itd.) z uwzględnieniem interoperacyjności w tym obszarze,
- zgodność ze standardami zwłaszcza w obszarze protokołów komunikacyjnych i formatów wymiany danych, np. METS, MARC, OAI-PMH (zarówno w kontekście importu jak i eksportu metadanych i obiektów cyfrowych),
- wsparcie dla długoterminowego przechowywania obiektów cyfrowych (np. okresowa kontrola spójności danych).

Poza oprogramowaniem umożliwiającym prezentację obiektów cyfrowych, warto przyjrzeć się również oprogramowaniu, które umożliwia zarządzanie procesem tworzenia cyfrowych kopii dokumentów przechowywanych w magazynach biblioteki. Prace związane z opracowaniem dokumentu, digitalizacją, tworzeniem wersji prezentacyjnych, korektą graficzną omówione w punkcie 3.4 muszą być sprawnie zarządzane, monitorowane i kontrolowane. W małych zespołach możliwe jest kontrolowanie i monitorowanie postępu prac przy użyciu typowych narzędzi biurowych (np. współdzielony w sieci arkusz kalkulacyjny Google Docs czy Microsoft Office Web Apps). Przy większych przedsięwzięciach warto rozważyć wykorzystanie rozwiązań służących do zarządzania i śledzenia wykonania zadań (tzw. issue tracking software, np. Mantis czy JIRA), lub też systemów dedykowanych do obsługi procesów digitalizacji, które

pozwolą na efektywną pracę i przepływ danych i informacji pomiędzy poszczególnymi pracownikami biblioteki cyfrowej.

Jeśli w bibliotece cyfrowej dla każdego zadania digitalizacji mamy do wykonania serię czynności, które przypisane są do poszczególnych osób, to możliwe jest zaprojektowanie tego w oprogramowaniu. Dzięki temu każda osoba będzie w pełni świadoma tego, za które czynności odpowiada i ile zadań ma aktualnie do wykonania. Przykładowo czynności związane z digitalizacją mogą być następujące: opracowanie merytoryczne, digitalizacja, konwersja, weryfikacja, udostępnienie online. W przypadku braku oprogramowania wszystkie te elementy muszą być koordynowane przez osoby odpowiedzialne za realizację poszczególnych czynności bądź też całego procesu. Oprogramowanie może ułatwić przepływ informacji, tzn. każdy z użytkowników po zalogowaniu do systemu widzi listę czynności do wykonania. Jeśli skanujący wykonał czynność digitalizacji i oznaczył ją w ten sposób w oprogramowaniu, to osoba wykonująca konwersję automatycznie widzi nową czynność do wykonania na swojej liście. Co więcej, czynności typu konwersja, czy nawet udostępnienie online mogą zostać, w przypadku użycia kompleksowego oprogramowania, zautomatyzowane. Na przykład: użytkownik oznacza, że wykonał czynność digitalizacji, w tym momencie system automatycznie pobiera pliki matki i automatycznie generuje wersję prezentacyjną z uwzględnieniem wyznaczonego profilu jej tworzenia (np. profil dla map). Po wygenerowaniu wersji prezentacyjnej następuje weryfikacja przez użytkownika systemu, a po jej wykonaniu system może automatycznie umieścić obiekt na portalu biblioteki cyfrowej, udostępniającym wersje prezentacyjne, wiążąc je z wcześniej przygotowanym opisem. Jak widać odpowiednio dobrane oprogramowanie do wspomagania przepływu prac digitalizacyjnych może znacznie uprościć działanie biblioteki cyfrowej i w znacznym stopniu zautomatyzować poszczególne czynności. Przykład dość szczegółowej analizy wybranych systemów do budowy bibliotek cyfrowych dostępnych na licencji open source przeprowadzonej pod kątem projektu konkretnej instytucji zawiera raport [NRGL 2009].

3.6. Promocja biblioteki cyfrowej

Przed udostępnieniem zasobów biblioteki cyfrowej w Internecie należy rozważyć kwestie związane z jej promocją i inicjacją. W tym kontekście należy wziąć pod uwagę możliwość kontaktu z potencjalnymi użytkownikami portalu w celu identyfikacji ich preferencji oraz zebrania opinii na temat finalnego kształtu stron internetowych. Jeśli obsługa pewnych obszarów biblioteki cyfrowej wymaga dodatkowych informacji, należy takie zamieścić i umożliwić łatwe ich odnalezienie.



Rys. 2. Komponent społecznościowy z portalu Facebook na stronach Biblioteki Cyfrowej Uniwersytetu Wrocławskiego (http://www.bibliotekacyfrowa.pl/).

Atrakcyjne mogą się też okazać funkcje społecznościowe portalu biblioteki cyfrowej – zaimplementowane bezpośrednio w portalu (np. lista ulubionych publikacji w bibliotece cyfrowej) lub powstałe na skutek wykorzystania funkcji dobrze znanych portali społecznościowych (por. Rysunek 2). O oficjalnym uruchomieniu biblioteki cyfrowej warto poinformować media (zarówno ogólne, jak i regionalne czy branżowe), przygotować ulotki i rozdystrybuować je wśród czytelników biblioteki. Dobrym pomysłem może być też zaproszenie na uroczystość otwarcia przedstawicieli różnego rodzaju władz, ekspertów z dziedziny bibliotek cyfrowych oraz doświadczonych pracowników technicznych.

3.7. Monitorowanie wykorzystania zasobów biblioteki cyfrowej

Uruchomienie biblioteki cyfrowej powinno wiązać się z rozpoczęciem monitorowania wykorzystania jej zasobów. W tym celu można skorzystać zarówno ze statystyk dostępnych w samym oprogramowaniu biblioteki cyfrowej (np. statystyki liczby wyświetlanych obiektów w danym miesiącu, statystyki dotyczące wyszukiwania), jak

i tych, które mogą być zbierane przez narzędzia ogólnego przeznaczenia typu AWStats (http://awstats.sourceforge.net/), czy Google Analytics (http://www.google.com/analytics/) (por. Rysunek 3). Skorzystanie z tych narzędzi jest stosunkowo proste i w zasadzie bezpłatne, a korzyści jakie wynikają z przeglądania generowanych przez nie raportów mogą być nieocenione. Na podstawie statystyk możliwe jest określenie trendów wykorzystania poszczególnych stron portalu biblioteki cyfrowej, co z kolei może dostarczyć cennych informacji nie tylko dotyczących potrzeb w modyfikacji interfejsu użytkownika, ale również potrzeb modyfikacji w kontekście selekcji dokumentów, które zostaną poddane digitalizacji.



Rys. 3. Przykładowe raporty Google Analytics dla Wielkopolskiej Biblioteki Cyfrowej.

Ponadto warto umożliwić użytkownikom kontakt z personelem biblioteki cyfrowej, tak by mogli oni wywierać wpływ na budowanie zasobów i udoskonalanie kształtu biblioteki cyfrowej. Informacje na temat funkcjonowania biblioteki cyfrowej można również uzyskać poprzez ankietowanie użytkowników. Informacje zebrane w ten sposób mogą również pomóc w zrozumieniu sposobu pracy użytkowników i ich potrzeb [Mazurek 2010].

4. Podsumowanie

W Polsce najpopularniejsze są dwa modele tworzenia biblioteki cyfrowej: instytucjonalny i konsorcjalny. Różnią się one liczbą zaangażowanych instytucji oraz stopniem ich odpowiedzialności za poszczególne obszary biblioteki cyfrowej. W obu przypadkach, dla pomyślnego uruchomienia usług biblioteki cyfrowej, konieczna jest odpowiednia organizacja pracy. Sprowadza się to do identyfikacji najważniejszych

obszarów działań biblioteki cyfrowej oraz umiejscowienia ich w przepływie danych i informacji związanym z pracami digitalizacyjnymi. W odniesieniu do niektórych obszarów (np. digitalizacji) niezbędna jest decyzja o wykonaniu pracy w ramach instytucji lub zleceniu do wykonania zewnętrznym podwykonawcom. Przed rozpoczęciem projektu budowy biblioteki cyfrowej należy w sposób świadomy i przemyślany zdefiniować jej wizję oraz cele przed nią stojące. Niezbędna będzie również decyzja co do formy tworzenia biblioteki cyfrowej – czy będzie miała ona charakter instytucjonalny, czy też konsorcjalny. Definicja procesu przepływu pracy oraz identyfikacja występujących w nim ról jest zasadniczym elementem w procesie tworzenia biblioteki cyfrowej i bardzo wpływa na efektywność funkcjonowania takiej biblioteki.

Kolejnym kluczowym etapem jest wybór a następnie adaptacja i konfiguracja oprogramowania, które ułatwi gromadzenie i udostępnianie cyfrowych zasobów oraz usprawni organizację prac digitalizacyjnych. Analiza dostępnych w oprogramowaniu funkcji ułatwi wybór odpowiedniego rozwiązania. Uruchamiając bibliotekę cyfrową warto zacząć również promować ją wśród potencjalnych użytkowników i zachęcić wszystkich zainteresowanych do odwiedzania jej stron.

Uruchomienie biblioteki cyfrowej to pierwszy krok w jej działalności. Z biegiem czasu warto weryfikować misję i wizję biblioteki cyfrowej, tak by odpowiadać na bieżące i zmieniające się przecież potrzeby użytkowników. Rozbudowa grup docelowych użytkowników, nowe formy prezentacji dokumentów, nowe typy dokumentów to tylko nieliczne możliwości rozwoju biblioteki cyfrowej. W kontekście upływu czasu istotną rolę odgrywać będzie długoterminowe przechowywanie obiektów źródłowych (plików matek) wraz z polityką ich migracji do nowszych formatów, tak by mogły być odczytane z użyciem kolejnych konfiguracji sprzętowo-programowych. Monitorowanie procesu digitalizacji, monitorowanie stopnia wykorzystywania obiektów cyfrowych są elementarnymi zadaniami w procesie utrzymania i rozwoju biblioteki cyfrowej, a reakcja na wnioski płynące z tych analiz jest kluczem do sukcesu każdej biblioteki cyfrowej.

Bibliografia

1. Fundacja Europeana (2010). *Europeana – statut domeny publicznej* [dok. elektr.] http://www.version1.europeana.eu/web/europeana-project/public-domain-charter-pl [odczyt: 02.08.2011].
2. Górny, Mirosław; Gruszczyński, Paweł; Mazurek, Cezary; Nikisch, Jan Andrzej; Stroiński, Maciej; Swędrzyński, A. (2003). *Zastosowanie oprogramowania dLibra do budowy Wielkopolskiej Biblioteki Cyfrowej.* „Zeszyty Naukowe Wydziału ETI Politechniki Gdańskiej" [I Krajowa Konferencja Technologie Informacyjne. Gdańsk], s. 109–117.
3. Mazurek, Cezary, Parkoła, Tomasz; Werla, Marcin (2007). *Rozwój bibliotek cyfrowych w Polsce.* [W:] *Partnerzy bibliotek: model komunikacji z otoczeniem. IX Ogólnopolska Konferencja Bibliotek Szkół Wyższych Niepublicznych* (2007). Łódź: Firma Księgarska Wiesława Juszczaka, s. 239-246.

4. Mazurek, Cezary, Stroiński, M., Węglarz, Jan (2001). *Znaczenie programu Pionier dla rozwoju bibliotek cyfrowych w Polsce.* [W:] *5 [pięć] lat Poznańskiej Fundacji Bibliotek Naukowych* (2001). Poznań: PFBN, s. 63-69.
5. Mazurek, Cezary; Stroiński, Maciej; Werla, Marcin; Węglarz, Jan (2009). *Infrastruktura bibliotek cyfrowych w sieci PIONIER.* [W:] Mazurek, Cezary; Stroiński, Maciej; Węglarz, Jan red. (2009). *Polskie Biblioteki Cyfrowe 2008: materiały z konferencji zorganizowanej w dniach 24-25 listopada 2008 r. przez Bibliotekę Kórnicką PAN, Poznańską Fundację Bibliotek Naukowych, Poznańskie Centrum Superkomputerowo-Sieciowe.* Poznań: Ośrodek Wydawnictw Naukowych, s. 9-14.
6. Mazurek, Jolanta (2010). *Bibliotekarskim okiem: jak nie ulec magii liczb, czyli co nam mówią statystyki biblioteki cyfrowej.* [W:] Mazurek, Cezary; Stroiński, Maciej; Węglarz, Jan red. (2011). *Polskie Biblioteki Cyfrowe 2010 : materiały z konferencji zorganizowanej w dniach 20-21 października 2010 r. przez Bibliotekę Kórnicką PAN, Poznańską Fundację Bibliotek Naukowych, Poznańskie Centrum Superkomputerowo-Sieciowe.* Poznań: Ośrodek Wydawnictw Naukowych, s. 89-97.
7. NRGL – National Repository of Grey Literature (2009). *Comparison of Selected Software Systems for Creation of Digital Libraries from the Field of Open Source for the Needs of the NRGL STL* [dok. elektr.] http://nrgl.techlib.cz/images/Open_source.pdf [odczyt: 02.08.2011].
8. Open Archives Initiative (2006). *Specification and XML Schema for the OAI Identifier Format* [dok. elektr.] OAI-PMU Implementation Guidelines http://www.openarchives.org/OAI/2.0/guidelines-oai-identifier.htm [odczyt: 02.08.2011].
9. Potęga, Joanna, Wróbel, Agnieszka (2010). *The Dublin Core Metadata Element Set, Ver. 1.1 a potrzeby i oczekiwania bibliotekarzy cyfrowych – analiza przypadków.* [W:] Mazurek, Cezary; Stroiński, Maciej; Węglarz, Jan red. (2010). *Polskie Biblioteki Cyfrowe 2009 : materiały z konferencji zorganizowanej dnia 9 grudnia 2009 r. przez Bibliotekę Kórnicką PAN, Poznańską Fundację Bibliotek Naukowych, Poznańskie Centrum Superkomputerowo-Sieciowe.* Poznań: Ośrodek Wydawnictw Naukowych, s. 71-78.
10. Poznańskie Centrum Superkomputerowo-Sieciowego (2011). *Dokumentacja schematu metadanych PLMET* [dok. elektr.] *Wiki Poznańskiego Centrum Superkomputerowo-Sieciowego* [odczyt: 02.08.2011]. Dostępny w WWW: http://confluence.man.poznan.pl/community/x/UQDC
11. Rychlewski, Jacek; Węglarz, Jan; Starzak, Stanisław [i in.]. (2000) *PIONIER: Polish Optical Internet, ISThmus2000, 11-13 kwietnia 2000.* Poznań: Instytut Informatyki Politechniki Poznańskiej, s. 19-28.
12. The Dublin Core Metadata Initiative (2010). *DCMI Metadata Terms* [dok. elektr.] DCMI Home: Dublin Core Metadata Initiative, http://dublincore.org/documents/dcmi-terms/ [odczyt: 02.08.2011].

Opis dokumentów

1. Struktury danych
2. Metadane w polskich bibliotekach cyfrowych
3. Odwzorowanie treści dokumentów elektronicznych
4. Opracowanie rzeczowe dokumentu
5. Korelacje między opisami w katalogach bibliotecznych i bibliotekach cyfrowych

Marek Nahotko, Piotr Myszkowski, Łukasz Mesek, Danuta Patkaniowska

W tekście przygotowanym przez zespół autorów przedstawiona została problematyka opisu formalnego oraz przedmiotowego danych począwszy od ich struktury, opisu metadanych wykorzystywanych w polskich bibliotekach cyfrowych, zagadnień związanych z odwzorowywaniem treści i wykorzystywanych dla charakterystyki dokumentów różnych języków informacyjno-wyszukiwawczych, aż po analizę porównawczą opisów odnajdywanych w katalogach tradycyjnych i cyfrowych bibliotek.

***

Marek Nahotko

1. Struktury danych

Według Jane Greenberg termin 'metadane' po raz pierwszy użyty został w 1969 r., a w druku ukazał się w 1973 r. [Greenberg 2005, s. 19]. Szybko przyjął się wśród bibliotekarzy, informatyków i przedstawicieli innych zawodów, związanych z przetwarzaniem informacji dla określenia danych o danych. Oznacza on stosowanie atrybutów metadanych służących do opisu, wskazywania relacji, określania jakości lub wskazywania innych charakterystyk opisywanego obiektu (czyli danych).

Tworzenie metadanych początkowo było przez bibliotekarzy odróżniane od katalogowania i dotyczyło wyłącznie dokumentów elektronicznych. Ze względu jednak na to, że bibliotekarze katalogowali zasoby elektroniczne na długo przed powstaniem Webu, a wiele schematów metadanych (włącznie z Dublin Core) można używać do

opisu obiektów nie tylko cyfrowych, ale również fizycznych, rozróżnienie na formaty danych bibliograficznych i schematy metadanych przestało funkcjonować. Cechą wspólną różnych definicji metadanych, dotyczących zasobów informacyjnych (materiałów udostępnianych przez biblioteki – tradycyjne i cyfrowe oraz archiwa, muzea i inne ośrodki informacji) jest nacisk na ich aspekt funkcjonalny, wyrażany przez określanie metadanych jako ustrukturyzowanych danych o danych. W tym przypadku metadane mogą być rozumiane jako ustrukturyzowane dane o opisywanym obiekcie, wspomagające wyznaczone dla niego funkcje; przy czym obiektem jest wówczas każda jednostka, dla której można wyznaczyć i zapisać dane kontekstualne.

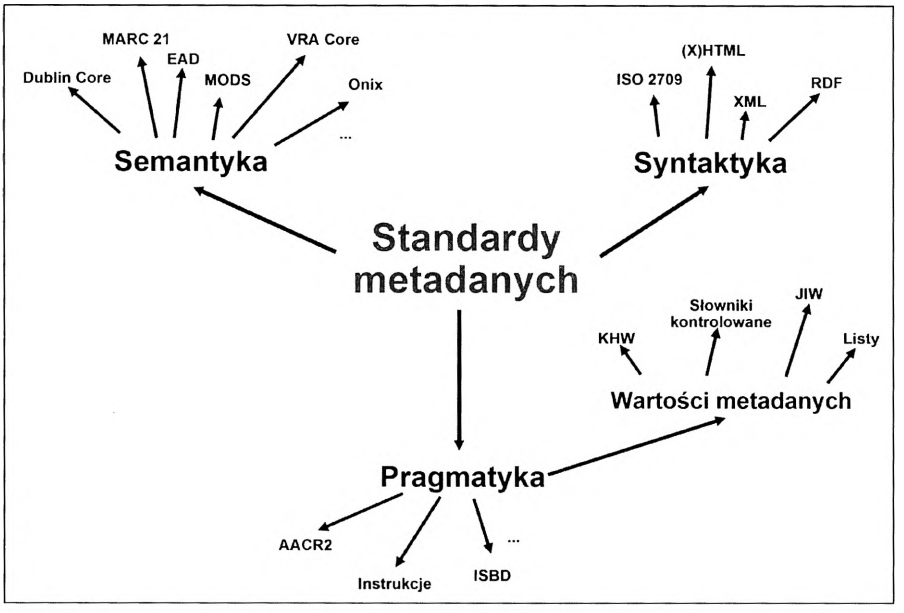
Metadane ustrukturyzowane, rozumiane w ten sposób, często przeciwstawiane są metadanym nieustrukturyzowanym, z którymi mamy do czynienia w indeksach wyszukiwarek internetowych. Nie jest to jednak podział ścisły, gdyż już od dawna wyszukiwarki nie funkcjonują na zasadzie tworzenia zwykłych indeksów odwróconych, będących efektem prostego cięcia tekstu na wyrazy [Nahotko 2010, s. 114-115]. Nazwa „metadane nieustrukturyzowane" może być więc myląca, gdyż wyszukiwarki również tworzą struktury w swoich zasobach indeksowych [Battelle 2006, s. 22]. Te skomplikowane, programistyczne narzędzia wyszukiwawcze gromadzą sporą ilość danych na temat indeksowanych stron (a więc ich metadanych), takich, jak: adresy URL, identyfikatory indeksowanych stron i dane o nich (np. długość strony), słowa ze strony wraz z informacją o stronie, na jakiej wystąpiło słowo, jego pozycji w dokumencie, rozmiarze czcionki, odnośnikach, tekście wokół i w obrębie odnośnika. Gromadzone są także dane dotyczące zachowań użytkowników, tzw. strumienie kliknięć. W tym celu wyszukiwarki obserwują i zapisują tak zwane intencje użytkownika – czego szuka, które wyniki z przedstawionej listy wybiera, jakie strony później są odwiedzane. Wyszukiwarka tworzy swojego rodzaju klasyfikację treści zasobów Web na podstawie zapytań użytkowników.

Najważniejsza jest jednak stała, głęboka analiza pozyskiwanych metadanych, nieznana w systemach bibliotecznych. Podstawą funkcjonowania współczesnych firm, tworzących światowe wyszukiwarki, jest analiza baz indeksowych, której efektem jest odkrywanie zależności statystycznych i algorytmizowalnych zjawisk, wykorzystywanych do zapewnienia jak najwyższej trafności odpowiedzi na zapytania użytkownika. Dzięki tego typu narzędziom i procesom statystycznym wyszukiwarka, na podstawie gromadzonych przez siebie metadanych oraz ich modeli i struktur powiązanych relacjami, które można określić mianem ontologii[1], staje się systemem samouczącym i dostosowującym do potrzeb użytkowników – im więcej osób wyszukuje w Internecie, tym wyszukiwanie staje się bardziej efektywne. Pomimo tego, że większość algorytmów stosowanych w wyszukiwarkach internetowych używa głównie metody statystycznej analizy tekstu, jakość ich działania jest wystarczająco wysoka dla efektywnego wyszukiwania informacji [Gawrysiak 2008, s. 658].

Po tym krótkim wstępie, przedstawiającym źródła metadanych we współczesnych systemach wyszukiwania informacji, można przejść do omówienia podstawowych

[1] Ontologie są tu rozumiane jako modele dziedzin wiedzy, odwzorowujące relacje w nich zachodzące, przy pomocy ujawnionej, relacyjnej struktury pojęć.

problemów metadanych ustrukturyzowanych, tworzonych przez specjalistów-bibliotekarzy. Przedstawiona zostanie ich semantyka, syntaktyka i pragmatyka (rys. 1) na przykładzie dwóch, podstawowych, stosowanych w Polsce standardów metadanych: MARC 21 i Dublin Core Metadata Element Set (DC).



Rys. 1. Standardy metadanych ustrukturyzowanych

1.1. Semantyka metadanych

Semantyka określa znaczenie gromadzonych metadanych. Wyrażana jest w odpowiednim ustrukturyzowaniu rekordu metadanych, podziale na pola i podpola (MARC 21) lub atrybuty i podatrybuty (DC). Semantyczne różnice pomiędzy wymienionymi dwoma schematami metadanych wynikają z różnic historycznych ich rozwoju. MARC rozwijany jest od lat 60. ubiegłego wieku. Kolejno dodawano struktury niezbędne dla tworzenia opisów poszczególnych rodzajów dokumentów, najpierw tekstowych, później także audiowizualnych. W podobny sposób próbowano tworzyć struktury metadanych dla stron Web oraz oprogramowania. Szybko okazało się jednak, że formę elektroniczną może posiadać każdy rodzaj dokumentu (dane), dla którego wcześniej stworzono odpowiednie struktury: elektroniczny może być zarówno dokument tekstowy (książka, czasopismo, artykuł), dźwiękowy (nagranie muzyczne lub niemuzyczne), obraz (nieruchomy lub film) i każdy inny. Wówczas opis tworzony według zasad dla określonego typu dokumentu zaczęto uzupełniać elementami specyficznymi dla dokumentu elektronicznego, z uwzględnieniem trybu dostępu: lokalnego lub zdalnego (zob. tab. 1).

Tabela 1. Zastosowanie pól MARC 21 do opisu dok. elektronicznego.

Pola MARC 21	Dokument o dostępie lokalnym	Dokument o dostępie zdalnym
256 Rodzaj i wielkość dokumentu (strefa specjalna)	**Zalecane**	**Obowiązkowe**
300 Opis fizyczny	**Stosowane** Nośnik fizyczny może być opisany	**Pomijane** Nośnik fizyczny nie może być opisany (charakterystykę dokumentu podaje się w polu 256)
538 Uwaga dot. systemu	**Obowiązkowe** Uwaga o wymaganiach systemowych	**Obowiązkowe** Uwaga o trybie dostępu
856 Lokalizacja elektroniczna	**Pomijane** Brak informacji do wypełnienia pola	**Obowiązkowe**

Sposób tworzenia schematu Dublin Core i innych, nowych schematów metadanych, był inny. Powstawały one z myślą o konkretnych zastosowaniach; albo dla określonego typu obiektów cyfrowych (np. DC stosowany jest głównie dla tekstowych obiektów cyfrowych, tzw. DLO[2], a MIX dla grafiki), albo dla wybranych zastosowań (metadane biblioteczne, muzealne – CIDOC, VRA, archiwalne – EAD, wspomagające handel, obsługę praw autorskich – ONIX itp.), albo do zastosowań specjalnych, np. schemat „konteneru" METS dla kodowania danych z wielu schematów metadanych dla ich wyszukiwania, przesyłania i archiwizowania. Taka metoda tworzenia struktur metadanych powoduje, że nie istnieje format, który mógłby zaspokoić wszystkie potrzeby związane z opisem sieciowych zasobów elektronicznych. Co więcej, struktury te muszą często uwzględniać potrzeby lokalne, na poziomie poszczególnych systemów informacyjnych. Ich dostosowanie do tych wymagań umożliwiają tzw. profile aplikacyjne[3].

MARC 21 często przeciwstawiany jest DC pod względem skomplikowania struktur obu schematów; w MARC, w samym rekordzie bibliograficznym wyróżnia się ponad 250 pól zmiennej długości (każde z nich dzielone na podpola, czasami w liczbie kilkudziesięciu), podczas gdy DC to... tylko 15 atrybuentów. Tak wielka różnica jest jednak pozorna. Po pierwsze właściwie od początku istnienia DC był stale wzbogacany i komplikowany, szczególnie od czasu, gdy pierwotne zamierzenie, aby rekordy DC tworzyli autorzy obiektów cyfrowych (a więc nie-katalogerzy), okazało się nierealne [Chopey 2005, s. 256]. Przełomowy był w tym zakresie rok 2008, w którym przyjęto tzw. DCMI Metadata Terms, opisujący ponad sto atrybutów bardzo różnego rodzaju [Brachfogel 2010, s. 58]. Wśród nich są także wspomniane profile aplikacyjne, zawierające kolejne elementy metadanych, pobieranych z różnych standardów,

[2] DLO – ang. Document-Like Object.

[3] DCMI przygotowało w ramach DC Metadata Terms cztery profile aplikacyjne: Dublin Core Collections Application Profile, Dublin Core Education Application Profile (w oprac.), Dublin Core Government Application Profile (w oprac.), Dublin Core Library Application Profile.

w tym także z MARC 21. Profil DC przeznaczony do zastosowań bibliotecznych zawiera na przykład 44 atrybuty. Dublin Core przestał już być zestawem zawierającym wyłącznie najbardziej podstawowe atrybuty, co zapewne wiąże się ze zmianą profilu użytkowników; nie tworzy się już go do użytku przez twórców stron Web, a nakierowany jest raczej na potrzeby bibliotekarzy (wykorzystujących go m.in. w bibliotekach cyfrowych), którzy są w stanie poradzić sobie z bardziej rozbudowanym schematem.

Podczas projektowania semantyki metadanych należy brać pod uwagę ich funkcjonalność. Według zasad FRBR schemat metadanych powinien zapewnić poprawną realizację czterech funkcji: identyfikację opisywanego obiektu, jego wyszukiwanie, wybór oraz uzyskanie dostępu [IFLA 1998, s. 84-86]. Dla jednoznacznej identyfikacji dokumentu, według polskich przepisów [Lenartowicz 1989, s. 3], wystarczy pięć elementów metadanych (opis na pierwszym poziomie szczegółowości). Według doświadczeń polskich bibliotekarzy, użytkownicy do wyszukiwania nie stosują nawet takiej liczby danych; jak wynika z badań nad serwisami informacyjnymi, oferującymi ustrukturyzowane metadane, użytkownicy zainteresowani są wyszukiwaniami według najprostszych kryteriów: autora i tytułu, najczęściej pomijając wszelkie inne kryteria formalne i rzeczowe[4]. Możliwości wyszukiwawcze systemu, w którym stosowane są metadane, tylko częściowo zależne są od struktur metadanych; w większym stopniu są one efektem rozwiązań informatycznych, przyjętych przez twórców systemu bibliotecznego. Z punktu widzenia formatu MARC do wyszukiwania przeznaczone są pola haseł w rekordzie bibliograficznym (1XX, 6XX, 7XX, większość 8XX). Oprócz tego systemy biblioteczne pozwalają na wyszukiwanie według zawartości innych pól opisu, w tym również pełnotekstowe wyszukiwanie w całym rekordzie bibliograficznym. Również w DC zawartość wszystkich atrybutów rekordu może być przeszukiwana.

Uzyskanie dostępu do wersji elektronicznej opisywanego obiektu możliwe jest z poziomu rekordów obu struktur metadanych. W MARC 21 służy do tego pole 856, a w DC atrybut Identyfikator, w których zapisywany jest (między innymi) adres internetowy. Podobnie jak w przypadku wyszukiwania, istotną rolę odgrywa tu sposób oprogramowania funkcji systemu bibliotecznego.

Obie struktury metadanych, stosowane w polskich bibliotekach cyfrowych, są aplikacjami różnych filozofii tworzenia rekordów bibliograficznych. Stąd próby ujednolicania tych struktur, poprzez dopasowywanie pól i atrybutów metadanych z góry są skazane na niepowodzenie. Możliwe jest jedynie stosowanie tablic przejścia i innych narzędzi służących konwersji metadanych, zwykle przebiegającej ze stratami. Jest wiele przyczyn takiej sytuacji. Jedną z nich jest to, że w jednym polu MARC 21 zapisywanych jest wiele różnych informacji. Przykładem może być zawartość pola 245, która może być konwertowana do wielu różnych atrybutów DC:

- Pole 245, podpole *a* przenosi się do atrybutu Tytuł,
- Pole 245, podpole *b* przenosi się do atrybutu Tytuł,

[4] Przykładowo por. dane dotyczące wyszukiwania w katalogu KaRo przedstawione przez Tomasza Wolniewicza: 98,5% wszystkich wyszukiwań prowadzonych jest według jednego lub dwóch kryteriów, którymi są najczęściej tytuł (62%), autor (18%), ISBN (16%), dla dwóch kryteriów: autor+tytuł (95% wyszukiwań według połączenia tych dwóch kryteriów) [Wolniewicz 2010, s. 159].

- Pole 245, podpole *c* przenosi się do atrybutu Twórca i/lub Współtwórca,
- Pole 245, podpole *f* przenosi się do atrybutu Data,
- Pole 245, podpole *g* przenosi się do atrybutu Data,
- Pole 245, podpole *h* przenosi się do atrybutu Format,
- Pole 245, podpole *k* przenosi się do atrybutu Rodzaj,
- Pole 245, podpole *n* przenosi się do atrybutu Opis lub Tytuł,
- Pole 245, podpole *p* przenosi się do atrybutu Opis lub Tytuł.

Dostosowywanie struktury DC do rozwiązań z MARC 21 może także prowadzić do tego, że wystąpienia atrybutów DC zawierają więcej danych, niż należy. Przykładem może być atrybut Wydawca, w którym umieszczana bywa cała zawartość pola 260 z MARC 21, czyli pełny adres wydawniczy [Jackson i in. 2008, s. 14]. Problemy mogą wynikać także z odmiennego nazewnictwa, na przykład opis fizyczny z pola 300 w MARC 21, w DC mieści się częściowo w atrybucie Rodzaj, a częściowo w atrybucie Format, natomiast atrybut Opis fizyczny nie istnieje [Hillman 2005].

1.2. Syntaktyka metadanych

Syntaktyka określa sposób fizycznego zapisu rekordów metadanych na nośniku maszynowym, a więc zasady ich kodowania. Musi ono być czytelne i jednoznaczne przede wszystkim dla oprogramowania komputerowego. Możliwość odkodowania zapisów bezpośrednio przez człowieka jest przydatna, ale nie ma znaczenia pierwszorzędnego. Sytuację komplikuje fakt, że wielu dostawców systemów bibliotecznych stosuje własne sposoby zapisu danych na nośnikach, często związanych ze stosowanym SZBD, natomiast systemy kodowania służą wówczas do definiowania sposobu wyświetlania i wymiany danych. Dotyczy to szczególnie systemów z formatem MARC, Dublin Core ma nieco inną specyfikę.

Do zapisu danych na nośnikach w formacie MARC stosowana jest syntaktyka będąca jedną z możliwych aplikacji ISO 2709 – standardu pomyślanego do ujednolicenia struktur danych przesyłanych na taśmach magnetycznych. Jeżeli nie ma rozbieżności na poziomie semantycznym, każdy system obsługujący bibliotekarską wersję ISO 2709 powinien bez większych kłopotów importować dane z każdej bazy danych MARC. Syntaktyka ta zapewnia w świecie bibliotekarskim wsadową wymianę danych.

Sytuacja gorzej wygląda, gdy zechcemy wyjść poza ten świat. Do niedawna nie było to konieczne – biblioteki posiadały monopol na dane bibliograficzne, inni, jeśli chcieli współpracować, musieli się dostosowywać. Od pewnego czasu sytuacja się zmieniła i obecnie biblioteki są tylko jednym z wielu źródeł danych. Wymusza to zdolność do współpracy. W środowisku sieci globalnych ktoś, kto nie potrafi dostosować się do powszechnie stosowanych standardów, przestaje się liczyć.

ISO 2709 poza bibliotekami praktycznie nie jest stosowany. Standardem w Internecie jest kodowanie danych (wszelkiego rodzaju) w XML, który zapewnia ich kodowanie wraz z określeniem znaczenia, a przy tym jest znacznie bardziej elastyczny

pod względem definiowania struktury danych; definiowanie to pozostawia się twórcy systemu.

Dublin Core jest schematem niezależnym od przyjętej syntaktyki. Rekordy DC mogą być kodowane w HTML, XHTML, XML, RDF i innych standardach. Można sobie nawet wyobrazić kodowanie DC w ISO 2709, chociaż chyba nikt tego nie próbował. DCMI udostępnia instrukcje kodowania wyrażeń DC w każdej z tych syntaktyk. Dzięki temu rekordy DC są częścią współczesnego środowiska sieci rozległych, mogą być gromadzone przez każdego, kto tego potrzebuje (np. przy użyciu OAI-PMH, jeśli lokalny system na to pozwala) i mogą podlegać dowolnym manipulacjom, np. scalaniu, obróbce statystycznej.

Istnieją próby unowocześnienia syntaktyki MARC przez zastosowanie XML w miejsce ISO 2709 [Tennant 2007, s. 387]. Idą one w dwóch kierunkach. Po pierwsze, próbuje się stworzyć zasady kodowania rekordów MARC w XML, czego efektem jest tzw. MARCXML. Do zakodowania treści rekordu stosuje się znaczniki znane z XML, pozostawiając jednocześnie etykiety pól i podpól oraz wskaźniki bez zmian (rys. 2). W takiej postaci rekordy MARC 21 mogą być wymieniane pomiędzy systemami informacyjnymi z zastosowaniem protokołu SRU/SRW[5].

```
<datafield tag="100" ind2=" " ind1="1">
    <subfield code="a">Sandburg, Carl,</subfield>
    <subfield code="d">1878-1967.</subfield>
</datafield>
<datafield tag="245" ind2="0" ind1="1">
    <subfield code="a">Arithmetic /</subfield>
    <subfield code="c">Carl Sandburg ; illustrated as an anamorphic adventure by Ted Rand.</subfield>
</datafield>
<datafield tag="250" ind2=" " ind1=" ">
    <subfield code="a">1st ed.</subfield>
</datafield>
<datafield tag="260" ind2=" " ind1=" ">
    <subfield code="a">San Diego :</subfield>
    <subfield code="b">Harcourt Brace Jovanovich,</subfield>
    <subfield code="c">c1993.</subfield>
</datafield>
<datafield tag="300" ind2=" " ind1=" ">
    <subfield code="a">1 v. (unpaged) :</subfield>
    <subfield code="b">ill. (some col.) ;</subfield>
    <subfield code="c">26 cm.</subfield>
</datafield>
```

Rys. 2. Fragment rekordu MARC zakodowanego w XML (http://www.loc.gov/standards/marcxml/).

Innym rozwiązaniem jest utworzenie nowego schematu metadanych, zachowującego w maksymalnym stopniu strukturę i filozofię MARC, ale ze zmodyfikowanymi etykietami (alfanumerycznymi). W efekcie powstaje nowy schemat metadanych, jakim jest MODS (rys. 3), utworzony i zarządzany przez Library of Congress. Pomiędzy MARC 21 a MODS możliwa jest bezstratna konwersja danych; w odwrotną stronę konwersja może powodować niewielką utratę danych.

[5] SRU/SRW (Search and Retrieve URL/Web Service) – usługa Web dla wyszukiwania z użyciem semantyki protokołu Z39.50.

```
<?xml version="1.0"?>
<mods xsi:schemaLocation="http://www.loc.gov/mods/ http://www.loc.gov/standards/mods/mods.xsd" xmlns="http://www.loc.gov/mods/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:xlink="http://www.w3.org/TR/xlink">
    <titleInfo>
        <title>Arithmetic /</title>
    </titleInfo>
    <name type="personal">
        <namePart>Sandburg, Carl</namePart>
        <namePart type="date">1878-1967</namePart>
        <role>
            <text>creator</text>
        </role>
    </name>
    <name type="personal">
        <namePart>Rand, Ted</namePart>
        <role>
            <text>ill.</text>
        </role>
    </name>
    <typeOfResource>text</typeOfResource>
    <originInfo>
        <place>
            <code authority="marc">cau</code>
            <text>San Diego</text>
        </place>
        <publisher>Harcourt Brace Jovanovich</publisher>
        <dateIssued>c1993</dateIssued>
        <dateIssued encoding="marc">1993</dateIssued>
        <edition>1st ed.</edition>
        <issuance>monographic</issuance>
    </originInfo>
    <language authority="iso639-2b">eng</language>
    <physicalDescription>
```

Rys. 3. Ten sam rekord (z rys. 2) w strukturze MODS (http://www.loc.gov/standards/marcxml/Sandburg/sandburgmods.xml).

Zmiana języka kodowania nie jest jednak łatwym procesem, ze względu na olbrzymią liczbę danych zakodowanych w ISO 2709, nagromadzoną w systemach bibliotecznych przez ostatnie 50 lat na całym świecie.

1.3. Pragmatyka metadanych

Poprawna struktura danych, umożliwiająca właściwą identyfikację i wyszukiwanie dokumentów oraz możliwość zapisu danych w tej strukturze na nośnikach maszynowych, nie rozwiązują wszystkich bibliotecznych problemów z metadanymi. Pozostaje kwestia jednolitego stosowania utworzonych struktur, zarówno na poziomie poszczególnych bibliotek, ale także we wszystkich bibliotekach, często bardzo odległych, przez bibliotekarzy mówiących różnymi językami, posługujących się różnymi alfabetami i funkcjonujących w różnych kulturach, dzięki czemu zasoby bibliotek stają się częścią jednolitego zasobu globalnego.

Bardzo istotne jest określenie sposobu zapisu danych w poszczególnych miejscach opisu (polach, atrybutach), wyznaczonych jego strukturą. W tym zakresie stosowane są przepisy katalogowania, takie jak AACR i ISBD oraz bardziej szczegółowe instrukcje, dokładnie opisujące sposób tworzenia metadanych, stosowane przez poszczególne biblioteki lub grupy bibliotek. Instrukcje te pozwalają względnie ujednolicić opis, dzięki czemu możliwe jest rozumienie znaczenia poszczególnych elementów opisu bez potrzeby znajomości języka naturalnego, a nawet alfabetu, w jakim sporządzono opis (istotną rolę odgrywają tu także znaki specjalne, przewidziane w przepisach).

Ze względu na wieloletnią współpracę w skali międzynarodowej pragmatyka związana ze stosowaniem semantyki MARC 21 jest bardzo rozbudowana. Znacznie gorzej pod tym względem wygląda pragmatyka Dublin Core. Wynika to również z faktu, że

twórcy DC z góry założyli znaczną elastyczność rozwiązań lokalnych. Również jednak tutaj obowiązują pewne zasady, jak na przykład zasada Dumb-Down, stosowana podczas uszczegółowiania semantyki schematu oraz zasada 1:1, nakazująca tworzenie odrębnych opisów dla różnych wersji dzieła (np. dla druku i jego skanu). Z braku bardziej szczegółowych instrukcji, w bibliotekach cyfrowych często przejmowane są zasady stosowane w bibliotekach tradycyjnych.

Pragmatyka metadanych musi oczywiście być ściśle związana z ich semantyką (i pośrednio syntaktyką). Z tego właśnie wynika m.in. denerwująca czasem potrzeba wpisywania w MARC 21 tego samego rodzaju danych w kilka miejsc rekordu. Może to wynikać np. z faktu, że te same dane w jednym polu służą do ich wyświetlania w formie zgodnej z ISBD, a w innym muszą mieć formę odpowiednią do automatycznego przetwarzania przez komputer.

Częścią pragmatyki metadanych są różnego rodzaju listy, słowniki i wykazy elementów pobieranych podczas tworzenia rekordów. Najbardziej oczywistymi przykładami tego typu narzędzi są kartoteki haseł wzorcowych formalnych oraz rzeczowych (tezaurusy, słowniki jhp, tablice klasyfikacji). Należą tu także proste listy alfabetyczne takich elementów danych, jak wykazy nazw krajów i języków. Do tej grupy zaliczyć można też standardy tworzenia dat itp. Są one stosowane zarówno w MARC 21, jak i w DC, jako część DC Metadata Terms.

Próby ujednolicania opisów w obu opisywanych formatach najłatwiej mogą być realizowane właśnie na poziomie pragmatyki metadanych. Polegają wówczas na uzgodnieniu formy metadanych w odpowiadających sobie miejscach struktury porównywanych schematów. Gorzej, gdy ze względu na różnice semantyczne w standardach metadanych takich odpowiadających sobie miejsc nie można wyznaczyć. Wówczas mogą pojawiać się próby dopasowywania elementów semantyki, co nieuchronnie prowadzi do złamania ogólnej idei stanowiącej podstawę schematu metadanych. W naszych warunkach zazwyczaj dochodzi do „siłowych" zmian w Dublin Core, podczas prób dostosowania go do lepiej znanych (przez bibliotekarzy) i rozpowszechnionych zasad stosowanych w MARC 21. Dochodzi do tworzenia niestandardowych struktur, niezgodnych z ogólnymi zasadami, poprzez krzyżowanie rozwiązań wynikających z realizacji różnych idei opisu dokumentów elektronicznych. Tymczasem wzbogacanie semantyki bez jednoczesnego zadbania o pragmatykę metadanych może czasem tylko zwiększyć zamieszanie w rekordach. Zamiast więc rozbudowy struktury o nowe elementy danych, czasem lepiej jest przygotować dobrą instrukcję wprowadzania ich wartości.

Co charakterystyczne, wszystkie te prace i pomoce nie zapewniają całkowitej jednolitości opisu nawet w pojedynczych bibliotekach, nie mówiąc oczywiście o sieciach bibliotecznych. Jednym z powodów są dość regularne zmiany w przepisach i instrukcjach, wynikające z różnych powodów, m.in. chęci dopasowania sposobu pracy do zmieniającego się środowiska. Czasem udaje się dopasować opisy wstecz do nowych instrukcji, czasem pozostają niezgodne z najnowszymi zasadami. Inny powód jest jeszcze trudniejszy do usunięcia – nawet najbardziej szczegółowe instrukcje zawsze pozostawiają jakiś obszar do interpretacji i samodzielnych rozstrzygnięć. Jest to zresz-

tą niezbędne, bo biblioteki to bardzo zróżnicowane instytucje, o często odmiennych potrzebach. Z tego powodu lokalne różnice w opisach są nie do uniknięcia.

Bibliografia

1. Battelle, John (2006). *Szukaj : jak Google i konkurencja wywołali biznesową i kulturową rewolucję*. Warszawa: PWN, 236 s.
2. Brachfogel, Agnieszka (2010). *Terminy metadanych DCMI – nowe zalecenie dla metadanych Dublin Core*. „Zagadnienia Informacji Naukowej" nr 1(95), s. 57-64.
3. Chopey, Michael (2005). *Planning and implementing ametadata-driven digital repository.* [W:] Smiraglia, Richard red. (2005). *Metadata: a cataloger's primer.* Binghampton: The Haworth Press, s. 255-287.
4. Gawrysiak, Piotr (2008). *Cyfrowe biblioteki a wyszukiwanie informacji.* „Przegląd Biblioteczny" R. 76 z. 4, s. 652-661.
5. Greenberg, Jane (2005). *Understanding metadata and metadata schemes.* [W:] Smiraglia, Richard red. (2005). *Metadata: a cataloger's primer.* Binghampton: The Haworth Press, s. 17-36.
6. IFLA (1998). *Functional requirements for bibliographic records.* München: KG Saur, 144 s.
7. Jackson, Amy [i in.] (2008). *Dublin Core meta data harvested through OAI-PMH.* „Journal of Library Metadata" vol. 8 nr 1, s. 5-21.
8. Lenartowicz, Maria (1989). *Przepisy katalogowania książek. Cz. 1: Opis bibliograficzny.* Warszawa: Wydaw. SBP, 115 s.
9. Nahotko, Marek (2010). *Komunikacja naukowa w środowisku cyfrowym.* Warszawa: Wydaw. SBP, 284 s.
10. Tennant, Roy (2007). *A modest proposal.* [W:] Kresh, Diane red. (2007). *The whole digital library handbook.* Chicago: ALA, s. 387-392.
11. Wolniewicz, Tomasz (2010). *Katalogi wirtualne na przykładzie polskiej wyszukiwarki KaRo.* [W:] Burchard, Maria; Grzędzińska, Kamila; Kasprzyk, Agnieszka red. (2010). *Rola katalogu centralnego NUKAT w kształtowaniu społeczeństwa wiedzy w Polsce.* Warszawa: Wydaw. SBP, s. 153-161.

2. Metadane w polskich bibliotekach cyfrowych

2.1. Dublin Core – modyfikacje schematu, interpretacja atrybutów

Polskie biblioteki cyfrowe w większości budowane są przy użyciu systemu *dLibra*, z domyślnie zaimplementowanym schematem Dublin Core, dlatego ten standard jest najczęściej używanym modelem opisu publikacji cyfrowych w Polsce (co pozostaje w zgodzie z zaleceniami Biblioteki Narodowej sformułowanymi 29 czerwca 2006 r. w Warszawie na II posiedzeniu Zespołu ds. digitalizacji, gdzie „przyjęto uchwałę o rekomendowaniu formatów MARC 21 i Dublin Core jako właściwych dla metadanych obiektów digitalizowanych [...]" [cyt. za Kolasa 2007]). Spośród 75 przeszukiwalnych z poziomu Federacji Bibliotek Cyfrowych repozytoriów 55 korzysta z oprogramowania *dLibra* i schematu Dublin Core (stan na październik 2011 r.). Pozostałe zrzeszone w FBC serwisy działają na oprogramowaniu *EPrints*, *DSpace*, *VTLS/Virtua* lub na własnych systemach. Publikacje udostępniane w tych bibliotekach opisane są w schematach znacznie uproszczonych w stosunku do DC, najczęściej ograniczonych do tytułu, autora, daty wydania i ewentualnie tematu utworu (wyjątkiem jest Akademicka Biblioteka Cyfrowa KRAKÓW, która opisuje zasoby w formacie MARC 21). Na poziomie uogólnienia przyjętym dla niniejszej publikacji metadane tych bibliotek można wyłączyć z dalszych rozważań.

Dublin Core, chociaż rekomendowany i najczęściej stosowany, nie jest wolny od wad. Niedoskonałości schematu, wielokrotnie sygnalizowane i omawiane, wynikają, paradoksalnie, z jego największej zalety – programowej prostoty. Okazuje się bowiem, że piętnastoelementowy zestaw atrybutów[6], przy zróżnicowanych typach zasobów prezentowanych w polskich bibliotekach cyfrowych, jest zbyt ubogi. Świadczą o tym zarówno głosy krytyki podnoszone w wielu publikacjach i wystąpieniach, jak i praktyczne zabiegi stosowane w poszczególnych bibliotekach cyfrowych. Problemy dotyczą wielu aspektów schematu. Zwraca się uwagę m.in. na kłopotliwe zawężenie niektórych pojęć (np. „tytuł", bez podziału na znane z praktyki bibliograficznej dookreślenia, takie jak podtytuły, nadtytuły, tytuły równoległe [Kolasa 2007]), a także kontrowersyjne dla tradycjonalistów pominięcie elementów „istotnych z punktu widzenia identyfikacji wydawnictwa, jak: kolejność wydania, miejsce wydania czy opis fizyczny" [tamże].

Wspomniane braki schematu znakomicie odzwierciedla statystyka modyfikacji podstawowego zestawu elementów. Oprogramowanie *dLibra* umożliwia bowiem ingerencję w oryginalny schemat w bardzo szerokim zakresie – pozwala na zmianę

[6] W *dLibrze* domyślny zestaw poszerzono o pole „Tagi", mające służyć wzbogaceniu opisu proponowanymi przez użytkowników i weryfikowanymi przez redaktorów znacznikami lub słowami kluczowymi [Parkoła 2008], jednak poziom jego użycia jest znikomy (ok. 20% bibliotek).

istniejących nazw atrybutów, dodawanie tzw. „podatrybutów" (swego rodzaju kwalifikatorów, elementów doprecyzowujących atrybut główny) oraz poszerzenie schematu o zupełnie nowe, niezależne elementy. Z 55 zbadanych na potrzeby niniejszego tekstu bibliotek cyfrowych w 45 wyjściowy zestaw atrybutów został zmodyfikowany (pominięto w tym zestawieniu zabiegi tak drobne jak zamiana „Twórcy" na „Autora"), w tym w 28 dodano nowe atrybuty. (Podobne badania przeprowadziły w 2009 r. Joanna Potęga i Agnieszka Wróbel [Potęga; Wróbel 2010], jednak ich przegląd obejmował 37 działających wówczas serwisów, a stale wzrastająca liczba bibliotek cyfrowych zrodziła konieczność powtórzenia badań, co nie tylko pozwoliło na zaktualizowanie obrazu sytuacji, ale także umożliwiło porównanie stanu ówczesnego z obecnym).

Tego typu modyfikacje schematu, choć lokalnie skuteczne (pozwalają doprecyzować opis publikacji według koncepcji przyjętej przez daną bibliotekę cyfrową), na poziomie ponadlokalnym tracą swoją funkcjonalność. W momencie eksportu metadanych poprzez protokół OAI-PMH[7] do serwisu agregującego, jakim jest FBC, a dalej – za jej pośrednictwem – do takich serwisów jak Europeana, DART-Europe czy ViFaOst, część informacji zostaje odrzucona (elementy spoza schematu) lub staje się nieczytelna (metadane zawarte w podatrybutach zostają co prawda przesłane, jednak nie zostaje zachowana etykieta podatrybutu – wpisy zostają eksportowane z rolą przypisaną atrybutowi głównemu).

Krótki przegląd najczęstszych modyfikacji daje obraz tego, co twórcy bibliotek cyfrowych chcieliby widzieć w swoim opisie bibliograficznym (czyli co uważają za informacje ważne), a czego nie daje im podstawowy schemat DCMES. Tabela 1 przedstawia najbardziej charakterystyczne zmiany w obrębie wybranych atrybutów DC oraz najczęściej dodawane nowe atrybuty (polskie nazwy atrybutów podano według wyjściowej nomenklatury w *dLibrze*; cyfry w nawiasach oznaczają liczbę podobnych modyfikacji; dla uproszczenia modyfikacje o podobnym znaczeniu, a różnych nazwach, przedstawiono pod jedną, częściej występującą nazwą, np. „Sygnatura", „Sygnatura oryginału" i „Identyfikator oryginału" jako „Sygnatura").

Tabela 1. Atrybuty i modyfikacje

Nazwa atrybutu DC	Zmieniona nazwa atrybutu	Podatrybuty
Tytuł		Wariant tytułu (4); Inny tytuł (2); Tytuł oryginału (1); Podtytuł (2)
Temat i słowa kluczowe	Hasło przedmiotowe (4)	Hasło przedmiotowe (5)
Opis	Adnotacje (3)	Abstrakt (1); Komentarz (1); Streszczenie (1)
Wydawca		Miejsce wydania (29); Drukarz (1)
Identyfikator zasobu	Identyfikator dokumentu cyfrowego (1)	ISBN ISSN (2); Sygnatura (3); Identyfikator obiektu cyfrowego (1)

[7] Open Archives Initiative Protocol for Metadata Harvesting – protokół służący do „selektywnego pobierania metadanych opisujących obiekty cyfrowe. Umożliwia on automatyczną wymianę informacji pomiędzy systemami gromadzącymi takie metadane a systemami chcącymi te dane uzyskać" [Werla 2008].

Źródło	Sygnatura (6)	Identyfikator dokumentu cyfrowego (1); Sygnatura (1)
Powiązania	Seria /cykl/czasopismo (1)	Seria (1); Katalog (2); Identyfikator obiektu cyfrowego (1)
Prawa	Prawa do dysponowania publikacją (5); Lokalizacja (1); Zarządzanie prawami (1); Status prawny (1)	Licencja (7); Właściciel praw (4); Prawa dostępu (5); Uwagi (1); Treść licencji (2); Digitalizacja (2); Lokalizacja oryginału (3)
Nowe atrybuty		
Tytuł w języku angielskim (2); Wydanie (2); Miejsce wydania (4); Seria (2); Sygnatura (3); Uwagi (6); Lokalizacja źródła (5); Sponsor digitalizacji (4); Digitalizacja (7); Lokalizacja oryginału (5); Identyfikator obiektu cyfrowego (2); Opis fizyczny (1)		

Kwestia indywidualnego poszerzania schematu Dublin Core nie jest jedynym problematycznym zjawiskiem, jakie ujawnia się po głębszym wglądzie w metadane w polskich bibliotekach cyfrowych. Analiza wartości wpisywanych w konkretne pola schematu w poszczególnych serwisach uwidacznia (czasem znaczne) różnice w wykorzystaniu standardu już na poziomie podstawowego zestawu elementów. W wielu wypadkach wiąże się to ze zjawiskiem poprzednim (nazwa atrybutu zostaje zmieniona zgodnie z wizją jego zagospodarowania), ale nawet pola o tych samych nazwach wypełniane są bardzo różnymi wartościami. W tabeli 2 przedstawiono przykłady zróżnicowania interpretacji poszczególnych atrybutów (przy zachowaniu oryginalnego nazewnictwa).

Tabela 2. Atrybuty i interpretacje

Nazwa atrybutu DC	Interpretacja
Temat i słowa kluczowe	słowa kluczowe (32); hasła przedmiotowe (5); hasła przedmiotowe i słowa kluczowe (przemieszane) (11)
Opis	opis fizyczny (22); adres bibliograficzny (np. dla artykułu) (4); swobodny tekst określający treść dokumentu (13); różnego rodzaju uwagi (o brakach, defektach) (22); informacje o serii (2); podtytuły (1)
Identyfikator zasobu	ISBN, ISSN (7); sygnatura (4); nieokreślone numery (2)
Źródło	sygnatura (23); sygnatura obiektu cyfrowego (2); odnośnik do katalogu bibliotecznego z opisem obiektu oryginalnego (9); nieokreślone numery (3); BRAK UŻYCIA ATRYBUTU (14)
Powiązania	serie (17); wersje mikrofilmowe (3); adres bibliograficzny artykułu (3); ISBN, ISSN (2); odnośnik do katalogu bibliotecznego z opisem obiektu oryginalnego (4); inne tytuły (4); odnośniki do innych publikacji (w tej samej bibliotece cyfrowej) (3); BRAK UŻYCIA ATRYBUTU (11)
Zakres	zakresy wiekowe (4); zakresy roczne (4); zakresy terytorialne (2); BRAK UŻYCIA ATRYBUTU (47)
Prawa	„Domena publiczna" (6); uwagi o ograniczeniach w dostępie (lub ich braku) (3); nazwa licencji (5); odnośniki do treści licencji (2); nazwy bibliotek (20); nazwiska (12) ; BRAK UŻYCIA ATRYBUTU (7)

Analiza modyfikacji schematu i różnic w interpretacji oryginalnych atrybutów ukazuje kilka wyraźnych tendencji i ujawnia szereg problemów. Zjawiska te wzajemnie się przeplatają, dlatego naturalne i celowe wydaje się wspólne ich podsumowanie:

- kwestie związane z **tytułem** dokumentu (kwalifikatory dla atrybutu „Tytuł", nowe pole schematu dla tytułu w języku angielskim, umieszczanie podtytułów w atrybucie „Opis") potwierdzają sygnalizowaną już potrzebę precyzowania tego elementu opisu;
- modyfikacje schematu w obrębie atrybutu „**Temat i słowa kluczowe**" to symptomy bardziej złożonego zjawiska, jakim jest relacja pomiędzy funkcjonującymi w katalogach bibliotecznych hasłami przedmiotowymi a stosunkowo nowymi dla bibliotekarzy słowami kluczowymi (więcej informacji na ten temat znajduje się w jednym z kolejnych podrozdziałów). Dość częsta praktyka importowania opisów z formatu MARC do DC właśnie w tym miejscu napotyka pewne trudności. Można zaobserwować kilka rozwiązań. Do najprostszych należy zmiana etykiety atrybutu na „klasyczne" „Hasła przedmiotowe" i pozostawienie zaimportowanych wpisów bez modyfikacji. Chociaż twórcy serwisów, w których stosuje się tę metodę, rezygnują z bardziej funkcjonalnych w środowisku sieciowym słów kluczowych, robią to świadomie i konsekwentnie. Mniej konsekwencji widać natomiast w bibliotekach, gdzie hasła przedmiotowe używane są przy zachowaniu oryginalnej etykiety. Jednak najmniej przemyślana wydaje się (dosyć częsta) praktyka mieszania słów kluczowych i haseł przedmiotowych w jednym atrybucie. Konsekwentne stosowanie słów kluczowych (przy tak zadeklarowanej nazwie atrybutu) można spotkać tylko w nieco ponad połowie serwisów;
- pola, które wydają się najbardziej niejednoznaczne, to „**Identyfikator zasobu**", „**Źródło**" i „**Powiązania**". Różnice w ich interpretacji, nagromadzenie kwalifikatorów i kilka przypadków zmiany nazwy atrybutu wprowadzają do tych elementów schematu duży bałagan. Pola te stają się miejscami do wpisania m.in.: sygnatury dokumentu oryginalnego, numeru ISBN czy ISSN, odnośników do katalogu bibliotecznego lub sygnatury obiektu cyfrowego. Dopełnieniem tego chaotycznego obrazu są przypadki umieszczania powyższych danych poza podstawowym schematem – w niezależnych atrybutach;
- elementem najbardziej pożądanym okazuje się nieobecne w DC „**Miejsce wydania**" (33 biblioteki cyfrowe wprowadziły je jako podatrybut „Wydawcy" lub samodzielny element, przy czym wiele serwisów, które tego nie uczyniło, uwzględnia je w atrybucie „Wydawca", wpisując w nim zawartość podpól $a i $b pola 260 formatu MARC 21, czyli strefy adresu wydawniczego);
- poza miejscem wydania do najczęstszych informacji, jakie twórcy bibliotek chcą przekazać czytelnikowi (a nie umożliwia im tego wyjściowy schemat), należą: miejsce, wykonawca i sponsor digitalizacji;
- atrybutem praktycznie pomijanym przez redaktorów bibliotek cyfrowych jest „**Zakres**", który używany jest w zaledwie 5 repozytoriach;
- najbardziej kontrowersyjnym elementem schematu jest atrybut „**Prawa**". W niemal połowie używających tego pola serwisów podawane są w nim nazwy bibliotek przechowujących oryginał z digitalizowanego dzieła. Tego typu praktyka od

lat wzbudza wiele słusznych protestów. Większość materiałów prezentowanych w polskich bibliotekach cyfrowych to utwory, do których wygasły autorskie prawa majątkowe. Jak się jednak okazuje, publikacje opisane jednoznacznie jako należące do domeny publicznej stanowią zaledwie 10% zasobów cyfrowych dostępnych za pośrednictwem FBC [Werla 2011a]. Wpisanie w polu „Prawa" nazwy instytucji, jaką jest biblioteka, sugeruje, że prawa te są w jakiś sposób zastrzeżone, a ich właścicielem jest właśnie podana instytucja. W większości wypadków jest to nieprawda. Pocieszające może być jednak porównanie referowanych tutaj badań ze wspomnianymi analizami Potęgi i Wróbel – wzrasta częstotliwość użycia pojęcia domeny publicznej i licencji (coraz częściej określanej w podatrybutach), natomiast nazwy bibliotek trafiają tam, gdzie wydają się bardziej na miejscu – do osobnych atrybutów lub podatrybutów o nazwach jednoznacznie mówiących, że chodzi o instytucję przechowującą oryginał.

Wspomniano dotychczas o głównej niedogodności wynikającej z modyfikacji i różnic w interpretacji standardu Dublin Core, jaką jest problem wymiany informacji z innymi serwisami. Podkreślano, że na poziomie lokalnym ingerencja w schemat nie stanowi większego problemu, co więcej, pozwala dostosować istniejące narzędzie do potrzeb i wizji konkretnej biblioteki cyfrowej, wynikających chociażby ze specyfiki prezentowanych materiałów czy indywidualnego poglądu na stopień ważności poszczególnych informacji. Jednak zbyt duża dowolność na tym polu może być dezorientująca także z perspektywy użytkowników. Czytelnik, który korzysta z wielu bibliotek cyfrowych, wiedząc, że używają one jednego standardu, ma prawo spodziewać się podobnych znaczeniowo wartości przynajmniej w polach o podobnej nazwie. Niejednoznaczność w tym zakresie, a co za tym idzie – konieczność „uczenia się" różnych wersji schematu, utrudnia korzystanie z zasobów bibliotek cyfrowych.

Na zakończenie tych rozważań warto zwrócić uwagę na jeszcze jedną kwestię, sygnalizowaną już w poszczególnych uwagach, ale wymagającą wyraźniejszego zaakcentowania w formie jednoznacznych wniosków. Otóż punktem wyjścia do przedstawionej analizy było stwierdzenie o niedoskonałości podstawowego schematu Dublin Core. W tym kontekście przeprowadzone w sposób przemyślany i konsekwentny modyfikacje zestawu atrybutów, nawet jeśli na niektórych poziomach rodzą problemy, jawią się jako działania celowe. Analiza dokonywanych zmian wykazała, że wiele z nich można uznać za w pełni uzasadnione (jako podyktowane chęcią wzbogacenia opisu). Jednak wśród przedstawionych zjawisk nie brakuje takich, które są efektem braku jasnej koncepcji, zwyczajnej niedbałości (np. mieszanie haseł przedmiotowych i słów kluczowych lub wpisy o zupełnie niezrozumiałym znaczeniu, zaznaczone w tabeli 2 jako „nieokreślone numery"[8]) czy braku analizy rozwiązań stosowanych w innych serwisach (wspomniane już podobne znaczeniowo wpisy w różnych polach wyj-

[8] W przedstawionej analizie pominięto zupełnie ewidentne pomyłki redaktorów, takie jak przypadki wpisywania informacji w niewłaściwych polach (sąsiadujących w *dLibrze*). Jest to problem dość częsty, co również świadczy o (braku) staranności i poziomie kontroli własnej pracy (okresowe przeglądanie wpisów w aplikacji „Zarządzanie biblioteką" umożliwia stosunkowo łatwe wychwycenie tego rodzaju omyłek).

ściowego schematu lub w polach dodatkowych). Przyjrzenie się metadanym bibliotek cyfrowych z tej perspektywy może ujawnić, gdzie implementację schematu poprzedził namysł nad jego pożądaną funkcjonalnością, a gdzie schemat wdrożono mechanicznie, bez zastanowienia się nad jego użytecznością w kontekście globalnym.

2.2. Potrzeba i próby standaryzacji

Problemy związane z użyciem DCMES zauważono już dawno i temat metadanych powraca w dyskusjach, na konferencjach, w publikacjach. Istniejące podstawowe wytyczne okazały się zbyt ogólne. Potrzeba wypracowania bardziej precyzyjnych standardów doprowadziła do powstania w 2007 r. *ePoradnika redaktora zasobów cyfrowych* – opracowanej na potrzeby Biblioteki Cyfrowej Uniwersytetu Wrocławskiego „interpretacji schematu Dublin Core wraz z materiałami pomocniczymi" [Domowicz i in. 2008]. Poradnik zawierał instrukcje zarówno na temat danych, jakie powinien zawierać każdy element DC, jak i sposobu ich zapisu. Uwzględniono w nim również odpowiadające każdemu atrybutowi pole schematu MARC. *ePoradnik* został dostrzeżony przez redaktorów innych bibliotek cyfrowych i stał się pomocą w pracy wielu z nich (jest też systematycznie poprawiany i aktualizowany – doczekał się już czterech wydań). Podobne zestawy wytycznych (często oparte na wrocławskim pierwowzorze) powstawały przy innych bibliotekach cyfrowych (np. Zachodniopomorskiej, Dolnośląskiej czy Śląskiej). Tego typu publikacje z pewnością poprawiają jakość metadanych w poszczególnych bibliotekach (zapewniają, przynajmniej w teorii, jednolitość zapisów w jednym serwisie), jednak nie standaryzują w sposób obligatoryjny interpretacji atrybutów dla wszystkich bibliotek, gdyż żadna z instrukcji nie ma rangi dokumentu normalizacyjnego i do takiej rangi nie pretenduje. Mieszczą się one raczej w konwencji „dobrych praktyk" i decyzja o skorzystaniu z ich pomocy leży po stronie twórców poszczególnych bibliotek cyfrowych (o tym, że nie jest to praktyka powszechna, świadczą przedstawione badania).

ePoradnik i inne podobne zestawy zaleceń nie obejmują zagadnienia indywidualnego rozbudowywania schematu DC. Niewystarczalność podstawowego zestawu elementów, potwierdzona analizami schematów wykorzystywanych w bibliotekach cyfrowych, jest jednak faktem i stanowi problem, któremu należy poświęcić należytą uwagę. Kwestią tą zajęto się między innymi podczas warsztatów na temat opracowania zasobów bibliotek cyfrowych zorganizowanych przez Poznańską Fundację Bibliotek Naukowych w dniach 7-8 września 2009 r. w Gnieźnie. Na warsztatach powołano m.in. zespół roboczy do spraw metadanych, który na podstawie analiz wykorzystywanych w polskich bibliotekach cyfrowych schematów, będących odzwierciedleniem potrzeb bibliotekarzy cyfrowych, rozpoczął pracę nad nowym zestawem metadanych. Swoją propozycję (w wersji roboczej) zespół przedstawił w grudniu 2009 r. w Poznaniu na szóstych warsztatach „Biblioteki cyfrowe" [Domowicz i in. 2009]. Zestaw zawierał 15 bazowych elementów poszerzonych o tzw. elementy podrzędne, częścio-

wo oparte na atrybutach DCMI Metadata Terms, a częściowo nowe, specyficzne dla polskich potrzeb.

Na znaczenie standaryzacji zwracają uwagę pracownicy Poznańskiego Centrum Superkomputerowo-Sieciowego – twórcy oprogramowania *dLibra* i operatorzy Federacji Bibliotek Cyfrowych. Podkreślają wpływ normalizacji metadanych na ich interoperacyjność, zwiększającą możliwość automatyzacji wielu procesów związanych ze wspomnianym już eksportem danych do dużych serwisów agregujących [Werla; Dudczak 2009a], a co za tym idzie – skalą widoczności i dostępności polskich zasobów cyfrowych na świecie. Sformułowali też własne zalecenia dotyczące zapisu metadanych, a także ewentualnego indywidualnego poszerzania schematów, do którego również radzili, tam gdzie to możliwe, używać standardowych kwalifikatorów DCMI Metadata Terms [Werla 2010; Werla; Dudczak 2009b]. W drugiej połowie 2010 r. zespół PCSS poszedł o krok dalej i podjął prace nad stworzeniem nowego schematu metadanych. Wyniki prac przedstawiono w styczniu 2011 r. w Warszawie podczas XVII edycji seminarium w cyklu „Digitalizacja – Problemy cyfryzacji dokumentów piśmienniczych w bibliotekach, muzeach i archiwach" [Werla 2011b].

PLMET, nowy schemat metadanych, będący profilem aplikacyjnym Dublin Core, zawiera 59 elementów (17 elementów podstawowych wraz z kwalifikatorami), na które składa się cały zestaw DCMES (15 atrybutów), 33 elementy wybrane ze schematu Dublin Core Metadata Terms (np. Tytuł odmienny, Abstrakt, Prawa dostępu, Licencja, Właściciel praw), 5 atrybutów standardu ETD-MS[9] (Informacje o stopniu naukowym, Uzyskany tytuł, Stopień studiów, Dyscyplina, Instytucja nadająca tytuł) oraz 6 elementów specyficznych dla polskich bibliotek cyfrowych (Tagi użytkowników, Miejsce wydania, Sponsor digitalizacji, Sygnatura, Digitalizacja, Lokalizacja oryginału) [Werla 2011c]. Na wyborze elementów w znacznym stopniu zaważyły – jak wynika ze słów twórców, ale również nawet pobieżnego przejrzenia schematu – wspomniane wyniki pracy zespołu do spraw metadanych oraz analiza przeprowadzona przez Joannę Potęgę i Agnieszkę Wróbel.

Do tej pory z wersji *dLibry* z wbudowanym nowym schematem korzystają dwa serwisy (Dominikańska Biblioteka Cyfrowa i Biblioteka Cyfrowa Politechniki Koszalińskiej), ale ich zasoby są na razie zbyt małe, by dokonywać jakiejkolwiek analizy wykorzystania nowego standardu. Jego implementacja we wszystkich serwisach (szczególnie już działających) jest raczej niemożliwa, dlatego PCSS poszło w innym kierunku – lokalne schematy metadanych są mapowane do PLMET na wyższym poziomie (szczegóły procesu można prześledzić w jednej z zakładek dokumentacji schematu [Werla 2011c]), co umożliwia agregację bardziej szczegółowych metadanych. Korzystanie ze zgromadzonych w ten sposób informacji z poziomu wyszukiwarki Federacji Bibliotek Cyfrowych nie jest jeszcze możliwe, ale prace nad udostępnieniem takiej funkcjonalności trwają. Sam schemat PLMET jest wykorzystywany w większym przedsięwzięciu – projekcie SYNAT[10], którego celem jest „stworzenie

[9] Interoperability Metadata Standard for Electronic Theses and Dissertations – profil aplikacyjny DC używany do opisu prac dyplomowych i dysertacji.

[10] http://www.synat.pl/

uniwersalnej, otwartej, repozytoryjnej platformy hostingowej i komunikacyjnej dla sieciowych zasobów wiedzy dla nauki, edukacji i otwartego społeczeństwa wiedzy". W ramach projektu dane agregowane przez PCSS są – za pośrednictwem PLMET – przetwarzane do postaci CIDOC CRM[11], a w efekcie z metadanych polskich bibliotek cyfrowych powstaje semantyczna baza wiedzy [Mazurek i in. 2012].

Warto jednak wrócić do samego schematu. Ze względu na znikomą liczbę wdrożeń trudno obiektywnie ocenić jego funkcjonalność (choć na forum Biblioteka 2.0[12] pierwsze głosy pojawiły się niemal natychmiast po prezentacjach koncepcji). Jak deklaruje Marcin Werla (w wypowiedzi na wspomnianym forum) „sam schemat ma już postać stabilną" i ewentualne zmiany będą się ograniczały do modyfikacji jego struktury (np. podziału danych na opisowe i administracyjne), ale dokumentacja schematu wciąż jest uzupełniana (co również uniemożliwia ocenę, gdyż zawarte w niej skrócone opisy niezbyt jasno precyzują funkcje poszczególnych pól), a w jej obrębie jest także miejsce na formułowane przez społeczność użytkowników wytyczne.

Wydaje się, że po chwilowym wybuchu entuzjazmu, nadziei na uporządkowanie metadanych polskich bibliotek cyfrowych, sprawa nowego schematu nieco przycichła. Samo porządkowanie ma miejsce, tyle że odbywa się na poziomie globalnym. Nadzieje wiązano jednak również z uporządkowaniem metadanych na poziomie lokalnym, standaryzacją widoczną i użyteczną dla czytelnika poszczególnych serwisów. Stworzenie schematu PLMET, będącego przecież odpowiedzią na zgłaszane przez twórców bibliotek cyfrowych potrzeby, było bez wątpienia krokiem w tym kierunku. Schemat daje bowiem nowe możliwości, szczególnie w kwestii poszerzenia podstawowego zestawu elementów Dublin Core. Proces, który jawi się jako nieunikniony, może być przeprowadzany według pewnego klucza (często przecież indywidualne modyfikacje, choć inaczej nazwane, dotyczą w zasadzie tego samego; bogaty zestaw elementów PLMET daje szeroki wybór potrzebnych atrybutów), a jego efekt może służyć użytkownikom. Być może kolejnym krokiem będzie stworzenie – na wzór wrocławskiego *ePoradnika* – precyzyjnej interpretacji nowego schematu. Mogłyby z niej skorzystać nie tylko nowo powstające serwisy, ale również biblioteki już działające – przynajmniej w zakresie ujednolicenia nazewnictwa i kwestii formalnych związanych z wypełnianiem atrybutów, które zawierają podobne znaczeniowo wpisy. Podczas Ogólnopolskiego Seminarium „Polskie biblioteki cyfrowe – dokąd zmierzamy?", które miało miejsce 9 grudnia 2011 r. w Bibliotece Uniwersyteckiej im. Jerzego Giedroycia w Białymstoku, Tomasz Parkoła zadeklarował, że „w 2012 r. PCSS ogłosi przetarg na opracowanie wytycznych tworzenia metadanych" [Parkoła 2011]. Pozostaje mieć nadzieję, że takie wytyczne zostaną przygotowane i będą konsekwentnie stosowane. To ostatnie będzie już zależało od samych twórców bibliotek cyfrowych.

[11] The CIDOC Conceptual Reference Model (CRM) – międzynarodowy model dokumentacji obiektów dziedzictwa kulturowego (http://www.cidoc-crm.org/).

[12] http://forum.biblioteka20.pl/

2.3. Słowo o typografii

Przy okazji rozważań na temat metadanych polskich bibliotek cyfrowych warto zwrócić uwagę na kwestię typografii – zagadnienie, na które uwagę zwraca się, niestety, zbyt rzadko. Na osobie zarządzającej treścią biblioteki cyfrowej (w tym metadanymi), określanej w terminologii sytemu *dLibra* zobowiązującym mianem redaktora, spoczywa odpowiedzialność za typograficzną poprawność widocznego na stronach serwisu tekstu. Opis publikacji sporządzany w schemacie Dublin Core nie musi (i nie powinien) być okaleczany z powodu pozornych ograniczeń klawiatury komputera. Pamiętając o kilku istotnych zasadach i używając obsługiwanego przez *dLibrę* standardu Unicode, można w prosty sposób ulepszyć opis, czyniąc zadość podstawowym zasadom typograficznym. Kwestia dotyczy zarówno samej treści opisu (poprawne typograficznie znaki interpunkcyjne), jak i sposobu jej wyświetlania.

Za przykład nagminnego błędu niech posłuży często stosowany w opisach bibliograficznych cudzysłów. Natrafienie na jego właściwy zapis – według zasad interpunkcyjnych języka polskiego („ "), francuskiego (« ») czy angielskiego (‘ ’) – należy do rzadkości. Najczęściej jest on zapisany w formie dostępnej bezpośrednio z klawiatury, jako prosty cudzysłów (" ") z zestawu znaków ASCII, niepoprawny z punktu widzenia typografii. Tymczasem wszystkie potrzebne znaki dostępne są w zestawie Unicode i można je wywołać prostą kombinacją klawiszy. Dotyczy to nie tylko cudzysłowów, ale także poprawnych pauz i półpauz (niemal wszędzie zastępowanych klawiaturowym minusem) oraz wielokropków (zapisywanych błędnie jako trzy kropki, podczas gdy w rzeczywistości wielokropek jest jednym znakiem).

Na estetykę wyświetlania tekstu na ekranie wpływa też użycie tzw. twardej spacji, która wstawiona zamiast zwykłej spacji uniemożliwia łamanie tekstu w niepożądanym miejscu (np. po jednoliterowych spójnikach czy przyimkach, po inicjale imienia czy przed kończącym linię skrótem „r.").

Poruszone tutaj sprawy nie są problemami największej wagi, jednak nie zmienia to faktu, że odrobina wysiłku i kilka prostych zabiegów, których można się nauczyć, zaglądając do powszechnie dostępnej literatury [Bringhurst 2008; Wolański 2008] lub korzystając z licznych poradników dostępnych w Internecie[13], pozwoli poprawić estetykę opisu, by godnie reprezentował instytucję, która go przygotowała.

Bibliografia

1. Bringhurst, Robert (2008). *Elementarz stylu w typografii*. Kraków: D2D.PL, 461 s.
2. Domowicz, Iwona [i in.] oprac. (2008). *ePoradnik redaktora zasobów cyfrowych* [dok. elektr.] Wrocław: Biblioteka Uniwersytecka we Wrocławiu. http://www.bibliotekacyfrowa.pl/dlibra/docmetadata?id=25069&showContent=true [odczyt: 2.10.2011].
3. Domowicz, Iwona [i in.] oprac. (2009). *Propozycja nowego schematu metadanych (wersja robocza). Wyniki pierwszego etapu prac nad schematem metadanych* [dok. elektr.]

[13] Np. http://taat.pl/typografia/

http://dl.psnc.pl/biblioteka/dlibra/docmetadata?id=233&from=&dirids=1&ver_id=&lp=1&QI= [odczyt: 5.10.2011].

4. Kolasa, Władysław Marek (2007). *„dLibra" Digital Library Framework – platforma do budowy bibliotek cyfrowych* [dok. elektr.] [W:] Woźniak-Kasperek, Jadwiga; Franke, Jerzy, red. *Biblioteki cyfrowe: projekty, realizacje, technologie*. Warszawa: Wydawnictwo SBP. http://bbc.uw.edu.pl/dlibra/docmetadata?id=4&dirds=1&tab=1 [odczyt: 12.10.2011].
5. Mazurek, Cezary [i in.] (2012). *Transforming a Flat Metadata Schema to a Semantic Web Ontology: The Polish Digital Libraries Federation and CIDOC CRM Case Study* [W:] Bembenik, Robert [i in.] red. *Intelligent Tools for Building a Scientific Information Platform (Studies in Computational Intelligence)*. New York: Springer. s. 153-177.
6. Parkoła, Tomasz (2008). *Podręcznik użytkownika środowiska dLibra (wersja 4.0)* [dok. elektr.] Poznań: PCSS. http://dlibra.psnc.pl/biblioteka/dlibra/docmetadata?id=133&from=publication [odczyt: 2.10.2011].
7. Parkoła, Tomasz (2011). *Rozwój polskich bibliotek cyfrowych* [dok. elektr.] Poznań: PCSS. http://bg.uwb.edu.pl/seminarium2011/materialy/2011-12-09-Bialystok-Parkola.pdf [odczyt: 5.10.2011].
8. Potęga, Joanna; Wróbel, Agnieszka (2010). *The Dublin Core Metadata Element Set, Ver. 1.1 a potrzeby i oczekiwania bibliotekarzy cyfrowych – analiza przypadków.* [W:] Mazurek, Cezary; Stroiński, Maciej; Węglarz, Jan red. *Polskie biblioteki cyfrowe 2009. Materiały z konferencji zorganizowanej dnia 9 grudnia 2009 roku przez: Bibliotekę Kórnicką PAN, Poznańską Fundację Bibliotek Naukowych, Poznańskie Centrum Superkomputerowo-Sieciowe*. Poznań: Ośrodek Wydawnictw Naukowych, s. 71-78.
9. Werla, Marcin (2008). *Biblioteka cyfrowa jako repozytorium OAI-PMH* [dok. elektr.] Poznań: PCSS. http://confluence.man.poznan.pl/community/display/KB/Biblioteka+cyfrowa+jako+repozytorium+OAI-PMH [odczyt: 2.10.2011].
10. Werla, Marcin (2010). *How to describe digital objects?* [dok. elektr.] Poznań: PCSS. http://dl.psnc.pl/biblioteka/dlibra/docmetadata?id=245&from=&dirids=1&ver_id=&lp=1&QI= [odczyt: 3.10.2011].
11. Werla, Marcin (2011a), *Domena publiczna i licencje Creative Commons w polskich bibliotekach cyfrowych* [dok. elektr.] Poznań: PCSS. http://dl.psnc.pl/2011/03/10/domena-publiczna-i-licencje-creative-commons-w-polskich-bibliotekach-cyfrowych/ [odczyt: 1.10.2011].
12. Werla, Marcin (2011b). *Metadane dokumentów w bibliotekach cyfrowych* [dok. elektr.] Poznań: PCSS. http://dlibra.psnc.pl/biblioteka/dlibra/docmetadata?id=284&from=&dirids=1&ver_id=&lp=1&QI= [odczyt: 3.10.2011].
13. Werla, Marcin (2011c). *Dokumentacja schematu metadanych PLMET* [dok. elektr.] Poznań: PCSS. http://confluence.man.poznan.pl/community/display/FBCMETGUIDE/Dokumentacja+schematu+metadanych+PLMET [odczyt: 7.10.2011].
14. Werla, Marcin; Dudczak, Adam (2009a). *ECDL 2009 Tutorial: Aggregation and Reuse of Digital Objects' Metadata from Distributed Digital Libraries* [dok. elektr.] Poznań: PCSS. http://dlibra.psnc.pl/biblioteka/dlibra/publication?id=278&tab=3 [odczyt: 3.10.2011].
15. Werla, Marcin; Dudczak, Adam (2009b), *Warsztaty nt. opracowania zasobów bibliotek cyfrowych* [dok. elektr.] Poznań: PCSS. http://dl.psnc.pl/biblioteka/dlibra/docmetadata?id=191&from=&dirids=1&ver_id=&lp=1&QI= [odczyt: 1.10.2011].
16. Wolański, Adam (2008). *Edycja tekstów : praktyczny poradnik : książka, prasa, www*. Warszawa: Wydawnictwo Naukowe PWN, s. 380.

3. Odwzorowanie treści dokumentów elektronicznych – Marek Nahotko

3.1. Sposoby odwzorowania treści

Z punktu widzenia systemu biblioteki cyfrowej (a więc pomijając specyfikę poszczególnych jiw), sposób odwzorowania treści dokumentów cyfrowych w tych bibliotekach zależy od podobnych czynników, co sposoby rejestracji opisu formalnego, przedstawione w poprzednich rozdziałach. Należą do nich: przyjęty schemat struktury metadanych (elementy semantyki), możliwości systemu informatycznego, obsługującego bibliotekę oraz czynnik najmniej zależny od struktur metadanych – rodzaj opisywanych zbiorów i tradycje w tym zakresie.

Wbrew pozorom wybór schematu metadanych może wpływać na sposób realizacji opracowania rzeczowego. W MARC 21 dla wprowadzania wyrażeń każdego rodzaju języków informacyjno-wyszukiwawczych wykorzystuje się inne grupy pól – dla języków przedmiotowych (tezaurusów) są to w rekordzie bibliograficznym pola 6XX (są to pola kontrolowane khw), dla klasyfikacji część pól 0XX, słowa kluczowe umieszcza się przeważnie w wybranych polach 9XX, przeznaczonych do lokalnego wykorzystania, a abstrakty i spisy treści w polu uwag 520 (adnotacja wyjaśniająca lub analiza wskazująca).

Inny sposób przedstawiania treści obiektu cyfrowego przewidziany został w Dublin Core. Istnieją tam trzy atrybuty służące temu celowi. W atrybucie Temat umieszcza się wyrażenia języków informacyjno wyszukiwawczych, a więc nie wyrażenia języka naturalnego. Nie tworzono żadnych ograniczeń co do rodzaju jiw, więc mogą to być zarówno hasła przedmiotowe, deskryptory, słowa kluczowe jak również symbole klasyfikacji. Atrybut Temat należy powtarzać dla każdej kolejnej użytej jednostki leksykalnej. W przypadku stosowania kliku jiw w jednym systemie należy zadbać, aby jednostki różnych języków były od siebie odseparowane (nie były wymieszane). Nieco myląca jest nazwa przyjęta w polskiej normie dla tego atrybutu. Wyrażenie 'Temat' w żadnym stopniu nie oddaje właściwej treści dla niego przeznaczonej i może wprowadzać w błąd bibliotekarzy, którzy mają prawo pytać o jakie tematy chodzi: czy o te tematy będące częścią haseł przedmiotowych? W takim razie gdzie umieszczać określniki? atrybut Opis służy do umieszczania uwag o treści w języku naturalnym, a więc tu jest miejsce między innymi dla abstraktów, streszczeń i spisów treści. Nie jest to natomiast pole do uwag wszelkiego rodzaju, mimo że niektóre biblioteki tak go traktują. Istnieje także atrybut Zasięg, który również powoduje trudności interpretacyjne, poczynając od jego nazwy, gdyż przy jego pomocy przedstawia się nie zasięg, a zakres

chronologiczny i geograficzny. Atrybut ten służy do zapisu miejsca (lokalizacji) i czasu będących elementami treści dokumentu, nie chodzi więc tu o miejsce ani o datę wydania opisywanego źródła. Atrybut ten może być przydatny np. do opisu zdjęć satelitarnych powierzchni Ziemi.

W pewnych sytuacjach można uznać za opisujące treść także inne atrybuty DC, takie jak Odbiorcy, Rodzaj, Format, Nośnik [Brachfogel 2010, s. 72].

Sieciowe środowisko dokumentów elektronicznych daje także nowe możliwości opisu treści, niewymagające pracy bibliotekarza. Chodzi o indeksowanie pełnotekstowe, zarówno na poziomie biblioteki cyfrowej, jak również wyszukiwarek globalnych.

Możliwości indeksowania pełnotekstowego związane są z czynnikiem wydawałoby się odległym od treści dokumentu, mianowicie z jego formatem. Brak jest technologii pozwalających na efektywne indeksowanie treści obiektów graficznych, a obiektem takim może być także zeskanowany tekst. W takim przypadku metadane są jedynym źródłem danych do indeksów wyszukiwarek. Teksty dostępne do indeksowania mogą być wykorzystywane w tym celu zarówno przez lokalne narzędzia na poziomie poszczególnych bibliotek cyfrowych (lub ich grup) i/lub wyszukiwarki globalne. W pierwszym przypadku biblioteka cyfrowa udostępnia użytkownikom możliwość wyszukiwania pełnotekstowego swoich zasobów.

Jednocześnie wszelkie teksty, udostępniane w bibliotece cyfrowej, powinny być udostępniane do indeksowania przez globalne wyszukiwarki. Dotyczy to zarówno metadanych, jak i danych, których są surogatem. Dzięki temu użytkownik ma szansę często odkryć w ogóle istnienie biblioteki cyfrowej, do której jest kierowany z Google. Jak wskazują najnowsze badania, 84% Amerykanów rozpoczyna poszukiwanie informacji od wyszukiwarki internetowej [Rosa i in. 2010, s. 32]. Jeżeli więc biblioteka cyfrowa i jej zasoby nie są widoczne w wynikach wyszukiwania przedstawianych przez wyszukiwarkę na dobrym miejscu (praktycznie na pierwszym ekranie), to zasoby te praktycznie są niedostępne. Nie powinno być technicznych trudności w udostępnianiu metadanych DC do indeksowania przez wyszukiwarki internetowe. Niestety, metadane MARC 21 w ISO 2709 znajdują się w ukrytym Webie, niedostępne do indeksowania. Stąd próby zastosowania XML do kodowania rekordów MARC, o których była mowa wcześniej.

Obok metod indeksowania i jiw stosowanych także w bibliotekach tradycyjnych, w bibliotekach cyfrowych wykorzystywane są także nowe, silniejsze narzędzia i sposoby opisu treści, wykorzystujące metody sztucznej inteligencji. Systemy te można przedstawić w trzech grupach:

1. Listy terminów, z kontrolowaną formą leksykalną lub synonimią, takie jak khw, glosariusze, wykazy nazw geograficznych i in.;
2. Systemy hierarchiczne, w tym jhp, klasyfikacje biblioteczno-bibliograficzne i fasetowe oraz taksonomie;
3. Systemy relacyjne, takie jak języki deskryptorowe, sieci semantyczne, ontologie (modele domen wiedzy) [Sosińska-Kalata 2008, s. 19].

Biblioteki cyfrowe mogą zaliczać się także do serwisów Webu 2.0, albo przynajmniej stosować jego narzędzia, które opierają się na współtworzeniu i współdzieleniu

treści cyfrowych przez użytkowników. Organizują oni zasoby cyfrowe za pomocą tzw. etykiet (tagów), swobodnie dobieranych wyrażeń. Zbiory etykiet często prezentowane są w postaci tzw. chmur tagów, w których wielkością czcionki określa się popularność wyrażenia. Na podstawie niekontrolowanego indeksowania przez „tłumy" internautów powstają t.zw. klasyfikacje społecznościowe, określane jako folksonomie. Stanowią one ciekawe uzupełnienie tradycyjnych jiw, stosowanych w bibliotekach cyfrowych.

Bibliografia

1. Brachfogel, Agnieszka (2010). *Terminy metadanych DCMI i możliwości ich wykorzystania w opisie rzeczowym*. „Zagadnienia Informacji Naukowej" nr 2(96), s. 69-76.
2. Rosa, Cathy [i in.] (2010). *Perceptions of libraries, 2010. Context and community*. Dublin, Ohio: OCLC, 108 s.
3. Sosińska-Kalata, Barbara (2008). *Kształtowanie się nowego paradygmatu organizacji zasobów wiedzy w społeczeństwie sieciowym?* [W:] Pietruch-Reizes, Diana; Babik, Wiesław red. (2008). *Wymiana informacji i rozwój profesjonalnych usług informacyjnych w edukacji, nauce i kulturze na rzecz społeczeństwa opartego na wiedzy*. Katowice: PTIN, s. 17-23.

3.2. Opracowanie rzeczowe dokumentu –

Danuta Patkaniowska

Celem opracowania rzeczowego jest przedstawienie zbiorów bibliotecznych od strony ich zawartości treściowej.

Opracowanie rzeczowe uwzględnia wszechstronność albo stopień wyspecjalizowania biblioteki. W zależności od rodzaju bibliotek, a także kryteriów wyszukiwania dokumentów przez użytkowników, różny może być stopień szczegółowości i głębokości opracowania rzeczowego.

W opracowaniu rzeczowym uwzględnia się układ tekstu książki, tematy i problemy pierwszoplanowe, równorzędne i poboczne, a także hierarchię i powiązania między nimi oraz cechy formalne, stylistykę, poziom, przeznaczenie dokumentu.

Opis rzeczowy zapisany w wybranym języku informacyjno-wyszukiwawczym tworzy charakterystykę wyszukiwawczą dokumentu. Zbiór charakterystyk wyszukiwawczych jest zebrany w katalogu rzeczowym lub komputerowej bazie danych [Stopa 2002].

Informacja rzeczowa książki elektronicznej ma swoją specyfikę i niejednokrotnie – tak w procesie opracowania, jak i końcowego efektu – różni się od informacji rzeczowej dotyczącej książki tradycyjnej.

3.2.1. Rodzaje języków informacyjno-wyszukiwawczych.

Różne języki informacyjno-wyszukiwawcze są sztucznymi systemami językowymi przeznaczonymi do opisu treści i formy dokumentu, zarówno tradycyjnego, jak i elektronicznego.

Klasyfikacja piśmiennictwa według dziedzin wiedzy stanowi najstarszy sposób budowy rzeczowej informacji bibliotecznej. Istotne są tu relacje hierarchiczne, „logiczny porządek" w grupowaniu piśmiennictwa według określonych działów wiedzy i klas tematycznych. Klasyfikacje systematyczne [Sosińska-Kalata 2002; Gawrysiak 2009], np. Klasyfikacja Deweya, Klasyfikacja Biblioteki Kongresu, Klasyfikacja Dziesiętna, szeroko wykorzystywane w tradycyjnych kartkowych katalogach bibliotecznych i w katalogach księgozbiorów podręcznych, znalazły również zastosowanie przy tworzeniu księgozbiorów w tzw. wolnym dostępie, w katalogach elektronicznych bibliotek (zwłaszcza technicznych) oraz w zarządzaniu informacją w Internecie.

Opracowanie dokumentów w językach haseł przedmiotowych [Ćwiekowa 1988; Sadowska 1991, 2000, 2003; Woźniak-Kasperek, Bierczyński 2006; Głowacka 2000, 2003] polega na wyodrębnieniu z dokumentu i opisaniu przedmiotów (tematów) głównych, pobocznych czy równorzędnych, a także określeniu ich własności, aspektów i ujęć w oparciu o reguły gramatyki pozycyjnej. Opis przedmiotowy składa się z haseł przedmiotowych prostych (tematów) oraz haseł przedmiotowych rozwiniętych, w których obok tematu znajdują się określniki: rzeczowe, chronologiczne, geograficzne, formy. Tematy i określniki tworzą charakterystyki wyszukiwawcze dokumentów i są szeregowane alfabetycznie.

W polskich elektronicznych katalogach bibliotecznych znalazły zastosowanie: język haseł przedmiotowych Biblioteki Narodowej (Biblioteka Narodowa, biblioteki publiczne, biblioteki pedagogiczne, biblioteki szkolne i niektóre biblioteki akademickie) oraz tworzony od 1993 r. w oparciu o kartotekę francuską RAMEAU i amerykański LCSH język haseł przedmiotowych KABA (większość bibliotek akademickich).

W bibliotekach cyfrowych stosuje się najczęściej języki słów kluczowych [Babik 2010], czyli słowa i wyrażenia dowolnie określające treść dokumentów.

Języki słów kluczowych to języki o notacji paranaturalnej, bez wyraźnie określonej w słowniku paradygmatyki, w których jednostki słownika nazywane są słowami kluczowymi. Posiadają gramatykę niepozycyjną tj. dowolny szyk jednostek leksykalnych w zdaniu. Cechuje je indeksowanie współrzędne.

Słowem kluczowym jest każde wyrażenie lub wyraz wybrane z tytułu lub tekstu dokumentu, wykorzystane do reprezentowania treści lub dowolne wyrażenie z pytania informacyjnego. Słowa kluczowe mogą także zawierać dopowiedzenia lub przybierać formę fraz, składających się z kilku wyrazów. Są szeroko stosowane przy wyszukiwaniu informacji w Internecie.

Języki deskryptorowe zawierają słownictwo specjalistyczne z poszczególnych dziedzin, zgromadzone w tezaurusach [Ścibor, Tomasik-Beck 1995; Woźniak-Kasperek 2005]. Są językami o notacji paranaturalnej, z określoną paradygmatyką, w których

jednostki słownika są nazywane *deskryptorami,* a gramatykę stanowią najczęściej reguły indeksowania współrzędnego.

Tezaurus stanowi normatywny, zamknięty w danym momencie wykaz deskryptorów (elementarnych jednostek leksykalnych używanych do indeksowania) oraz askryptorów (wyrażeń nieużywanych, przeważnie synonimów). Istnieją relacje hierarchiczne i kojarzeniowe między deskryptorami.

Tezaurusy są wykorzystywane w Internecie do automatycznego indeksowania zasobów stron WWW, analizy współwystępowania słów, wyszukiwania z wykorzystaniem relacji kojarzeniowych.

Innymi językami informacyjno-wyszukiwawczymi są języki kodów semantycznych i języki syntagmatyczne, które zawierają kody (symbole) literowe lub cyfrowe używane w celu oznaczenia pojęć ogólnych.

Do języków informacyjno-wyszukiwawczych należą też języki opisu formalnego czyli języki opisu bibliograficznego np. stosowane w bibliografiach oraz języki cytowań bibliograficznych.

3.2.2. Języki informacyjno-wyszukiwawcze stosowane w polskich bibliotekach cyfrowych. Tagowanie

W bibliotekach cyfrowych stosuje się grupowanie dokumentów według dziedzin wiedzy. Przy poszczególnych dokumentach często spotyka się określenie tzw. Subject Area czyli dziedziny, do której należy dokument. Typologia materiałów cyfrowych – zróżnicowana w zależności od rodzaju i wielkości biblioteki – jest oparta najczęściej na podziale dziedzinowym, podziale według formy/typu dokumentu oraz według rodzaju zasobu.

Cyfrowa Biblioteka Narodowa Polona posiada podział formalny oraz podział według kolekcji tematycznych. Podobny podział mają biblioteki cyfrowe oparte na oprogramowaniu dLibra zgrupowane w Federacji Bibliotek Cyfrowych. Widoczny jest tu podział formalny na czasopisma (często są wyodrębnione czasopisma regionalne) oraz książki, w tym na: dziedzictwo kulturowe (rękopisy, stare druki, zbiory muzyczne, zbiory kartograficzne, zbiory ikonograficzne, dokumenty życia społecznego, książki 1800-1945) oraz publikacje współczesne, materiały dydaktyczne oraz kolekcje specjalne. Przykładami kolekcji specjalnych są: Prawnicza i Ekonomiczna Biblioteka Cyfrowa (Biblioteka Cyfrowa Uniwersytetu Wrocławskiego), Historia UW i BUW oraz księgozbiór Warszawskiego Towarzystwa Przyjaciół Nauk (e-biblioteka Uniwersytetu Warszawskiego), Historia informatyki oraz Historia nauki i techniki (Biblioteka Cyfrowa Politechniki Warszawskiej oraz Politechniki Lubelskiej) i in.

Czasami podział formalny biblioteki cyfrowej jest bardzo ogólny na: Książki oraz Czasopisma np. w Bibliologicznej Bibliotece Cyfrowej.

W Jagiellońskiej Bibliotece Cyfrowej istnieje następujący podział formalny: czasopisma, czytelnia naukowa, czytelnia naukowa UJ, dokumenty życia społecznego, książki (od 1800 r.), muzykalia, rękopisy, stare druki, zbiory ikonograficzne, zbiory kartograficzne. W odniesieniu do książek wyd. od 1800 r. do czasów współczesnych

zastosowano klasyfikację systematyczną opartą na tradycyjnym Katalogu Systematycznym BJ.

W niektórych bibliotekach cyfrowych znalazły zastosowanie języki haseł przedmiotowych. Język haseł przedmiotowych Biblioteki Narodowej (jhp BN) zastosowano w Cyfrowej Bibliotece Narodowej POLONA, Bałtyckiej Bibliotece Cyfrowej, Jeleniogórskiej Bibliotece Cyfrowej, Śląskiej Bibliotece Cyfrowej i in. Język haseł przedmiotowych KABA (jhp KABA) znalazł zastosowanie w Polskiej Bibliotece Internetowej, e-bibliotece Uniwersytetu Warszawskiego, Bibliotece Cyfrowej Politechniki Łódzkiej i in. Słowa kluczowe oparte na języku haseł przedmiotowych KABA posiada Jagiellońska Biblioteka Cyfrowa. W e-bibliotece Uniwersytetu Warszawskiego oraz w krakowskiej Akademickiej Bibliotece Cyfrowej istnieje możliwość osobnego wyszukiwania dokumentów według haseł jhp KABA i według słów kluczowych.

Charakterystyczny dla bibliotek cyfrowych jest brak precyzyjnej terminologii metadanych. Wyświetlające się na ekranie „słowa kluczowe" w niektórych bibliotekach cyfrowych opartych na oprogramowaniu dLibra to często nie „słowa kluczowe" sensu stricto, ale słowa i wyrażenia występujące w językach haseł przedmiotowych. W Śląskiej Bibliotece Cyfrowej obok dowolnych słów kluczowych występują tematy jhp BN. W Dolnośląskiej Bibliotece Cyfrowej, w Bibliotece Cyfrowej Politechniki Lubelskiej, w Bibliotece Cyfrowej Uniwersytetu Marii Curie-Skłodowskiej i in. znajdujemy słowa kluczowe wraz z częściowo zmodyfikowanymi i skróconymi hasłami jhp KABA.

Język deskryptorowy oparty na tezaurusie znalazł zastosowanie w Bibliotece Cyfrowej Politechniki Krakowskiej. Innym językiem deskryptorowym, zawierającym polską terminologię medyczną jest język MeSH (wersja polska), tworzony przez biblioteki medyczne.

Języki słów kluczowych występują w większości bibliotek cyfrowych polskich bibliotek akademickich np. Wielkopolskiej Bibliotece Cyfrowej, Bibliotece Cyfrowej Politechniki Warszawskiej, Bibliotece Cyfrowej Uniwersytetu Łódzkiego, a także w Księgozbiorze Wirtualnym FIDES oraz w Pedagogicznej Bibliotece Cyfrowej.

Biblioteki cyfrowe posiadają, liczące od kilkuset do kilkuset tysięcy terminów, indeksy słów kluczowych, ułożone w porządku alfabetycznym. Indeksy te pokazują, jak często słownictwo słów kluczowych przybiera formę słownictwa niekontrolowanego, a zapis terminologii jest niejednolity i często dowolny. Niektóre niekonsekwencje w indeksach słów kluczowych to: stosowanie synonimów w odniesieniu do dokumentów o podobnej treści oraz niejednolity zapis terminów np. *bitwa pod Grunwaldem* oraz *Grunwald 1410;* zapis tych samych terminów dużą i małą literą; niekonsekwencja w tworzeniu nazw osobowych i geograficznych (nazwy z dopowiedzeniami obok nazw bez dopowiedzeń; zapis nazw osobowych zaczynających się od imienia obok zapisu nazw osobowych zaczynających się od nazwiska); zapis dat cyframi arabskimi obok zapisu dat cyframi rzymskimi i in.

W bibliotekach cyfrowych w ujęciu alfabetycznym występują kategorie tematyczne, które są zarówno nazwami dziedzin wiedzy, jak i nazwami działów gospodarki, kategorii osób i nazwami pojedynczych wydarzeń, osób, tytułów, instytucji i organizacji, przedmiotów, pojęć abstrakcyjnych i in. Większość tematów to tematy

jednowyrazowe, rzadziej – wielowyrazowe. W indeksach słów kluczowych obok pojedynczych terminów istnieją terminy z dopowiedzeniami, określnikami rzeczowymi, chronologicznymi, geograficznymi i określnikami formy.

W katalogu Skarby Dziedzictwa Narodowego katalog tematyczny jest wzbogacany o nowe hasła przedmiotowe, które w bazie określono jako 'tematy jednostkowe' oraz 'tematy łączone'. W katalogu tym istnieje jakaś forma kontroli słownictwa (np. istnienie odsyłaczy całkowitych), niemniej jednak trudno do końca określić, czy są to hasła przedmiotowe, czy raczej słowa kluczowe.

Czasami obok indeksu słów kluczowych w języku polskim istnieje indeks słów kluczowych w języku angielskim np. w bazie BazTech. Czasami terminologia angielska występuje w jednym indeksie z terminologią polską np. w Wielkopolskiej Bibliotece Cyfrowej. Z kolei w Bibliotece Wirtualnej Matematyki oraz w Bibliotece Cyfrowej Katedry Lingwistyki Formalnej UW istnieją wyłącznie słowa kluczowe w języku angielskim.

W poznańskiej Akademickiej Bibliotece Internetowej występuje indeks tematyczny składający się z następujących dziedzin: Chemia; Etyka informatyczna; Film i Telewizja; Filologia polska; Filozofia; Historia sztuki; Media cyfrowe; Politologia; Software Engineering. Biblioteka ta nie posiada słów kluczowych, a możliwość wyszukiwania ogranicza się do wyszukiwania wyrażeń w tytułach i nazwach autorów.

W Bibliotece Cyfrowej Politechniki Krakowskiej nie ma indeksu słów kluczowych, natomiast istnieje podstawowe i rozszerzone (we wszystkich polach opisu) wyszukiwanie zasobów, także według słów kluczowych opartych na tezaurusie. Niektóre biblioteki cyfrowe np. Biblioteka Cyfrowa Politechniki Śląskiej, Cieszyńska Biblioteka Wirtualna, Cyfrowa Biblioteka Diecezjalna w Sandomierzu w ogóle nie posiadają opracowania rzeczowego.

Wirtualna Biblioteka Literatury Polskiej nie posiada słów kluczowych, ale zawiera inne elementy opisu rzeczowego: noty biograficzne, tekstowe noty o utworach, ilustracje, szczegółowe opisy bibliograficzne.

W ramach polityki współtworzenia opisów rzeczowych przez użytkowników bibliotek cyfrowych, wiele z nich oferuje tagowanie czyli możliwość dodania tagów (dowolnych słów kluczowych). Tagowanie czy etykietowanie dokumentów powoduje powstanie folksonomii, niekontrolowanej struktury informacji zbudowanej przez nieprofesjonalistów [Woźniak-Kasperek 2008; Babik 2010, s. 183-186]. Gdyby rozwiązanie to miało wzbogacać istniejący w bibliotece cyfrowej ustrukturyzowany jiw to wprowadziłoby do niego mnóstwo zamieszania i przyniosło więcej szkody niż pożytku, natomiast tagowanie dobrze sprawdza się przy zastosowaniu w repozytoriach cyfrowych języka swobodnych słów kluczowych. Podobnie jest z jakością informacji zawartej w tagach – może nieść ze sobą bardzo celne, stworzone przez specjalistów w danej dziedzinie określenia, pozwalające na uszczegółowienie zawartości treściowej publikacji, lecz z drugiej strony mogą okazać się bezwartościowymi hasłami zwiększającymi szum informacyjny.

3.2.3. Jhp KABA a swobodne słowa kluczowe – możliwości translacji.

Język haseł przedmiotowych KABA jest językiem informacyjno-wyszukiwawczym, współtworzonym od 1993 r. przez polskie biblioteki akademickie i stosowanym w katalogu NUKAT oraz katalogach bibliotek naukowych.

Do ważnych cech jhp KABA należą: nieograniczony zasób słownictwa, liczącego obecnie ponad 480 tys. terminów, kompatybilność z RAMEAU i LCSH, szczegółowość i sformalizowanie języka oraz kontrola słownictwa poprzez kartotekę haseł wzorcowych.

Jhp KABA składa się z tematów (wyrażeń autosyntaktycznych) oraz określników (wyrażeń synsyntaktycznych). Tematami jhp KABA są: osoby (też fikcyjne), dzieła autorskie i anonimowe, instytucje i organizacje, zdarzenia historyczne, jednostki geograficzne i fizjograficzne, kategorie osób oraz grupy etniczne i narodowościowe dyscypliny naukowe i dziedziny wiedzy, kierunki i style literackie i artystyczne, obiekty fizyczne, przedmioty, materiały; działania i procesy, jednostki abstrakcyjne, własności przedmiotów; zjawiska, stany, relacje. Niektóre tematy występują z dopowiedzeniami (słownymi lub cyfrowymi), niektóre przyjmują postać wyrażeń wieloelementowych.

Własności, aspekt i ujęcie przedmiotów są określane poprzez określniki: rzeczowe, chronologiczne (swobodne i związane), geograficzne, formy.

Jhp KABA jest językiem sformalizowanym; obowiązują szczegółowe zasady stosowania tematów w poszczególnych dziedzinach, zasada łączliwości tematów z poszczególnymi określnikami geograficznymi i chronologicznymi oraz zasada koordynacji czyli kolejności stosowania określników po tematach.

Tematy w jhp KABA posiadają relacje ekwiwalencji (terminy odrzucone) oraz relacje hierarchiczne i skojarzeniowe (terminy nadrzędne, terminy podrzędne, terminy skojarzone), a także system odsyłaczy, noty stosowania, czasami definicje terminu i uwagi.

Jako przykład translacji ze sformalizowanego jiw jakim jest jhp KABA na język swobodnych słów kluczowych może posłużyć próba dokonana w Bibliotece Jagiellońskiej dla potrzeb Jagiellońskiej Biblioteki Cyfrowej (JBC).

Podjęto się tego zadania ze względu na następujące okoliczności:

1. translacja danych z Komputerowego Katalogu Zbiorów Bibliotek UJ do oprogramowania Dublin Core nie obejmuje kartotek haseł wzorcowych i indeksu jhp KABA, które umożliwiają wyszukiwanie tematyczne;
2. badania użytkowników Internetu oraz części jego zasobów jakim są biblioteki cyfrowe pokazują, iż preferują oni wyszukiwanie dokumentów za pomocą słów kluczowych.

Jest zrozumiałe, iż przekład tak sformalizowanego języka, jakim jest jhp KABA, na swobodne słowa kluczowe, jest zadaniem trudnym i wymagającym dookreślenia wielu szczegółów przy tworzeniu tablicy przekładu.

Sformułowano następujące zasady tworzenia słów kluczowych:

1. Słowa kluczowe są oparte na terminologii jhp KABA; dotyczy to zwłaszcza tematów (nazwy osobowe, tytuły, nazwy geograficzne, nazwy korporatywne,

nazwy pospolite). W większości hasła przedmiotowe proste są przejmowane w tej samej formie, w jakiej występują w jhp KABA.

2. Tematy – nazwy osobowe, tytuły, nazwy korporatywne, nazwy geograficzne, nazwy pospolite są przejmowane jako słowa kluczowe w formie występującej w jhp KABA (wraz z dopowiedzeniami), jeśli forma ta jest zrozumiała dla użytkownika.
3. W przypadku skomplikowanych lub obcojęzycznych terminów istnieje możliwość stosowania ich polskich odpowiedników i tworzenia słów kluczowych w oparciu o terminy odrzucone (odsyłacze) w jhp KABA.
4. Określniki rzeczowe, określniki geograficzne i określniki formy mogą być stosowane jako dodatkowe (obok tematów) słowa kluczowe.
5. Hasła przedmiotowe rozwinięte w jhp KABA, zbudowane z tematu i określnika rzeczowego, przejmuje się w większości w niezmienionej formie; funkcję separatora między tematem i określnikiem w słowie kluczowym tworzy przecinek.
6. W przypadku skomplikowanych lub niezrozumiałych haseł przedmiotowych rozwiniętych w jhp KABA słowa kluczowe mogą być budowane w oparciu o frazy tworzone z terminów przyjętych lub odrzuconych w jhp KABA.
7. Jeśli w haśle przedmiotowym rozwiniętym występuje określnik chronologiczny, słowo kluczowe tworzy się łącznie z tym określnikiem.
8. Zastosowanie określników chronologicznych swobodnych i związanych (występujących wyłącznie z pewnymi tematami). Przy tworzeniu słowa kluczowego istnieje możliwość zastąpienia skomplikowanej chronologii prostszym określnikiem chronologicznym swobodnym.
9. W języku swobodnych słów kluczowych – jak w jhp KABA – rozróżnia się tematy formalne i tematy związane z opracowaniami literatury.
10. Istotna jest konsekwencja w tworzeniu terminologii słów kluczowych; dla utrzymania tej konsekwencji i kontroli słownictwa tworzy się – na bieżąco – słownik słów kluczowych.

Dla dodatkowego dookreślenia i zaklasyfikowania tematycznego dokumentu obok języka swobodnych słów kluczowych zastosowana została klasyfikacja systematyczna (według dziedzin wiedzy), oparta na klasyfikacji systematycznej tradycyjnego Katalogu Systematycznego BJ. Ma to na celu zapewnić użytkownikowi dodatkowe narzędzie wyszukiwawcze w postaci katalogu, będącego alternatywą dla okna wyszukiwarki, a sprawdzającego się szczególnie w przypadku starszych użytkowników oraz poruszających się w obrębie jednej dziedziny wiedzy.

3.2.4. Dyskusje nad przyszłością języków przedmiotowych i opracowania rzeczowego dokumentów w bibliotekach cyfrowych

Rozwój technologii, rozwój bibliotek cyfrowych oraz zmiana oczekiwań użytkowników wpływają na fakt, iż toczą się dyskusje nad przyszłością języków informacyjno-

-wyszukiwawczych i opracowania rzeczowego dokumentów w bibliotekach cyfrowych.

Szybkiemu rozwojowi bibliotek cyfrowych towarzyszy problem nieprecyzyjnej terminologii metadanych oraz brak kontroli słownictwa, służącego do opisu treści dokumentów. Problemy te niejednokrotnie sygnalizowano i omawiano na konferencjach bibliotecznych, a także w licznych publikacjach [Kisilowska,Woźniak-Kasperek 2006, s. 75-76; Woźniak-Kasperek 2004, 2008, 2009; Sosińska-Kalata 2004; Lapis 2002; Nahotko, 2004, 2010].

Istnieje potrzeba stworzenia jednolitych standardów opracowania dokumentów w bibliotekach cyfrowych, ujednoliconej terminologii oraz systemu koordynacji działań w tym zakresie. Mówiła o tym na konferencji *Uniwersum piśmiennictwa wobec komunikacji elektronicznej* Jadwiga Woźniak-Kasperek, podkreślając, jak „daleka od doskonałości" jest „organizacja informacji, nawigacji, systemu etykietowania i wyszukiwania informacji" oraz funkcjonalność współczesnych bibliotek cyfrowych [Woźniak-Kasperek 2009].

Istotny jest wybór języka informacyjno-wyszukiwawczego oraz oprogramowania w bibliotekach cyfrowych. Większość oprogramowania np. dLibra nie tworzy kartotek haseł wzorcowych ani nie umożliwia konwersji tych kartotek z katalogów bibliotecznych.

Biblioteki cyfrowe liczą się, co zrozumiałe, z oczekiwaniami użytkowników, którzy, podobnie jak użytkownicy całego Internetu, preferują wyszukiwanie dokumentów według swobodnych słów kluczowych. Konsekwencją jest niepełna, niejednoznaczna i nieujednolicona informacja rzeczowa oraz niekontrolowane – lub tylko częściowo podlegające kontroli – słownictwo języka słów kluczowych.

Biblioteki cyfrowe nie zawsze są połączone z katalogami bibliotecznymi, które dają użytkownikowi większe możliwości wyszukiwawcze.

Z drugiej strony – języki haseł przedmiotowych, z ich rozbudowaną składnią, długimi „zdaniami", szczegółowymi zasadami i sztucznością słownictwa wydają się dziś „przestarzałe", za bardzo skomplikowane, za trudne dla współczesnego użytkownika.

Zbliżeniu tych języków do potrzeb nowych technologii oraz oczekiwań użytkowników służą przeprowadzane modyfikacje, które zmierzają w stronę uproszczenia i ujednolicenia słownictwa oraz gramatyki, bliskości z wyrażeniami języka naturalnego, a także ich dostosowania do standardów i norm międzynarodowych oraz coraz nowocześniejszych systemów komputerowych. Modyfikacje polegają także na skróceniu długich „zdań" tematów i określników. Nie zawsze jednak modyfikacje języków haseł przedmiotowych są korzystne; wiążą się z dodatkowymi kosztami bibliotek, na przykład z koniecznością ponownego katalogowania przedmiotowego opracowanych wcześniej dokumentów. Niektóre modyfikacje prowadzą do odejścia od zasad języków haseł przedmiotowych, burzą logikę i strukturę tych języków. Pojawia się problem: czy nie lepiej tworzyć nowe języki informacyjno-wyszukiwawcze dla nowych dokumentów (języki deskryptorowe, języki oparte na słowach kluczowych, tezaurusy z poszczególnych dziedzin, języki współtworzone przez bibliotekarzy i użytkowni-

ków), pozostawiając dotychczasowe języki haseł przedmiotowych w katalogach bibliotecznych w niezmienionej formie?

J. Woźniak-Kasperek wskazuje na dwa rozwiązania, które mogą zbliżyć języki haseł przedmiotowych do użytkownika:

1. zaopatrywanie haseł przedmiotowych rozwiniętych o „nienaturalnej" budowie w odsyłacze całkowite, uzupełniające, wyliczające i orientacyjne,
2. stworzenie zmodyfikowanego interfejsu wyszukiwawczego do istniejących baz z opisami przedmiotowymi, który sam układałby z pytań hasła przedmiotowe [Kisilowska,Woźniak-Kasperek 2006, s. 79].

Osobny problem to opracowanie rzeczowe zbiorów specjalnych w bibliotekach cyfrowych. Biblioteki cyfrowe udostępniają różne rodzaje dokumentów od książek, przez rękopisy, inkunabuły i stare druki, zbiory muzyczne, kartograficzne, ikonograficzne (fotografie, plakaty, obrazy in.), dokumenty życia społecznego, obiekty muzealne, archiwalia. Specyfika każdego z tych zbiorów wymaga ustalenia osobnych reguł opracowania rzeczowego.

Kolejną zmianą w opracowaniu rzeczowym jest opracowanie w bibliotekach cyfrowych – obok całych dokumentów – fragmentów tekstów. Z pewnością w przyszłości tendencja ta będzie się rozwijać.

Innym rozwiązaniem związanym ze zmianą oczekiwań czytelników jest możliwość współtworzenia produktów i usług informacyjnych przez bibliotekarza i użytkownika, wspólne opracowanie treściowe dokumentu. Biblioteki cyfrowe zachęcają czytelnika: 'Zaproponuj słowa kluczowe, które twoim zdaniem dobrze opisują publikację'. Zacieranie różnic między producentem a odbiorcą, widoczne na przykład w tworzeniu niektórych źródeł informacji (np. w encyklopedii Wikipedia) może stać się również przyszłością bibliotek. Niektóre niebezpieczeństwa z tym związane to m.in. współistnienie terminologii potocznej i naukowej w katalogach, brak kontroli słownictwa, niejednolite opracowanie rzeczowe tych samych lub podobnych treściowo publikacji, współkatalogowanie dokumentów przez specjalistów i nieprofesjonalistów, nieuchronność pomyłek, dowolność tworzonych opisów rzeczowych, błędy rzeczowe lub formalne itp. Mnożenie dowolnego, niekontrolowanego słownictwa może spowodować obniżenie poziomu i jakości opracowania rzeczowego, kwestionować wiarygodność informacji oraz przyczyniać się do szumu informacyjnego [Dudczak; Parkoła 2010].

Zmiany w opracowaniu rzeczowym wiążą się z nowymi kompetencjami bibliotekarzy i koniecznością szkoleń w tym zakresie.

Nowymi wyzwaniami dla opracowania rzeczowego i kształcenia bibliotekarzy są:

- współpraca między bibliotekarzami opracowania rzeczowego, informatykami oraz pracownikami oddziałów informacji naukowej;
- określenie wykładowców i organizatorów szkoleń; czy nadal będą nimi bibliotekarze-praktycy, czy może bibliotekoznawcy i pracownicy uczelnianych instytutów informacji naukowej? Czy szkolenia będą organizowane przez biblioteki? uczelnie? konsorcja bibliotek cyfrowych? Bibliotekę Narodową lub/i Centrum NUKAT?

- wprowadzenie nowych form szkoleń np. e-learning;
- komunikacja z użytkownikiem poprzez email, Skype itd.; dostarczanie informacji o treści dokumentu bezpośrednio do użytkownika i współpraca z nim (etykietowanie, komentarze i opinie czytelników, rola moderatorów itd.);
- określenie kompetencji bibliotekarzy w dziedzinie opracowania rzeczowego (wiedza biblioteczna, wiedza specjalistyczna, wiedza z dziedziny informacji naukowej, wiedza informatyczna itd.);
- kształcenie bibliotekarzy w ramach studiów podyplomowych, kursów specjalistycznych w celu nabycia nowych umiejętności itp.

Kolejnym aspektem jest różnorodność informacji rzeczowej współtworzonej przez bibliotekarzy, wydawców i użytkowników. Istotne jest nawiązanie współpracy bibliotek z wydawcami i księgarzami. Cenny dla użytkownika jest – zawarty w serwisach wydawniczych i dokumentach cyfrowych – aparat naukowy (przypisy, aktualne opracowanie krytyczne dokumentu, zwłaszcza starego dokumentu itp.). Wydawcy zamieszczają informację prospektywną i retrospektywną związaną z daną publikacją, planami wydawniczymi itd., zamieszczają fragmenty książki, opinie krytyków, recenzje, rekomendacje zakupu innych tytułów, rekomendacje książek o podobnej tematyce. Tego typu informacje także mogą stanowić część bibliotecznej informacji rzeczowej o dokumencie.

Dla użytkownika ważne jest, aby zasób był dostępny bez ograniczeń. Czytelnik chce szukać informacji z jednego miejsca i w jednym czasie, przeszukiwać jednocześnie: katalogi biblioteczne, biblioteki cyfrowe, bazy danych, serwisy wydawnicze i księgarskie. Zastosowanie nowych technologii, przeglądarek i wyszukiwarek stwarza użytkownikowi nowe możliwości wyszukiwawcze, także w zakresie informacji rzeczowej np. aktywny spis treści, aktywny indeks, przeszukiwanie tekstu przez tzw. słowa kluczowe (także w streszczeniach), opcja tzw. *Look inside*; możliwość połączenia książki z plikami audio, wideo np. filmami i in.

Użytkownik otrzymuje informację rzeczową z różnych źródeł: bibliotek cyfrowych, serwisów księgarskich i wydawniczych, katalogów bibliotecznych, portali internetowych i repozytoriów wiedzy. Problemem jest wiarygodność informacji rzeczowej, zwłaszcza w sytuacji, gdy treści mają hipertekstualny, interaktywny i multimedialny charakter [Skórka 2006; Kocójowa 2005], a efektowności informacji nie zawsze towarzyszy jej rzetelność.

Charakterystycznymi tendencjami współczesnej informacji rzeczowej są: różnorodność i rozproszenie informacji; malejąca rola katalogów rzeczowych, które pełnią funkcję uzupełniającą; nowe sposoby tworzenia informacji rzeczowej; fragmentaryczność, niepełność i niewiarygodność informacji, współtworzenie informacji rzeczowej przez specjalistów i nieprofesjonalistów: bibliotekarzy, księgarzy, wydawców i użytkowników. Przed specjalistami w zakresie opracowania rzeczowego oraz architektury informacji stoją zadania nie tylko tworzenia nowej informacji, ale i przywrócenia wiarygodności informacji rzeczowej, budowanej przez różne podmioty.

Bibliografia

1. Babik, Wiesław (2010). *Słowa kluczowe*. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego, 241, [1] s.
2. Ćwiekowa, Jadwiga (1988). *Opracowanie tematyczne piśmiennictwa : język haseł przedmiotowych: katalog przedmiotowy*. Warszawa: Państwowe Wydawnictwo Naukowe, 298, [1] s.
3. Dudczak, Adam; Parkoła, Tomasz (2010). *Rola użytkowników w tworzeniu zasobów bibliotek cyfrowych* [W:] Mazurek, Cezary; Stroiński, Maciej; Węglarz, Jan red. (2009). *Polskie biblioteki cyfrowe 2009. Materiały z konferencji zorganizowanej dnia 9 grudnia 2009 roku przez: Bibliotekę Kórnicką PAN, Poznańską Fundację Bibliotek Naukowych, Poznańskie Centrum Superkomputerowo-Sieciowe*. Poznań: Ośrodek Wydawnictw Naukowych, s. 45-51.
4. Gawrysiak, Piotr (2009). *Klasyfikacja : narzędzie zarządzania i wyszukiwania informacji*. Warszawa: MOST Press, 101 s.
5. Głowacka, Teresa (2003). *Analiza dokumentu i jego opis przedmiotowy*. Warszawa: Wydawnictwo SBP, 92, [1] s.
6. Głowacka, Teresa (red.) (2000). *Język haseł przedmiotowych KABA : zasady tworzenia słownictwa*. Warszawa: Wydawnictwo SBP, 207, [1] s.
7. Kisilowska, Małgorzata; Woźniak-Kasperek, Jadwiga (2006). *Czy katalogowanie ma przyszłość? Dwugłos względnie uporządkowany* [W:] *Biblioteki XXI wieku. Czy przetrwamy? : II Konferencja Biblioteki Politechniki Łódzkiej, Łódź, 19-21 czerwca 2006 r. : materiały konferencyjne*. Łódź: Politechnika Łódzka, s. 71-82.
8. Kocójowa, Maria (red.) (2005). *Profesjonalna informacja w Internecie*. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego, 254 s.
9. Lapis, Karina (2002). *Efektywne metody pozyskiwania i przekazywania informacji za pośrednictwem witryny internetowej* [W:] Krzemińska, Wanda; Nowak, Piotr red. (2002) *Przestrzenie informacji*. Poznań: Sorus, s. 47-51.
10. Nahotko, Marek (2004). *Metadane : sposób na uporządkowanie Internetu*. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego, 201 s.
11. Nahotko, Marek (2010). *Komunikacja naukowa w środowisku cyfrowym : globalna biblioteka cyfrowa w informatycznej infrastrukturze nauki*. Warszawa: Wydawnictwo SBP, 283, [1] s.
12. Sadowska, Jadwiga (1991). *Instrukcja tematowania i katalogu przedmiotowego*. Warszawa: Wydawnictwo SBP, 100 s.
13. Sadowska, Jadwiga (2000). *Hasła przedmiotowe w teorii Adama Łysakowskiego i praktyce „Przewodnika Bibliograficznego"*. Warszawa: Biblioteka Narodowa, 196 s.
14. Sadowska, Jadwiga (2003). *Język haseł przedmiotowych Biblioteki Narodowej : studium analityczno-porównawcze*. Warszawa: Biblioteka Narodowa, 244 s.
15. Skórka, Stanisław (2006). *Użytkownicy systemów hipertekstowych : strategie poszukiwania informacji w edukacyjnym serwisie internetowym*. Kraków: Wydawnictwo Naukowe Akademii Pedagogicznej, 204 s., [5] k. tabl. kolor.
16. Sosińska-Kalata, Barbara (2002). *Klasyfikacja : struktury organizacji wiedzy, piśmiennictwa i zasobów informacyjnych*. Warszawa: Wydawnictwo SBP, 261, [1] s.
17. Sosińska-Kalata, Barbara (2004). *Nowe tendencje w rozwoju języków informacyjno-wyszukiwawczych* [W:] Chruścińska, Jadwiga; Kubisz, Ewa red. (2004). *Opracowanie rzeczowe piśmiennictwa : materiały z seminarium, Warszawa, 26 marca 2003 r.* Warszawa: Wydawnictwo CEBID, s. 5-16.

18. Stopa, Adam (2002). *O treści książek: opracowanie rzeczowe piśmiennictwa*. Warszawa: Wydawnictwo CEBID, 196 s.
19. Ścibor, Eugeniusz; Tomasik-Beck, Joanna (1995). *Metodyka budowy tezaurusów*. Warszawa: IINTE, [2], 175 s.
20. Woźniak-Kasperek, Jadwiga (2004). *Tendencje w teorii i praktyce języków informacyjno-wyszukiwawczych*. [W:] Stefańska-Matuszyn, Maria red. (2004) *Opracowanie przedmiotowe : osiągnięcia naukowe i praktyka, Warsztaty: 31.03.–1.04.2004*. Warszawa: WSE-I, s. 3-24.
21. Woźniak-Kasperek, Jadwiga (2005). *Podstawy budowy tezaurusa : poradnik*. Warszawa: Wydawnictwo SBP, 100, [4] s.
22. Woźniak-Kasperek, Jadwiga (2008). *System organizacji wiedzy czy folksonomia?* [W:] Zybert, Elżbieta Barbara; Grabowska, Dorota red. (2008) *Książka, biblioteka, informacja w kręgu kultury i edukacji : praca zbiorowa*. Warszawa: Wydawnictwo SBP, s. 117-124.
23. Woźniak-Kasperek, Jadwiga (2009). *Funkcjonalność i dostępność bibliotek cyfrowych* [W:] Migoń, Krzysztof; Skalska-Zlat, Marta red. (2009) *Uniwersum piśmiennictwa wobec komunikacji elektronicznej*. Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego, s. 393-405.
24. Woźniak-Kasperek, Jadwiga, Bierczyński, Piotr red. (2006). *Opracowanie rzeczowe : praca zbiorowa*. Warszawa: Wydawnictwo SBP, 243, [1] s.

4. Korelacje między opisami w katalogach bibliotecznych i bibliotekach cyfrowych

Analizując zasoby cyfrowe i ich opisy bibliograficzne w formacie Dublin Core, trudno zorientować się co było podstawą ich utworzenia. Niejednokrotnie daje się zauważyć, iż redaktor samodzielnie opisywał dokument na podstawie uzyskanego obrazu cyfrowego. Z uwagi na fakt iż wielu spośród redaktorów nie jest bibliotekarzami, odbija się to na jakości opisów. Oczywistym rozwiązaniem, które nasuwa się bibliotekarzowi pracującemu z kopiami cyfrowymi dokumentów jest przejęcie utworzonego opisu bibliograficznego dla oryginału i dostosowanie go do opisu kopii cyfrowej. Takie rozwiązanie zostało przewidziane przez twórców platformy dLibra, ale wydaje się rozwiązaniem mało popularnym i wykorzystywanym w niewielu bibliotekach cyfrowych. Zgodnie ze wspomnianą już wcześniej rekomendacją formatów MARC 21 i Dublin Core jako właściwych dla metadanych obiektów digitalizowanych słusznym założeniem jest, aby tworzyć opis dla dokumentu cyfrowego na podstawie opisu bibliograficznego oryginału. Osiągnięta zostaje spójność wynikająca z prostej zależności: tak jak podstawą dla kopii cyfrowej jest oryginał tak na podstawie opisu bibliograficznego oryginalnego dokumentu w formacie MARC 21 tworzony jest opis kopii cyfrowej w formacie Dublin Core.

Do głównych zalet tworzenia opisów w formacie Dublin Core na podstawie opisów w formacie MARC 21 jest ich jednolitość wynikająca z jasnych i ogólnie respektowanych zasad stosowanych przy tworzeniu opisów w bibliotekach. Poza tym wykorzystywana jest praca całego zastępu profesjonalnych katalogerów, którzy materiał zbadali dokładnie z autopsji i wprowadzili w obrót informację, którą należy tylko umiejętnie wykorzystać, konwertując do formatu właściwego dla kopii już opracowanego obiektu.

4.1. Konwersja z formatu MARC21 do Dublin Core

Podstawą konwersji z formatu MARC 21 do formatu Dublin Core jest tablica przejścia, która w swojej najprostszej postaci została przedstawiona w tabeli 1:

Tabela 1. Najprostsza wersja tablicy konwersji z formatu MARC 21 do formatu Dublin Core[14]

Pole formatu MARC 21	Atrybut formatu Dublin Core
041 \a	język
100 \a, \d, 110, 111	autor, współtwórca

[14] Pełna wersja tablicy konwersji dostępna na stronie http://www.loc.gov/marc/marc2dc.html

245 \a, \b 130, 210, 240, 242, 246, 730, 740	Tytuł, podtytuł, tytuł równoległy, tytuł ujednolicony, tytuł odmienny
260 \a	miejsce wydania
260 \b, \f	wydawca
260 \c	data
300	format
310 \a	częstotliwość
440, 490, 500, 505, 520	opis
521	odbiorca
600, 610, 611, 650, 651, 655	temat
700 \e, 710, 711	współtwórca
770, 772, 775, 776, 777, 780, 785, 787	powiązania
856 \u	identyfikator

Aby zastosować tablicę przejścia w praktyce, wystarczy opracować odpowiedni dokument tekstowy z bardzo prostym zapisem:

Format=
Creator=100:${a} ${d};110;111;
Title=245:${a} ${b} ${n} ${p};
Title_version=246;
Subject=600;610;611;650;651;655;
Description=440;490;500;505;
Systematic=595
Publisher=260b;260f;
Place=260a;
Contributor=700:${a} ${b} ${d} ${c} ${e};710:${a} ${b} ${n} ${d} ${c};711:${a} ${n} ${d} ${c} ${e};
Date=260: ${c} ${g};
Type=;
Identifier=;
Language=041a
Relation=;
Coverage=;

Porównując tabele konwersyjne ze stosowanym w praktyce plikiem tekstowym od razu rzuca się w oczy, że nie wszystkie pola zostały do konwersji wykorzystane. Nie jest to obligatoryjne, a tu został przedstawiony tylko przykład. Każdy z redaktorów sam określa jakie pola zostaną przejęte i jakie atrybuty w związku z tym zostaną zastosowane. Jest to zależne od dokumentu jaki w danym momencie opracowuje. Odmienny będzie zestaw pól i atrybutów dla inkunabułu, a inny dla pojedynczego numeru XIX-wiecznego czasopisma.

4.2. Dostęp do dokumentów z różnych poziomów metadanych

Oczywistym jest stwierdzenie, że najpopularniejszym źródłem wyszukiwania informacji w obecnym momencie jest Google. Ta sytuacja zaistniała kilka lat temu i raczej nie zmieni się w najbliższym okresie. Według statystyk za 2007 r. ranking najpopularniejszych na świecie wyszukiwarek kształtował się następująco:

1. Google – 41 mln – 62,4%
2. Yahoo – 8,5 mln – 12,8%
3. Baidu.com – 3,4 mln – 5,2%[15]

W maju 2011 r. wyniki rankingu wyszukiwarek w Stanach Zjednoczonych nie odbiegały od tych sprzed 4 lat:

1. Google – 63,3%
2. Yahoo! – 19,6%
3. Microsoft – 13,1%[16]

W grudniu 2011 r. wyniki rankingu w Stanach Zjednoczonych kształtowały się na następującym poziomie:

Google – 65,9%
Microsoft – 15,1%
Yahoo! – 14,5%[17]

Z tego krótkiego zestawienia widać jak wyszukiwarka Google stopniowo zdobywa rynek amerykański, spychając konkurencję w cybernetyczną niszę. Dla porównania wyniki rankingu wyszukiwarek w Polsce z ostatniego tygodnia 2011 r. jeszcze bardziej ukazują dominację Google.

Google – 96,37%
MSN (Bing) – 1,46%
Netsprint – 1,21%[18]

Powszechne stosowanie wyszukiwarek jako narzędzia wyszukiwawczego zmusza twórców bibliotek cyfrowych do takiego konstruowania metadanych aby wyniki wyszukiwawcze znalazły się na pierwszej stronie z wynikami zapytania. Brak wyniku wśród pierwszych 10 skazuje dokument na „niebyt informacyjny". Nie znaczy to, że najważniejszym zadaniem redaktorów i bibliotekarzy jest pozycjonowanie stron internetowych, natomiast jest to niewątpliwie aspekt nad którym warto się skupić w dalszych pracach nad rozwojem bibliotek cyfrowych.

Użytkownicy zainteresowani dostępem do zasobów pełnotekstowych na szczęście nie traktują Google jako wyrocznię i przeszukują Internet głębiej, wchodząc także na strony katalogów bibliotecznych, gdzie formatem opracowania dokumentów jest

[15] http://www.comscore.com
[16] http://www.comscore.com/Press_Events/Press_Releases/2011/6/comScore_Releases_May_2011_U.S._Search_Engine_Rankings
[17] http://www.comscore.com/Press_Events/Press_Releases/2012/1/comScore_Releases_December_2011_U.S._Search_Engine_Rankings
[18] Źródło: Gemius SA, gemiusTraffic, 26/12/11 - 01/01/12.

MARC 21, nieprzeszukiwalny przez roboty wyszukiwarek internetowych. Zastosowanie konwersji z formatu MARC 21 do formatu Dublin Core, oprócz odsłonięcia treści opisu dla wyszukiwarek, niesie ze sobą jeszcze jedną bardzo ważną zaletę. Opracowanie rekordu bibliograficznego dla oryginału widoczne jest nie tylko z poziomu katalogu biblioteki taki rekord tworzącej, ale również z ogólnopolskiego Katalogu Rozproszonego Bibliotek Polskich (KaRo). Konsekwentne katalogowanie dokumentów, których opisy trafiają do KaRo i ich uzupełnianie o odnośniki do bibliotek cyfrowych udostępniających ich kopie cyfrowe doprowadzi do sytuacji, kiedy każdy zdigitalizowany dokument będzie miał swój opis w ogólnodostępnej bazie danych. Co więcej, odnośnik przenosi użytkownika za pomocą jednego kliknięcia do pełnotekstowej wersji dokumentu, niejednokrotnie wyposażonej w wydzieloną, możliwą do przeszukiwania i edytowania warstwę tekstową. Świadomość istnienia dokumentów o wolnym dostępie jest coraz większa, co można zaobserwować po systematycznie przyrastającej liczbie użytkowników bibliotek cyfrowych.

4.3. Sprzężenie zwrotne w informacji o dokumencie

Z praktyki bibliotecznej wynika, że jest pewna grupa użytkowników, dla których dostarczenie wersji pełnotekstowej dzieła, z wykonanym OCR nie jest najważniejsze. Potrzebują oni metadanych, czyli informacji zebranych i uporządkowanych przez specjalistę na temat dokumentu, który został im w pełni udostępniony. Dotyczy to szczególnie czasopism, przy opracowaniu których bibliotekarz wykonuje ogromną pracę obejrzenia całego zasobu, niejednokrotnie ukazującego się przez dziesiątki lat, uwzględniając zmiany tytułu, zmiany redaktorów, charakter czasopisma, itp. Jest to praca porównywalna z tworzeniem monografii na temat danego dokumentu, przy czym bibliotekarze niejednokrotnie z wszelkich opracowań monograficznych korzystają, tworząc rekord bibliograficzny.

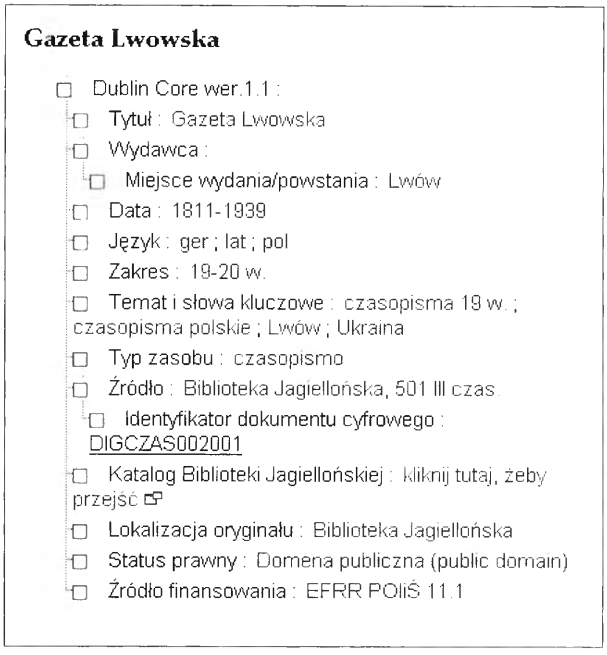
Dla porównania podany został opis czasopisma z katalogu bibliotecznego w formacie MARC 21 i jego postać po konwersji do formatu Dublin Core.

Jak widać różnica pomiędzy tymi opisami jest ogromna. Użytkownik aby osiągnąć stan wiedzy z pierwszego przykładu musiałby spędzić długie godziny przeglądając cały zasób czasopisma. Stąd bardzo istotne dla niego informacje bez odpowiedniego narzędzia mogą być niedostępne, o ile ograniczy się do wyszukiwania jedynie za pomocą wyszukiwarki internetowej i wreszcie trafi do biblioteki cyfrowej. Zastosowanie odnośnika od opisu zdigitalizowanego dokumentu do pełnego opisu w katalogu bibliotecznym daje użytkownikowi możliwość zapoznania się z profesjonalnie stworzonym i pełnym treści zestawem metadanych.

Terminem sprzężenie zwrotne w informacji o dokumencie określa się tu odnośnik umieszczany w polu 856 przekierowujący użytkownika z rekordu bibliograficznego do pełnotekstowej wersji dzieła i atrybut powiązania (w tym przypadku opatrzony etykietą „Katalog Biblioteki Jagiellońskiej") odnoszący do pełnego rekordu bibliograficznego w katalogu bibliotecznym. Jest to rozwiązanie o tyle proste, co skuteczne.

Rekord MARC

Pole	Ws 1	Ws 2	Dane
001			vtls000576734
003			UJ
005			20110610042800.0
008			031007d18111939un ir|p| | |000 |0pol c
035			\a zz2003973578
039		9	\a 201106100428 \b VLOAD \c 201106100423 \d VLOAD \c 201106061239 \d kazmier \c 201012080425 \d VLOAD \y 200804231227 \z babraj
040			\a WR O/DA \c WR O/DA \d WR O/DA \d KR U/BBs \d KR U/BB \d KR U/mel
041	0		\a pol \a ger \a lat
245	0	0	\a Gazeta Lwowska.
246	1	3	\a Gazeta Lwowska. \p Oddział polityczny \f 1819, t.2-1826, t.1
246	1		\i Nota o prenumeracie w nr 66=131 (10 czerw. 1822): \a Gazeta Lwowska t. 1. Wiadomości polityczne
246	3		\i Od red. : \a Gazeta Lwowska Polska
246	3		\i Memoryał F. Krattera: \a Lemberger Pohlnische Zeitung
260			\a Lwów : \b [s. n.], \c 1811-1939.
300			\c 26 cm.
310			\a Dz., \b 1849-1939
321			\a Wych. we wtorek i piątek, \b 1811-1815
321			\a 4 x tyg. ; w poniedziałek, środę, czwartek, sobotę, \b 1816-1818
321			\a 3 x tyg. ; we wtorek, czwartek, sobotę, \b 1831, nr 118-1848
321			\a 3x tyg. ; pon., śr., piątek, \b 1819-1831, nr 117
362	0		\a 1811, nr 1 (2 kwiet.)- R.129, nr 198 (2 wrzes. 1939).
500			\a Ostatni znany numer 198 z 2 wrzes. 1939.
500			\a Od 1819, no 119=238 (18 paźdz.) redakcya F. Krattera; od 1823, nr 78 (7 lip.) redakcya J. Bensy; od 1827, nr 1 (3 stycz.) red. Mikołaia Michalewicza; od 1835 redakcyja J. N. Kamińskiego ; od 1848, nr 1 (3 stycz.) red. M. Srzemiawa Sartyni ; od 1862, 1 lip. red. Adolf Rudyński ; od 1873, nr 80 red. Władysław Łoziński, od 1883, nr 123 (1 czerw.) red. Adam Krechowiecki, od 1918, nr 269 (14 grudz.) red. Stanisław Rossowski.
500			\a Od 1816: drukiem Józefa Pillera.
500			\a Od R.39 (1850): format 40 cm, od R.57 (1867): 45 cm, od R.60 (1870): 47 cm.
500			\a Do czerw. 1862 niemiecka gazeta "Lemberger Zeitung" (nie zawierająca dodatków) wychodzi w tej samej redakcji; od 1 lip. 1862 rozdział.
505	1		\a Od R. 57, nr 1 (2 stycz. 1867) w paginacji gazety, w postaci zredukowanej rubryki wychodzi Dziennik Urzędowy.
510	2		\a Bibliografia Gazety Lwowskiej 1811-1846, nr 61 / Władysław Staniszewski // Przewodnik Naukowy i Literacki . - 1911, 02-1914, 08.
510	2		\a Stulecie gazety lwowskiej 1811-1911. T.1, cz.1, Historya Gazety Lwowskiej. - Lwów, 1911 / pod. red. Wilhelma Bruchnalskiego.
515			\a Oznaczanie rocznika od R.38 (1849)-R.41 (1852) a następnie korekta: R.43 (1853)-.
515			\a Od R.63, nr 126 (3 czerw.1873) zanik paginacji ciągłej ; każdy numer osobno liczbowany.
525			\a Od nr 15 (15 maja 1811) [egz. Wa N], o osobnej paginacji (ca 2 s., z czasem urasta do 12 s.): Nadzwyczayny Dodatek do Gazety Lwowskiey nr ... ; w l. 1819-1822 przyłączany wg daty wychodzenia albo do Gazety Lwowskiej albo do Rozmaitości ; ca od nr 99 (28 sierp. 1826) przeważnie pod zmiertowanym tytułem : Dodatek Nadzwyczayny do Gazety Lwowskiey nr ... ; często tekst rozpoczyna się na odwrocie ostatnie strony arkusza Gazety ; tyt. równol. Ausserordentliche Beilage zur Lemberger Zeitung, np. nr 68 (9 czerw. 1848) [kr U].
525			\a W 1811stałe dwa dodatki do każdego numeru, stricte w paginacji ciągłej gazety, pt.: Dodatek1-szy do nru ... Gazety Lwowskiey [tekst pol.] i Dodatek 2-gi do nru ... Gazety Lwowskiey [tekst urzędowy niem, pol., łac.].
530			\a Zasób w wersji cyfrowej dostępny online.
546			\a Tekst gł. pol.; tekst dodatkow przeważnie niem., łac. \b Tekst niem. pisany gotykiem
546			\a Od R. 57(1867) tekst Dziennika Urzędowego pol., niem. ukr. \b Tekst niem., ukr. pisany gotykiem.
580			\a Od 1819 do 1822 wprowadzenie w nagłówku numeracji ogólnego zbioru, ponieważ "w każdym tygodniu ukazują się 3 numery Wiadomości politycznych i trzy Rozmaitości" [1819, nr 1 <og. zb. nr 1> -149 <og. zb. nr 299> ; 1820 ,nr 1 -148 <og. zb. nr 296> ;1821, nr 1 (3 stycz.)-150 <og. zb. nr 298> ; 1822, nr 1 (2 stycz.)-151 <og.zb. nr 300>].
594			\a JBC EFPR 08'2010
650			\a Polityka swiatowa \y 19 w. \v czasopisma.
651			\a Lwów (Ukraina) \y 19 w. \v czasopisma.
651			\a Polska \x sytuacja gospodarcza \y 19 w. \v czasopisma.
655			\a Czasopisma polskie \y 19 w.
655			\a Czasopisma polskie \z Ukraina \z Lwów \y 19 w.
700	1		\a Kratter, Franciszek \d (ca 1760-1838). \e Red.
700	1		\a Łoziński, Władysław \d (1843-1913). \e Red.
700	1		\a Krechowiecki, Adam \d (1850-1919). \e Red.
700	1		\a Rossowski, Stanisław \d (1861-1940). \e Red.
700	1		\a Benza, Antoni \d (1787-1859). \e Red.
700	1		\a Michalewicz, Mikołaj \d (1792-1846). \e Red.
700	1		\a Kamiński, Jan Nepomucen \d (1777-1855). \e Red.
700	1		\a Sartyni, Mateusz Eustachy \d (1796-1884). \e Red.
700	1		\a Rudyński, Adolf \d (1819-1873). \e Red.
740	0	2	\a Dziennik Urzędowy
770	0		\t Amtsblatt zur Lemberger Zeitung no ...= Dziennik Urzędowy do Gazety Lwowskiej nr ... \g 1848-1866 \w =x001121068
770	0		\t Dodatek do Gazety Lwowskiey nr ... : doniesienia urzędowe \g 1812, nr 38-1848, nr 68 \w =x001109150
770	0		\t Dodatek Tygodniowy \g 1851-t.10, nr 52 (1860) \w zz2003915981
770	0		\t Dodatek Osobny przy Gazecie Lwowskiej \g 1861,nr 1-1862, nr 31 [egz. Kr U] \w =x001160532
770	0		\t Dodatek Nadzwyczajny do nr ... Gazety Lwowskiej \g 1867, [nr 1]-nr 12 \w =x001160640
770	0		\t Dodatek Tygodniowy do Gazety Lwowskiej \g 1868, nr 1-1869, nr 26 \w =x001160719
770	0		\t Dodatek Miesięczny do Gazety Lwowskiej \g R. 1, t. 1, z. 1 (stycz. 1872)-R. 2, t. 1, z. 1, 64 s. (stycz. 1873) \w bb2001159313
770	0		\t Przewodnik Naukowy i Literacki \g 1873, t.1 [luty], s. 65-R.19 (1921) \w da2001150046
770	0		\t Narodna Časopis \g R. 1 (1891)-1900 \w =x002079881
787	1		\t Rozmaitości : przydatek do numeru ... gazety Lwowskiej \g 1817-1859 \w zz2002955952
856	4	1	\u http://jbc.bj.uj.edu.pl/dlibra/dlibra/publication/1533?tab=1



Rys. 1. Opis bibliograficzny czasopisma w katalogu bibliotecznym w formacie MARC 21 oraz w bibliotece cyfrowej w formacie Dublin Core

Podsumowując, opis zdigitalizowanej wersji dokumentu w bibliotece cyfrowej ma tak samo istotne znaczenie jak opis oryginalnego dokumentu w katalogu bibliotecznym. Nie jest konieczne stworzenie dokładnego opisu w formacie MARC 21, czy ustanowienie kryteriów w tablicy przejścia, choć przyczynia się to do korelacji danych pomiędzy biblioteką tradycyjną i repozytorium przy tej bibliotece stworzonym, ale zastosowanie jednolitego schematu w formacie Dublin Core wydaje się koniecznością dla stworzenia jednorodnego opisu funkcjonującego zarówno w Federacji Bibliotek Cyfrowych jak i w poszczególnych repozytoriach. Kwestia zastosowania języków informacyjno-wyszukiwawczych, doboru słów kluczowych, czy stosowania tagów pozostaje otwarta i leży po stronie bibliotekarzy i redaktorów świadomych procesów wyszukiwawczych zachodzących we współczesnej przestrzeni cyfrowej i odpowiedzialnie tworzących repozytoria w których wersje cyfrowe dokumentów będą łatwe do odnalezienia.

Bibliografia

1. comScore Releases December 2011 U.S. Search Engine Rankings (2011). [dok. elektr.] http://www.comscore.com/Press_Events/Press_Releases/2012/1/comScore_Releases_December_2011_U.S._Search_Engine_Rankings [odczyt: 03.01.2012].
2. comScore Releases May 2011 U.S. Search Engine Rankings (2011). [dok. elektr.] http://www.comscore.com/Press_Events/Press_Releases/2011/6/comScore_Releases_May_2011_U.S._Search_Engine_Rankings [odczyt: 03.08.2011].
3. comScore Releases U.S. Search Engine Rankings (2007). [dok. elektr.] http://www.comscore.com [odczyt: 30.02.2008].
4. Gemius SA. (2012). *Wyszukiwarki – silniki* [dok. elektr.] http://www.ranking.pl/pl/rankings/search-engines.html [odczyt: 5.01.2012].
5. Library of Congress Development and MARC Standards Office (2008). *MARC to Dublin Core Crosswalk* [dok. elektr.] http://www.loc.gov/marc/marc2dc.html [odczyt: 30.09.2011].

Formaty dokumentów w bibliotekach cyfrowych

Władysław Marek Kolasa

Opracowanie zawiera zwięzłą charakterystykę formatów dokumentów używanych w bibliotekach i repozytoriach cyfrowych. W kolejnych rozdziałach omówiono związane z tym problemy teoretyczne, techniczne oraz zaprezentowano najważniejsze specyfikacje i zastosowania. Pracę otwiera rozdział na temat zmian formatów w cyklu tworzenia publikacji cyfrowej (formaty przejściowe) oraz zwięzła charakterystyka formatów archiwalnych. W dalszej części szczegółowo omówiono formaty prezencyjne, w podziale na tekstowe, graficzne, hybrydowe, audio, wideo i inne.

1. Wprowadzenie

Format dokumentu w potocznym rozumieniu zazwyczaj kojarzony jest przez użytkowników z rozszerzeniem pliku, który wskazuje na aplikację potrzebną do jego uruchomienia. Takie pojmowanie jest oczywiście słuszne, lecz niepełne. Do precyzyjnego określenia jego zakresu niezbędne jest zatem zdefiniowanie dokumentu elektronicznego. W literaturze istnieje wiele sposobów jego definiowania, np. informatologiczne[1], prawne[2], czy administracyjne[3], jednak z technicznego punktu widzenia

[1] Dokument elektroniczny (komputerowy, cyfrowy) – dokument z sygnałem utrwalonym na nośniku elektronicznym. Mogą to być komunikaty rożnych języków naturalnych w subkodach akustycznych i graficznych, inne komunikaty graficzne jak obrazy, wykresy, animacje, filmy, komunikaty dźwiękowe – (Bojar 2002, s. 50).

[2] Dokument elektroniczny – każdy przedmiot lub zapis na komputerowym nośniku informacji, z którym jest wiązane określone prawo albo który ze względu na zawartą w nim treść stanowi dowód prawa, stosunku prawnego lub okoliczności mającej znaczenie prawne – *Kodeks Karny*, Dz.U. nr 140/1997, poz. 553.

[3] Dokument elektroniczny – zbiór danych wprowadzonych lub przechowywanych na dowolnym nośniku przez system informatyczny lub podobny układ, które mogą być odczytane lub wyświetlone przez osobę lub przez tego rodzaju system lub układ, a także wszelkiego rodzaju prezentacja i wszelkiego rodzaju przedstawienie tych danych w formie drukowanej lub innej – *Decyzja Komisji Europejskiej* nr 2004/563/WE.

odpowiednie wydaje się określenie zapisane w pierwszym arkuszu normy *Informacja i dokumentacja – zarządzanie dokumentami* (PN-ISO 15489-1:2006). Dokument elektroniczny jest tam zdefiniowany jako: *dokument istniejący w postaci elektronicznej, dostępny za pośrednictwem techniki komputerowej*. Definicja akcentuje zatem jego dwie dystynktywne cechy: trwałość i możliwość lektury. Istotne jest więc, aby dane były zapisane i przechowywane na elektronicznym nośniku i aby można je było odtworzyć w formie umożliwiającej ich lekturę. Nie ma przy tym znaczenia sam nośnik informacji (np. pendrive, dysk twardy, serwer, sieć) ani rodzaj zawartych danych (tekst, grafika, multimedia), istotny jest jedynie fakt, aby dane te mogły zostać wielokrotnie odtworzone i aby mógł zachodzić akt percepcji. Z praktycznego punktu widzenia warto poczynić jednak drobną dygresję: w niniejszym opracowaniu będą nas interesowały wyłącznie dokumenty w formie skończonej (zamkniętej), przeznaczone do publicznego rozpowszechniania – czyli publikacje elektroniczne. Takie ujęcie wychodzi naprzeciw bibliotecznemu pojmowaniu zbiorów i pozwoli uniknąć problemów, jakie rodzą nowe typy dokumentów (np. blogi, posty).

Przy bliższej analizie rozważanego pojęcia można wskazać na kilka jego cech szczegółowych. Każdy dokument elektroniczny składa sie z czterech abstrakcji: pierwszą jest zawartość, czyli jego warstwa znaczeniowa (semantyczna); drugą – struktura logiczna, odpowiadająca z jakich jednostek dokument jest złożony (oraz z jakimi powiązany); trzecią – struktura fizyczna (czyli **format danych**) oraz czwartą – tzw. kontekst, rozumiany jako metadane, określające w jakiej postaci dokument będzie prezentowany. Kluczowe znaczenie dla dalszych rozważań ma struktura fizyczna, którą można określić, jako wynik przetworzenia informacji zawartych w dokumencie elektronicznym na dane w układzie binarnym. Zasadniczym polem dalszych rozważań będą więc problemy kodowania i innych transformacji jakim podlega dokument oraz stosowane w tym zakresie standardy. Warto dodać, że operacje te są odbywają się w środowiskach języków programowania wysokiego poziomu, zazwyczaj w aplikacjach macierzystych (np. Adobe Acrobat). Z informatycznego punktu widzenia zagadnienia formatu danych można jeszcze podzielić się na dwie grupy: pierwsza obejmuje atrybuty fizyczne (np. algorytmy kompresji, kodowane), druga – atrybuty typograficzne (np. formatowanie, czcionka).

Jakkolwiek na użytek praktyki stworzono liczne typologie formatów danych, jednak brakuje schematu uniwersalnego. W przypadku bibliotek cyfrowych rolą taką może jednak pełnić wykaz typów MIME [MIME Media Types 2002], specyfikacja zarządzana przez Internet Assigned Numbers Authority (IANA) i wspierana przez Konsorcjum W3C. Standard ów odnosi sie wprawdzie do protokołów komunikacyjnych, lecz równolegle precyzyjnie definiuje także związane z nimi formaty i wskazuje na oficjalne specyfikacje. Wśród dziewięciu głównych grup wykaz MIME wylicza m.in. formaty: dźwiękowe (audio), graficzne (image), tekstowe (text), wideo (video) oraz związane z programami (application). Z punktu widzenia repozytoriów cyfrowych warto zwrócić uwagę na kilka najważniejszych; w grupie tekstowej będą to m.in.: czysty tekst (plain text), HTML i XML; w grupie audio: MPEG (np. MP3), Windows Media Audio, RealAudio i WAV; w grupie graficznych: GIF, JPG, PNG,

TIFF, DJVu; w grupie wideo: MPG-1, MP4, QuickTime, Windows Media Video; spośród innych (application): JavaScript, Adobe Flash, strumień (octet-stream) oraz PDF. Warto podkreślić, że wskazane specyfikacje odwołują się także do licznych standardów ogólnych (np. kodowania znaków, barwy i in.), które są odrębnie definiowane i opisane. Szczegółowych informacji na ich temat należy poszukiwać w odpowiednich źródłach (np. International Organization for Standardization, The Unicode Consortium, International Color Consortium i in.). Istotną rolę w poprawnym przetwarzaniu i interpretacji standardów przez systemy teleinformatyczne pełnią nadto osadzone w dokumentach metadane techniczne, które również mogą być odrębnie zdefiniowane, np. standardy zapisu informacji w plikach graficznych IPTC (International Press Telecommunications Council) [IPTC: http://www.iptc.org/] czy EXIF (Exchangeable Image File Format) [EXIF.org: http://exif.org/specifications.html]. Trendem ostatnich lat jest natomiast otwarte podejście do tej kwestii, czyli definiowanie kontenera metadanych w formie XML, co sprawia, że może być on dołączony do opisywanego obiektu na wiele sposobów i zawierać różne elementy; takie rozwiązanie przyjęto np. specyfikacji MPEG-7, dedykowanej dla opisu treści multimedialnej [MPEG-7 Overview: http://mpeg.chiariglione.org/standards/mpeg-7/mpeg-7.htm].

2. Formaty w cyklu tworzenia publikacji cyfrowej

W praktyce bibliotek cyfrowych wybór formatu zależy od kilku kwestii, wśród których czołową rolę odgrywa stopień przetworzenia i przeznaczenie materiału. Zazwyczaj wyróżnia się trzy etapy cyklu przetworzenia i odpowiadające im formaty. Na etapie pierwszym, który obejmuje: przygotowanie, digitalizację, korekty, montaż, obróbkę i porządkowanie materiał podlega wielu przekształceniom i jest zachowywany w rozmaitych formatach **pośrednich** (przeważnie natywnych); istotną rolę odgrywa tu nie tyle sam format ile poprawny system wersjonowania plików i ich notacji (nazewnictwa). W etapie drugim finalny wersja przetworzonych i zmontowanych plików staje się podstawą wersji archiwalnej (ang. master format), którą należy przygotować w odpowiednim formacie **archiwalnym**. Podstawową funkcją tej kopii jest długotrwałe przechowywanie publikacji z zachowaniem możliwie najwyższych parametrów. Na etapie trzecim w oparciu o kopię master tworzy się publikację w wersji przeznaczonej do prezentacji w sieci, czyli publikację w formacie **prezencyjnym** (ang. end-user format). Ostatnia wersja jest zazwyczaj owocem kompromisu pomiędzy trzema parametrami: niewielkimi rozmiarami, akceptowalną jakością i wysokim poziomem zorganizowania. Aby osiągnąć ów kompromis należy skorzystać z któregoś z formatów oferujących skuteczne algorytmy kompresji stratnej (np. JPG dla grafiki rastrowej). Należy podkreślić, że dobór formatu prezencyjnego jest zadaniem dość trudnym, gdyż uzyskane efekty zależą od rodzaju publikacji, w szczególności takich jej parametrów jak: ilość, głębia i rozmieszczenie kolorów, rozmiary fizyczne, rodzaj prezentowanej treści, a nawet rodzaju użytych znaków (np. krój czcionki).

3. Formaty archiwalne

Formaty archiwalne (FA) w porównaniu do formatów prezencyjnych nie są ani zbyt skompilowane, ani nie wymagają większej wiedzy na etapie użytkowania. Z drugiej strony jedynie kilka standardów aspiruje do pełnienia tej roli (głownie PDF/A i TIFF). Jeśli – przypomnijmy – rolą kopii master jest długotrwałe przechowywanie publikacji z zachowaniem możliwie najwyższych parametrów, stąd łatwo o wniosek, że jedną ważną cechą jest wierne odwzorowanie oryginału. Aby osiągnąć ten cel deponuje się zwykle pliki archiwalne z zastosowaniem maksymalnych atrybutów, częstokroć stosując wielkości nadmiarowe. Plastycznie ilustruje to przykład grafiki rastrowej: jakkolwiek maksymalna głębia dla w modelu RGB wynosi 24 bity (true color), skany do celów archiwalnych wykonuje się w głębi 36 lub 48-bitowej, stosując jednocześnie wyższą rozdzielczość (np. nie 300, lecz 600 lpi [line per inch]); identyczne podejście przyjmuje się zresztą do dźwięku i obrazy (próbkowanie nadmiarowe) oraz innych rodzajów danych. Efektem zastosowania wspomnianych danych nadmiarowych są gigantyczne rozmiary plików master, gdyż ich objętość rośnie bądź proporcjonalnie do zmiany parametrów (np. głębi lub próbkowania), bądź proporcjonalnie do ich kwadratu, np. gdy zwiększamy rozdzielczość, a nawet czasem wykładniczo. Problem objętości nabiera szczególnego znaczenia, jeśli uwzględnić fakt, że zachowanie wysokiej jakości pociąga za sobą konieczność użycia plików bez kompresji lub stosowania algorytmów kompresji bezstratnej. W efekcie kopie archiwalne mają bardzo duże objętości i do przechowywania wymagają specjalnych urządzeń (zwykle macierzy dyskowych). Wymagania stawiane kopiom archiwalnym szczegółowo opisują różne dokumenty, np. *Decyzja Nr 25 Naczelnego Dyrektora Archiwów Państwowych* [2005] lub odpowiednie fragmenty dokumentacji PREMIS, rozwijanej przez OCLC i Bibliotekę Kongresu USA [PREMIS maintenance activity: http://www.loc.gov/standards/premis/], szczególnie wprowadzenie i słownik [Caplan 2009; *Data Dictionary for Preservation Metadata* 2011].

Wśród kilku formatów stosowanych do zachowywania macierzy archiwalnych główną rolę dla grafiki rastrowej pełni specyfikacja **TIF** (Tagged Image File Format). O jej przydatności zadecydowały następujące cechy formatu: obsługa wszystkich rodzajów głębi i przestrzeni kolorów (RGB, CMYK), możliwość bezstratnej kompresji (algorytm LZW), możliwość przechowywania ścieżek i kanałów alfa oraz profili koloru i komentarzy tekstowych, a nade wszystko to, że jest ona standardem międzyplatformowym [Specyfikacja TIF rev. 6.0: http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf]. Jedynym problemem, jaki pojawia się przy okazji stosowania TIF, jest pytanie czy stosować pliki wielostronicowe? Praktyka uczy jednak, aby nie korzystać z tej możliwości, gdyż na dłuższą metę może to powodować błędy w przetwarzaniu, nie wspominając o problemach z ekstrakcją pojedynczych stron. Jedynym mankamentem stosowania TIF jest zatem niski poziom zorganizowania macierzy (każdy plik osobno). Aby zapobiec możliwym problemom należy przyłożyć wiele uwagi do właściwej notacji i systemu nazewnictwa, gdyż muszą one być poprawnie obsługiwane w różnych systemach (Windows, Unix, MacOs i in.). Należy więc zadbać, aby

każda jednostka była zdeponowana w odrębnym folderze i aby mechaniczne sortowanie plików było zgodne z naturalną kolejnością stron finalnej publikacji. Druga – i bodaj ważniejsza kwestia – wiąże się z wykazem dozwolonych znaków. Ze względów bezpieczeństwa do nazywania plików należy stosować tylko znaki z bezpiecznej puli. W praktyce są to litery alfabetu łacińskiego (najlepiej małe), cyfry oraz znaki „-" (minus) i „_" (podkreślenie). Poprawne sortowanie z kolei wymaga, aby nazwy plików były równej długości, a indeks cyfrowy był uzupełniony zerami (np.: gazeta_polska_1936_033_00014).

Drugim formatem, który kwalifikuje się do długoterminowego przechowywania danych jest **PDF/A** (Portable Document Format w wersji A), czyli specjalna implementacja PDF uregulowana w 2005 r. normą ISO 32000-1 [PDF Standards: http://pdf.editme.com/PDFA] oraz jej polski odpowiednik [*Zarządzanie dokumentacją. PN-ISO 19005-1* 2007). Plik PDF pełni w tym zastosowaniu rolę kontenera, tzn. pozwala na zaimportowanie i osadzenie w pliku dowolnych danych (z wyjątkiem audio/wideo). Zasadniczo stosuje się go w dwu przypadkach: do grafiki rastrowej zamiast TIF (wówczas pliki zyskują wyższy poziom organizacji – strony, bookmarki) oraz dla dokumentów testowych (najczęściej biurowych, w tym listów). W ostatnim zastosowaniu PDF/A sprawdza się doskonale, gdyż podczas importu zachowuje wszelkie niezbędne komponenty (w tym np. użyte czcionki i obrazy), gwarantując, że będą one miały identyczny wygląd nawet na innej platformie. Jest to wyjście naprzeciw istotnemu problemowi zgodności wersji dokumentów biurowych, które – jak uczy doświadczenie – z uwagi na znaczną labilność wersji w ramach tego samego programu (np. MS Word) i dużej liczbie formatów natywnych (różne edytory) są bardzo trudne do poprawnej archiwizacji. Jedyny problem, jaki dotyczy zastosowania PDF/A wiąże się z odzyskiem dokumentu w trybie do edycji. Operację taką można bezstratnie wykonać dla formatów: TIF, JPG, PNG, TXT, RTF, Word, XML, HTML; w przypadku innych dokumentów będzie zachodzić konieczność konwersji, która może prowadzić do utraty części informacji (np. formatowania).

Archiwizacja innych rodzajów danych, w szczególności multimediów wymaga odrębnych standardów. W przypadku profesjonalnym filmów właściwym rozwiązaniem jest format **DPX** (Digital Picture Exchange), który zachowuje maksymalne parametry ścieżek, m.in. informacje dla każdego kanału o gęstości optycznej w 10 bitach i pełną informację o użytych barwach; jego wadą są jednak bardzo duże pliki wyjściowe, wymagające dedykowanych nośników. Standard DPX został opracowany przez Society of Motion Picture and Television Engineers (SMPTE) i opublikowany w 2003 r. jako norma ANSI/SMPTE (268M-2003) [Cineon Image File Format Draft: http://www.cineon.com/ff_draft.php]. W archiwach dźwiękowych można natomiast użyć standardu **ALC** (MPEG-4 Audio Lossless Coding), który opisuje wydana w latach 2005-2006 norma ISO/IEC 14496-3. W przypadku z archiwów z multimediami, dla których wymagania DPX nie mogą być spełnione rolę formatu master dla multimediów mogą pełnić popularne kontenery wideo lub audio, np. **AVI** (Audio Video Interleave) lub **WAV** (wave form audio format) z użyciem kodeków bazstratnych (np. FLAC – Free Lossless Audio Codec) lub bez kompresji.

4. Formaty prezencyjne

Formaty prezencyjne (FP) są grupą największą i jednocześnie najbardziej znaną, gdyż to z nimi obcują bezpośrednio użytkownicy serwisów i repozytoriów. Zasadniczym celem FP jest prezentacja dokumentów w sieci Web, z czego wynikają dalsze konsekwencje. A zatem winny one: po pierwsze – być niezależne od platformy (Windows, MacOS, Unix, Android, inne); po wtóre – prezentować się identycznie (lub podobnie) na rozmaitych urządzeniach wyjściowych (ekrany, konsole, drukarki, projektory); po trzecie – mieć niewielkie rozmiary; po czwarte – oferować akceptowalną jakość oraz po piąte – cechować się wysokim poziomem zorganizowania (czyli powiązania poszczególnych stron, części). Dwie pierwsze cechy (międzyplatformowość i przenośność) sprawiają, że już na wstępie liczba możliwych FP staje ograniczona do wąskiej grupy specyfikacji zaprojektowanych *ad hoc* do celów wymiany i prezentacji danych (np. TXT, HTML, RTF, JPG, PDF). Ich poprawną obsługę gwarantują poszczególne platformy bądź poprzez dedykowane wersje wtyczek i programów, bądź wspierają wprost na poziomie sterowników i bibliotek. Z drugiej strony postulat ten powoduje, że z grupy FP należy wykluczyć niemal wszystkie formaty dedykowane dla aplikacji, czyli formaty natywne (np. PSD, DOCX). Trudniej natomiast osiągnąć jednocześnie dwa kolejne postulaty: niewielkie rozmiary i wysoką jakość; są to bowiem parametry antynomiczne, ponieważ wyższa jakość powoduje zwykle wzrost objętości, i odwrotnie: niższa – spadek. Ostatecznie wiec należy szukać kompromisu pomiędzy niewielkimi rozmiarami a akceptowalną jakością. Konsensus ten umożliwia zastosowanie odpowiednich algorytmów kompresji, które przeważnie mają charakter stratny (np. JPEG, MPEG). Znaczy to, że dokumentu zapisanego z kompresją stratną nie można bez uszczerbku przywrócić do formy pierwotnej. Mimo tej wady kompresja stratna stała się dominującym sposobem rozwiązywania wskazanego dylematu, zaś ciężar odtworzenia formy pierwotnej przeniesiono na format archiwalny. Ostatnim problemem FP jest postulat wysokiego poziomu zorganizowania; w praktyce chodzi o mechanizm nawigacyjny w obrębie dokumentu, jeśli jest on wieloczęściowy lub ma znaczne rozmiary. Problem jest rozwiązywany rozmaicie. W dokumentach tekstowych (TXT, HTML, RTF) lub hybrydowych (PDF) następstwo części jest niejako „zaszyte" w dokumencie lub realizowane metodą hiperłączy i zakładek. W przypadku grafiki rastrowej niezbędny jest mechanizm zewnętrzny, czyli zastosowanie obudowy HTLM (rozwiązanie przestarzałe) bądź użycie nawigacji skryptowej (JavaScript), która odwołuje się do nazw plików. Jeszcze inną metoda prezentacji możliwą do zastosowania dla dowolnych danych jest renderowanie dokumentu za pomocą odpowiednich wtyczek (np. Flash); sposób ten z uwagi na duże możliwości i skalowalność stał się szczególnie przydatny w portalach społecznościowych (np. FaceBook).

Biorąc po uwagę wcześniejsze ustalenia i wspomniane postulaty można zaproponować następującą typologię formatów prezencyjnych i wskazać ważniejsze specyfikacje używane przez biblioteki cyfrowe.

Tabela 1. Typy dokumentów

Typ dokumentu	Wybrane formaty prezencyjne i ich stosowanie w bibliotekach cyfrowych	
	Zalecane	Nie zalecane
Tekstowe – nieadnotowane	TXT (UTF-8)	TXT (inne kodowanie)
Tekstowe – adnotowane	HTML, RTF, ODT (UTF-8)	DOC, XLS, inne natywne
Graficzne – rastrowe	JPG, GIF, PNG, JPG2000, TIF (G4)	PSD, inne natywne
Graficzne – wektorowe	Konwersja na rastrowe	CDR, DWG, inne natywne
Hybrydowe	PDF, DJVU	
Audio	MP3, WMA, AAC, MOV, Ogg	Audio CD, WAV
Wideo	MPEG-2, MPEG-4, FLV, MOV, WMV	VCD, DVD, SVCD, DivX, AVI
Animacje	SWF, ActionScript	
Specjalne – mapy	Zoomify	
Specjalne – nuty	Sibelius	
Aplikacyjne	Konwersja na hybrydowe lub audio, wideo	Strumień binarny
Mobilne	LIT, EPUB, PRC, Adobe Digital, PBD	

4.1. Formaty tekstowe

Formaty tekstowe należą do najstarszej formy dokumentów, gdyż istnieją niemal od zarania ery komputerów. Najprostszą ich formą jest tzw. czysty tekst (plain text, **TXT**), czyli plik znakowy nie zawierający formatowania (inaczej: tekst nieadnotowany). Dokumenty tego rodzaju są nieskomplikowane, gdyż składają się wyłącznie ze znaków alfanumerycznych, przestankowych oraz dwu znaków sterujących (koniec wiersza, tabulacja). Ich zaletą jest mała objętość, gdyż liczba bajtów jest w przybliżeniu równa liczbie użytych znaków. Jedynym problemem jaki wiąże się z tymi plikami jest zagadnienie kodowania. Pierwotnie stosowano standard ASCI, gdzie znaki kodowano w 7 bitach (0-127), co było wytaczające dla języka angielskiego. Dla innych języków zgodnie z normą ISO/IEC 8859 (16 arkuszy) rozszerzano ASCI do 8 bitów i na dodatkowych pozycjach (128-255) definiowano odpowiednie diakrytyki (język polski był zdefiniowany w arkuszu 2 [tzw. latin 2]). Równolegle funkcjonowały też inne standardy kodowania (np. IBM, Windows). Stan ów był bardzo niewygodny i już w 1991 r. podjęto prace harmonizacyjne. Aktualnie niekwestionowanym standardem obejmującym wszystkie języki świata jest UNICODE (norma ISO/IEC 10646) [JTC1/SC2/WG2 – ISO/IEC 10646 – UCS – http://std.dkuug.dk/JTC1/SC2/WG2/]. Standard pozwala zakodować wszystkie języki świata i wykorzystuje do tego celu zmienną liczbę bajtów (7 do 32). Najczęściej wykorzystywana jest tablica UTF-8; standard ten od połowy pierwszej dekady XXI wieku zagościł w także w systemach operacyjnych i popularnych aplikacjach biurowych wypierając inne specy-

fikacje. Współcześnie czysty tekst raczej rzadko wykorzystywany jest do prezentacji zasobów w bibliotekach cyfrowych (zwykle jako wariant publikacji); niemniej w starszych serwisach był formą dominującą (np. *Gutenberg Project* czy *Internet Archive*).

Poszerzoną wersją czystego tekstu jest tekst adnotowany (text enriched), czyli wzbogacony o możliwości formatowania lub inne funkcjonalności. Taki zestaw opcji umożliwiają w praktyce wszystkie edytory teksów, np. dostępne w popularnych pakietach: MS Office (doc, docx), OpenOffice (sxw), WordPerfect (wpd), Works (doc). Jednak żaden z wymienionych formatów nie aspiruje do miana standardu prezencyjnego, gdyż każdy ma charakter natywny, tzn. w pełni poprawnie funkcjonuje jedynie w określonym programie, a nierzadko wersji. Na potrzeby wymiany powstała specyfikacja **RTF** (Rich Text Format) – opracowana w 1987 r. przez Microsoft i sukcesywnie rozwijana; najnowszą wersję standardu (1.9.1) ogłoszono w 2008 r. przy okazji debiutu edytora Word 2007 [Rich Text Format (RTF) Specification, version 1.9.1 – http://www.microsoft.com/]. RTF w odróżnieniu od czystego tekstu definiuje i przechowuje w plikach m.in. informację o użytych czcionkach, ich ułożeniu, atrybutach (np. kolorze), rozłożeniu akapitów, marginesów, tabel oraz parametry osadzonych obiektów (np. rysunków). Użyty tekst może być zapisany w dowolnej stronie kodowej. Mimo wskazanych zalet trudno zaliczyć RTF do formatów dobrze funkcjonujących w internecie. Złożyły się na to cztery powody: pliki RTF mają stosunkowo duże objętości (10-15 razy większe niż TXT); obiekty inne niż tekst (tabele, grafika) są w często zniekształcane podczas odczytu; pliki RTF nie są interpretowane przez przeglądarki oraz – co najważniejsze – mogą zawierać elementy potencjalnie niebezpieczne (makra, wirusy). Drugim formatem o zbliżonej funkcjonalności jest **ODT** (OpenDocument [Text]) – czyli otwarty formatu plików pakietów biurowych [Online Community for the OpenDocument OASIS Standard – http://opendocument.xml.org/], rozwijany przez niezależną organizację OASIS (Organization for the Advancement of Structured Information Standards) i uregulowany normą ISO/IEC 26300: 2006. Specyfikacja wyrosła na bazie XML-owego formatu pakietu OpenOffice i oferuje zestaw opcji zbliżony do RTF. ODT jest specyfikacją stosunkowo młodą, lecz bliższa analiza przekonuje, że jest on obciążony licznymi wadami (podobnie jak RTF), to powoduje, że jego zastosowanie w sieci Web jest ograniczone.

Do grupy adnotowanych formatów tekstowych należy także zaliczyć hipertekst, np. HTML (HyperText Markup Language) [HTML 4.01 Specification – http://www.w3.org/TR/html4/] wraz z wariantami (np. arkusze stylów – CSS, XSL) i języki pochodne, szczególnie XHTML (Extensible Hyper Text Markup Language) oraz inne specyfikacje towarzyszące, którymi zarządza World Wide Web Consortium (W3C) [World Wide Web Consortium (W3C) – http://www.w3.org/]; nadto wersje specjalne, np. dedykowane dla urządzeń mobilnych, np. WML (Wireless Markup Language) nadzorowane przez Open Mobile Alliance (OMA) [Open Mobile Alliance – http://www.openmobilealliance.org/]. Wspólnym rysem hipertekstu jest poziom uboższy względem RTF poziom formatowania (np. brak justowania, dzielenia wyrazów, zaawansowanych tabel, zagnieżdżeń), a jednocześnie lepsza obsługa elementów osadzonych, np. grafiki. Hipertekst nie jest jednak obciążony wadami RTF: pliki są

zawsze małe, mniej podatne na przenoszenie wirusów i – co najważniejsze – obsługiwane wprost przez przeglądarki. Jedynym mankamentem jest jedynie labilność wyglądu strony, gdyż przeglądarki w zależności od lokalnych ustawień, wielkości ekranu i użytych komponentów (np. czcionki) inaczej wyświetlają zawartość. Czasami jest to cecha korzystna (tekst dopasowuje się do możliwości ekranu), lecz przeważają aspekty negatywne, gdyż zmienność parametrów i geometrii sprawia, że trudno zapewnić wiarygodność udostępnianego dokumentu. Strony w języku HTML i jego odmianach już od początku XXI są wieku niemal wyłącznie pisane z zastosowaniem kodowania UTF-8, często używają języków skryptowych (np. JavaScript) i wykorzystują XML do reprezentacji swej struktury.

W odpowiedzi na powyższe słabości dokumentów tekstowych (w szczególności w zastosowaniu do e-administracji) [Matejko; Januszewski 2006; Rozporządzenie Rady Ministrów z dnia 11 października 2005 r.] upowszechniło się w ostatnich latach specyficzne podejście do tej kwestii. W nowoczesnych systemach teleinformacyjnych dokumenty przechowuje się w **formacie strukturalnym** i na życzenie generuje odpowiedni format prezencyjny. W takim systemie pierwotny dokument tekstowy zapisany jest w trzech plikach: dane (XML), struktura (XSD), wygląd (XSLT), na podstawie których na życzenie jest generowany plik finalny: PDF, HTML, RTF lub DOC. Zaletą tego rozwiązania jest duża otwartość na nowe standardy prezencyjne, a nade wszystko oddzielenie danych od warstwy prezentacji, którą najczęściej jest urzędowy formularz.

4.2. Formaty graficzne

Grafikę w systemach informatycznych można reprezentować w formie wektorowej bądź rastrowej. W pierwszym modelu obiekty są tworzone z figur i krzywych matematycznych, w drugim – rysowane za pomocą punktów „pikseli" na płaszczyźnie w układzie kartezjańskim. Pierwszy rodzaj charakteryzuje się tym, że obrazy można dowolnie powiększać, bez utraty jakości; w drugim zaś granicę ową określa rozdzielczość, gdyż po przekroczeniu pewnej wartości granicznej powiększenia obraz staje się coraz mniej czytelny. Łatwo wydedukować, że każdy z tych sposobów do zakodowania informacji wymaga innych ilości danych. Zdecydowanie mniej w danych definiujących wymagają obrazy wektorowe, gdyż ilość parametrów niezbędnych do zdefiniowania figur geometrycznych jest niewielka (np. dla okręgu: współrzędne środka, promień, grubość krawędzi i jej kolor). W modelu rastrowym natomiast zakodować należy współrzędne i atrybuty każdego piksela – stąd rozmiar pliku jest większy i rośnie proporcjonalnie do kwadratu rozdzielczości oraz wprost proporcjonalnie do innych parametrów (np. głębi koloru). Najistotniejsza jednak różnica pomiędzy grafiką wektorową a rastrową dotyczy warstwy prezentacji, a ściślej przenośności. Obrazy rastrowe dają się łatwo zapisać w formatach wymiennych i można je wyświetlać w rozmaitych programach, z kolei obrazy wektorowe funkcjonują poprawnie jedynie w aplikacjach macierzystych, w których zostały stworzone (np. CorelDraw i AutoCad), czyli w formatach natywnych (odpowiednio: CDR, DWG). Cecha ta powodu-

je, że na potrzeby prezentacji wykorzystuje się wyłącznie grafikę rastrową, zaś finalne wersje obrazów wektorowych konwertuje do postaci rastrowej.

Głównymi parametrami opisującymi grafikę rastrową są: rozmiary obrazu (wysokość i szerokość liczona w pikselach); rozdzielczość (czyli liczba pikseli przypadająca na jednostkę długości), liczona zwykle w pikselach na cal (ppi – pixels per inch) oraz parametr głębi koloru (color depth) określający paletę użytych kolorów, czyli liczbę bitów potrzebną do reprezentacji koloru danego piksela. W praktyce stosuje się cztery modele głębi: 1-bitową (czerń i biel), 2-8-bitową (odcienie szarości [4-256 odcieni]), zaś kolor można reprezentować w palecie indeksowanej (2-8 bitów [4-256 kolorów]) lub w modelu RGB (6-24 bitów [64-16777216 kolorów]. Zagadnienia koloru obejmują więcej parametrów, które można analizować w ramach stosowanej przestrzeni barwnej. Dominującym sposobem reprezentacji koloru jest model RGB, czyli barwa addytywna. W przestrzeni tej każdy kolor uzyskiwany jest z mieszania trzech barw podstawowych (kanałów): R – red (czerwonej), G – green (zielonej) i B – blue (niebieskiej). Jeżeli głębię koloru dla każdego kanału wynosi 8 bitów, to pozwoli to na uzyskanie $2^{8+8+8}=2^{24}$ barw o 256 poziomach jasności, czyli ponad 16 milionów (true color). Powszechny jest także model CMYK, gdzie kolor uzyskiwany metodą subtraktywną, czyli nakładania barw: C – cyan (turkus), M – magenta (karmazyn), Y– yellow (żółty), K – Black (czarny). Obie przestrzenie są równie popularne, lecz mają odrębny zakres stosowania: RGB zwykle wykorzystuje się do prezentacji na monitorach, CMYK natomiast w działalności poligraficznej. Dodatkowym atutem tych rozwiązań jest możliwość bezstratnej konwersji barw RGB/CMYK, np. dla głębi 24-bitowej wg wzoru: R=255-C, G=255-M, B=255-Y (K – nie podlega konwersji). Prócz przestrzeni RGB/CMYK barwę opisuje jeszcze uniwersalny model HSV (Hue, Saturation, Value), gdzie: H – hue (barwa światła), S – saturation (nasycenie koloru), V – Value (jasność, czasami określana jako B – brightness). Model ten nawiązuje do sposobu, w jakim barwę widzi ludzkie oko. H – to odcień, czyli cecha jakościowa barwy (np. zielony, niebieski, różowy), S – udział barwy czystej w kolorze (np. jasnoniebieski); V – intensywność barwy w zależności od wartości światła odbitego. Według tego modelu wszelkie barwy wywodzą się ze światła białego, gdzie część widma zostaje wchłonięta, a część odbita od oświetlanych przedmiotów. Prócz wyliczonych uniwersalnych abstrakcji można jeszcze wskazać kilka innych parametrów o węższym zakresie stosowania, np. kanał alfa (czyli barwa przezroczysta), typowa dla palety indeksowanej) oraz szereg opcji używanych do manipulacji na barwie, np. kontrast czy korekcja gamma.

Wśród licznych formatów stosowanych do prezentacji grafiki rastrowej (istnieje ich ok. 250) [http://en.wikipedia.org/wiki/Comparison_of_graphics_file_formats] szczególnie użyteczne w różnych zastosowaniach są cztery specyfikacje: JPEG, GIF, PNG i TIF. Najbardziej popularny jest pierwszy z nich. **JPEG** (lub JPG) jest formatem stratnym, który sprawdza się szczególnie dla obiektów złożonych z dużej liczby barw i dużej liczby szczegółów (np. pejzaże i portrety), charakteryzujących się płynnymi przejściami barw oraz brakiem lub małą ilością ostrych krawędzi i drobnych detali. Największą zaletą formatu jest mały rozmiar pliku wyjściowego przy zachowa-

niu akceptowalnej jakości oraz obsługa pełnej przestrzeni barwnej (true color); wadą zaś jednokierunkowa utrata jakości podczas kompresji. Algorytm kompresji JPG przekształca obraz w trzech etapach: najpierw każdy piksel (dla kolejnych kanałów RGB z osobna) zostaje przekształcony na skalę jasności (luminancję) oraz 2 barwy (chrominancje); następnie obraz jest dzielony na bloki 8x8 pikseli i poddany transformacji DTC (discrete cosine transform), która sprawia, że pikselom zostają przypisane uśrednione wartości barw (zmiennoprzecinkowe); na końcu następuje etap kwantyzacji, czyli transformacja wartości zmiennoprzecinkowych na całkowite (tylko tu powstają straty). Warto zauważyć, że skuteczność algorytmu zależy od zawartości obrazu, stąd niemal w każdym przypadku należy empirycznie dobrać optymalne ustawienia. W praktyce proces konwersji sprowadza się do doboru kompresji (0-100%) i obserwacji utraty jakości (granica akceptacji); jedynie w zastosowaniach profesjonalnych pomocne mogą być inne parametry: czynnik próbkowania i optymalizacja tabeli Huffmana – czyli parametry drugiego kroku kompresji oraz wygładzanie (antyaliasing) i progresja (sukcesywne wyświetlanie obrazu w trakcie ładowania). Warto dodać, że JPG doskonale obsługują popularne schematy metadanych (EXIF, IPTC i XMP), zaś ich wadą jest brak obsługi kanału alfa [*JPEG image compression FAQ* 1999]. Standard JPG ma wśród innych specyfikacji pozycję ustabilizowaną. Prace nad nim rozpoczęto stosunkowo dawno, bo już w 1983 r.; najpierw pod egidą ISO, potem w ramach specjalnej grupy (Joint Photographic Experts Group) [Joint Photographic Experts Group – http://www.jpeg.org/], gdzie dostępne są wszystkie specyfikacje i raporty techniczne. Pierwsza specyfikacja powstawała w latach 1991-1998 (ISO/IEC IS 10918-1), zaś ostatnia, w której zawarto szereg poprawek ogłoszono w 1999 r. (ISO/IEC IS 14495-1). Rok później na bazie dotychczasowych doświadczeń opracowany nowy standard **JPG2000** (norma ISO/IEC 15444-1), w którym wykorzystano nową metodę transformacji (DWT – kafelkową) oraz wbudowano możliwość stosowania kompresji bezstratnej. Format ten nie zyskał jednak większej popularności, gdyż opcja kompresji bezstratnej tylko nieznacznie przewyższa popularny algorytm LZW stosowany np. w formacie TIFF.

Drugim popularnym formatem rastrowym jest **GIF** (Graphics Interchange Format), którego wykorzystanie daje najlepsze efekty w obrazach złożonych z małej liczby barw i zawierających duże przestrzenie jednokolorowe (np. obrazy konwertowane z grafiki wektorowej); sprawdza się także przy obiektach zawierających ostre krawędzie i drobne detale (np. czcionki). GIF uznawany jest za format stratny, co tylko częściowo jest prawdą. Stratność wynika z faktu, że obsługuje on tylko kolor 8-bitowy (256 barw) w jednym bloku (ramce). Teoretyczne możliwe jest wiec podzielenie obrazu na bardzo małe bloki, co prowadziłoby do poszerzenia palety. Jednak większość programów graficznych stosuje jedynie jedną ramkę, co ostatecznie sprawia, że w praktyce GIF obsługuje maksymalnie 256 barw. Poza ograniczeniem palety GIF nie zniekształca obrazu, gdyż w obrębie bloku kompresja ma charakter bezstratny (LZW). Ciekawą jego własnością jest obsługa przezroczystości, czyli kanału alfa (transparent background color). Funkcjonalność ta sprawia, że w ramach formatu można tworzyć proste animacje (animowane GIFy) i obrazy z przezroczystym tłem (np. dla znaków

wodnych). Druga cecha: przeplot (interlacing) – pozwala wyświetlać grafikę warstwa po warstwie (4 fazy), w praktyce sprawia to, że na początku obraz wydaje się rozmyty, a w trakcie ładowania nabiera ostrości. GIF został stworzony w 1987 r. przez firmę CompuServe, a w 1989 r. doczekał się aktualizacji, dodającej między innymi możliwości przechowywania w nim animacji (Gif89a) [Graphics Interchange Format, Version 89a – http://www.w3.org/Graphics/GIF/spec-gif89a.txt]. Przy opracowywaniu specyfikacji, firma zaniedbała problem użytego algorytmu kompresji LZW, który był chroniony patentami firm IBM oraz UniSys. W połowie lat 90. XX w. nastąpił regres, gdyż UniSys zapowiedziała, że będzie ścigać firmy używające ich chronionej technologii i pobierać od nich opłaty licencyjne. Stan niepewności trwał aż do 2006 r., gdy ostatecznie patenty wygasły. Równolegle w odpowiedzi na istniejące zagrożenie podjęto prace nad następcą GIFa – formatem **PNG**. Nowy standard oferuje funkcjonalność zbliżoną do GIF, choć na lego korzyść przemawia możliwość stosowania pełnej palety barwnej (true color). W praktyce jednak PNG nie zastąpił GIFa, ani nie zagroził JPG i TIF, gdyż oferowane opcje kompresji stratnej dają nieco gorsze efekty niż GIF/JPG, zaś algorytm bezstratny okazuje się tylko nieznacznie lepszy od TIF. PNG sprawdza się jednak na wielu specyficznych obszarach (np. zrzuty ekranowe lub obiekty o dużych rozmiarach), jest ponadto formatem domyślnym w środowisku Linux i rekomendowanym przez W3C do zastosowań internetowych. Od 1995 r. nad rozwojem formatu czuwa Internet Engineering Steering Group (IESG), wspierana przez Konsorcjum W3C; od 2003 r. PNG został uregulowany normą ISO/IEC 15948 [Portable Network Graphics. Specification (Second Edition) – http://www.w3.org/TR/2003/REC-PNG-20031110/].

W grupie formatów prezencyjnych o wąskim zakresie stosowania należy także umieścić **TIF (G4)**. Został on wprawdzie zaliczony do formatów archiwalnych (por. omówienie w rozdz. 3), niemniej jego funkcjonalność jest znacznie szersza. W jednym zastosowaniu TIF okazuje się bezkonkurencyjny: obrazy czarno-białe (1-bitowe) z bezstratną kompresją CCITT Group 4 (w skrócie G4). W tym zastosowaniu obrazy są ok. 6-, 15-krotnie mniejsze od oryginału i jednoczenie znacznie mniejsze niż w innych formatach (GIF, PNG, JPG). Funkcjonalność ta sprawia, że TIF (G4) doskonale sprawdza się jako kontener na skany bitonalne. Warto jednak zaznaczyć, że uzyskanie dobrej jakości obrazu o głębi 1-bitowej możliwe jest bezpośrednio tylko dla grafiki kreskowej lub nieciągłotonalnej. W przypadku obrazów ciągłotonalnych (gdy przejście między barwami odbywa się stopniowo) należy podczas konwersji głębi zastosować któryś z algorytmów rozpraszania (ditheringu), np. algorytm Floyda-Steinberga, wzór lub rozpraszane adaptacyjne. W efekcie obrazy ciągłotonalne zostaną zamienione na odpowiednio zagęszczone kropki lub linie, które będą imitować pierwotne barwy. Dodajmy, że technika ta jest podstawą przygotowania obrazów do druku, z wykorzystaniem tzw. technologii rastra. Obrazy z rozproszeniem cechuje efekt mory (moiré), którą ludzkie oko postrzega jako deformację przy pewnych wielkościach powiększania obrazu, tzn. gdy między pikselami dochodzi do interferencji. Warto też dodać, że do oglądania obrazów tego typu należy używać przeglądarek, które obsługują opcję antyaliasingu (tj. programowego rozmycia), gdyż wzór rastra

składa się z małych ostrych elementów, które dla ludzkiego oka sprawiają wrażenie dyskomfortu i powodują, że obraz wydaje się blady i rozmyty.

4.3. Formaty hybrydowe

Formaty hybrydowe to grupa specyfikacji, która łączy cechy formatów tekstowych i graficznych. W praktyce można wskazać dwa takie rozwiązania: PDF i DJVu. Pierwsze w systematykach MIME należy do grupy aplikacyjnej, drugie do graficznej.

PDF (Portable Document Format) [dokumentacja na witrynie Adobe: http://www.adobe.com/pl/products/acrobat/standards.html oraz PDF Standards: http://pdf.editme.com/] jest implementacją języka PostScript, który z czasem wzbogacono o nowe możliwości (elementy hipertekstowe i aplikacyjne). W najszerszym sensie PDF to specyfikacja, która w zależności od wersji może pełnić rozmaite funkcje: prezentacji, przenoszenia, drukowania i przechowywania treści tekstowo-graficznych. Dokumenty PDF cechują się pełną niezależnością od platformy (w każdym środowisku prezentują się identycznie) oraz pełną niezależnością od źródła (zawierają w pliku wszelkie komponenty niezbędne do prawidłowego działania, jak czcionki i obrazy); mają nadto wbudowany mechanizm programowego powiększenia (zoom). Twórcą formatu była firma Adobe, która w latach 1993-2007 r. samodzielnie rozwijała jego kolejne wersje (od 1.1 do 1.7 i odpowiadające mu wersje Acrobata: 1.0-9.0), które sukcesywnie zyskiwały nowe funkcjonalności, np. podpis elektroniczny (wersja 1.3 – 2000), opcje 3D (1.6 – 2004); w 2008 r. zaszła istotna zmiana, gdyż firma ogłosiła publiczne otwarcie formatu i przekazała jego specyfikację organizacji Association for Information and Image Management (AIIM), która opracowała odpowiednią normę ISO 32000-1:2008 i przejęła obowiązki koordynacyjne. Pierwotnie zastosowaniem PDF była poligrafia i rozwijany z tą myślą format ewoluował do wersji X (PDF/X), która od 2001 r. stałą się podstawą normy ISO 15929 i ISO 15930. Kolejnym polem zastosowania było kontenerowanie, tj. długotrwałe przechowywanie danych – wersja X (PDF/X – ISO 19005: 2005) oraz potrzeby inżynieryjne – wersja E (PDF/E – ISO 24517: 2008). W 2010 r. powstała specyfikacja na potrzeby transakcyjne VT (PDF/VT – ISO 16612-2), czyli do zastosowań przy wymiana danych w formularzach PDF, np. w e-administracji. Wciąż w trakcie rozwoju jest natomiast najistotniejsza z punktu widzenia bibliotek cyfrowych – wersja PDF/UA (projekt ISO/DIS 14289-1) [PDF Universal Accessibility Committee: http://pdf.editme.com/PDFUA] – czyli PDF uniwersalny na potrzeby prezentacji (universal accessibility).

PDF jest formatem bardzo złożonym. W plikach może być zawarty zarówno tekst, jak i grafika (rastrowa, wektorowa), a także elementy aktywne. Specyfikacja potrafi bardzo precyzyjnie obsłużyć wszelkie związane z tym kwestie. Dokumenty PDF umożliwiają zatem pełną obsługę specyficznych takich problemów, jak: kodowanie znaków (w tym UTF-8), rodzaje czcionek (Type1, TrueType, OpenType), atrybuty graficzne (przestrzenie barwne, profile itp.), formularze, strukturę XML, metadane, podpis elektroniczny i in. Z punktu widzenia bibliotek cyfrowych warto zwrócić uwagę na trzy kwestie: PDFy powstałe na bazie dokumentów tekstowych (adnotowanych

lub nieadnotowanych) zyskują na wiarygodności i stają się pełni przenośne, bowiem pokonany został problem labilności wyglądu strony (na każdej platformie prezentują się identycznie); zaś dokumenty graficzne zyskują wyższy poziom zorganizowania (strony, bookmarki). Specyfikacja PDF w odróżnieniu od formatów tekstowych i graficznych stwarza też spore możliwości ochrony dostępu do treści (szyfrowanie, hasła, certyfikaty i in.), które sprawiają, że dokumenty te doskonale nadają się do serwowania treści udostępnianej w sposób limitowany. Powstaje też – podkreślmy – możliwość prezentowania grafiki wektorowej, bez ich konwersji do form rastrowych. Jedynym problemem, jaki może się pojawić w wyżej wymienionych zastosowaniach to wzrost objętości PDF względem formy pierwotnej. Aby zredukować objętość plików należy podczas konwersji zastosować specjalne opcje redukcji parametrów osadzonych grafik (np. dowsampling) lub skorzystać ze specjalnych narzędzi umożliwiających optymalizację pliku (optimize). PDF ma ponadto wiele innych funkcjonalności (np. warstwę komentarzy), jednak ich rola w bibliotekach cyfrowych jest niewielka. Odrębnym wątkiem związanym z plikami PDF jest ich edycja. Jakkolwiek jest ona możliwa w odniesieniu do tekstu i grafiki (Adobe Acrobat, Adobe Bridge), stosuje się ją raczej rzadko, i niemal wyłącznie w zastosowaniach poligraficznych. Na potrzeby bibliotek cyfrowych najczęściej generuje się pliki (zwykle drukarkami wirtualnymi) i w razie potrzeby dokonuje ich montowania. Czasami też wykonuje się edycję w trybie ograniczonym: dodawanie warstwy OCR, bookmarki, kadrowanie, metadane, obracanie i reformatowanie stron itp. Watro dodać, że w niektórych kręgach użytkowników zamiast PDF stosowany jest czysty **PostScript** (PS) [PostScrip language reference 1999]. Jest on szczególnie popularny wśród matematyków i chemików, czyli w środowiskach, które używają niestandardowych znaków oraz dedykowanych edytorów. Od lat prym w tej roli dzierży TeX (LaTeX) [LaTeX – A document preparation system – http://www.latex-project.org/] lub jego klony. Efektem pracy w tym środowisku są albo dokumenty w formacie natywnym (DVI), albo wyeksportowane do PostScriptu. Warto na zakończenie dodać, że aplikacje do tworzenia PDF dzielą się na dwie duże grupy: komercyjne, oparte na rozwiązaniach Adobe oraz bezpłatne, bazujące na kompilatorach PostScriptu (np. GhostScript) [Home Page for Ghostscript – http://pages.cs.wisc.edu/~ghost/].

Drugim standardem hybrydowym jest **DJVu** – technologia przeznaczona do kompresji stratnej dokumentów skanowanych. W odróżnieniu od PDF, który jest kontenerem, DJVu opiera się na metodzie segmentacji obrazu i wydzieleniu warstw; dodatkowo istnieje możliwość osadzenia niesformatowanego tekstu. W uproszczeniu proces kompresji DJVu polega na tym, że w zależności od typu skanu obraz, jest analizowany i dzielony na kilka warstw tła o niższej rozdzielczości i warstwę treści o wyższej rozdzielczości. Podczas wyświetlania warstwy nakładają się na siebie i tworzą iluzję ostrego obrazu. Istotną cechą omawianego formatu jest bardzo wysoki współczynnik kompresji, znacznie przewyższający inne algorytmy – co sprawia, że pliki DJVu osiągają rozmiary najmniejsze z możliwych (nawet 10-20 razy mniejsze od JPG czy GIF). W odróżnieniu od PDF używanie technologii DJVu jest znacznie trudniejsze, gdyż do uzyskania dobrego efektu wymaga określenia wielu parametrów (np. dotyczących

segmentacji), a także biegłego operowania profilami i innymi zmiennymi. Prace nad technologia DJVu rozpoczęto w 1996 r. w koncernie AT&T.; po 2000 r. jej rozwój odbywa się równolegle w dwu środowiskach: komercyjnym (LizardTech Inc., od 2008 Caminova, Inc.) oraz społecznym na licencji (projekt DJVuLibre na licencji GNU GPL) [DJVULIbre – http://djvu.sourceforge.net/], gdzie czołową rolę odgrywa jeden z twórców specyfikacji, były pracownik AT&T Léon Bottou. Dwutorowość rozwoju sprawia, że użytkownicy mają do wyboru w pełni funkcjonalne aplikacje zarówno komercyjne (np. DocumentExpress), jak i ich bezpłatne odpowiedniki (DJVuLibre) oraz szereg przeglądarek i narządzi [DJVU.org – http://djvu.org/resources/]. Szerzej na temat formatu DJVu poświęcony jest odrębny rozdział w niniejszej monografii.

4.4. Formaty audio

Dźwięk jest w systemach informatycznych reprezentuje sygnał dyskretny, który jest odwzorowaniem fali akustycznej. Wskazana transformacja jest realizowana w specjalnych urządzeniach (przetwornikach analogowo-cyfrowych), które ciągły przekształcają sygnał analogowy w nieciągły sygnał cyfrowy; zatem głównym parametrem charakteryzującym jakość dźwięku cyfrowego jest częstotliwość próbkowania (sampling). Parametr ten winien być co najmniej dwa razy wyższy od pasma słyszalnego. Ponieważ ludzkie ucho słyszy dźwięki w zakresie ok. 16 Hz do 18 kHz, próbkowanie winno wynosić co najmniej 36 kHz; w praktyce dla jakości CD stosuje się wartość nieco wyższą – 44,1 kHz. Drugim parametrem, który w jeszcze większym stopniu wpływa na jakość odtwarzanego dźwięku, jest przepływność (bitrate). Parametr ten wskazuje na liczę bitów potrzebną do cyfrowej reprezentacji jednej sekundy nagrania i jest wyrażany w bitach na sekundę (b/s=bps) i ich krotnościach (kbps, Mbps); im wyższa jego wartość, tym lepsza jakość (np. dla MP3 przyjmuje się najczęściej 128 kBps, rzadziej 192 kBps). Przy zbyt agresywnej kompresji, co skutkuje zmniejszeniem wartości bitrate sygnał staje się zniekształcony przez tzw. artefakty, czyli dźwięki, których nie ma w pliku oryginalnym. Aby obejść ów niekorzystny efekt w większości formatów stosuje się zmienną bitrate (VBR – variable bit rate), co wpływa na zmniejszenie rozmiarów plików. W fali dźwiękowej (i jej reprezentacji cyfrowej) możemy wyróżnić jeszcze inne składniki widma dźwięku i ich jednostki, lecz nie ich związek z formatami zapisu jest niewielki; należą tu m.in.: wysokość (Hz), głośność (dB), składowe, wysokość i barwa; podobnie niewielki wpływ wywierają standardy i modele, które wpływają jednie na liczbę równoległych ścieżek, np. mono (1), stereo (2), Dolby surround (4-5), Dolby Digital (5+1) i in. Z uwagi na znaczne rozmiary nieskompresowanego pliku z sygnałem dźwiękowym większość formatów stosuje rozmaite algorytmy kompresji (kodeki). W efekcie na opis formatu audio zawsze składa się kontener multimedialny (container) oraz stosowany kodek (codec).

Wśród ok. 40 formatów do zapisu dźwięku [Comparison of audio formats – http://en.wikipedia.org/wiki/Comparison_of_audio_formats] najpopularniejszym formatem stratnym audio stosowanym w internecie jest **MP3** (MPEG-1 Audio Layer 3), który doskonale sprawdza się w każdym zastosowaniu. Standard ów charakteryzu-

je się przede wszystkim znakomitym współczynnikiem kompresji, który w porównaniu z WAV jest 12-, 13-krotnie mniejszy. Istota kodowania MPEG jest dość złożona, w dużym uproszczeniu polega ona dzieleniu sygnału na tzw. ramki (frame), w obrębie których dokonywana jest kompresja (maskowanie i usuwanie pewnych danych); ponadto kodowane są różnice między sąsiednimi ramkami. Sam mechanizm kompresji MPEG był stworzony z myślą o filmach, ale praktyka pokazała, że sprawdza się także doskonale w funkcji kodeka audio. Prace nad formatem rozpoczęto w 1991 r.; głównym promotorem działań była grupa Motion Picture Expert Group (MPEG) [MPEG.org – http://www.mpeg.org/], wspierana przez naukowców z Instytutu Frauenhofera z Erlangen [Fraunhofer-Institut für Sichere Informationstechnologie – http://www.sit.fraunhofer.de/]. W krótkim czasie standard okrzepł i już w 1993 r. został utrwalony jako norma ISO/IEC 11172-3. Podejmowane później próby prac na następcą standardu (np. MP3Pro) nie powiodły się. Jedynym wyjątkiem, który ostatnio zanotował umiarkowane powodzenie był powstały w 1997 r. format **AAC** (Advanced Audio Coding), oferujący lepszą jakość dźwięku przy podobnym rozmiarze danych. Standard jest oparty na kompresji MPEG-4 Part 14 i opisany w normie ISO/IEC 13818-7. Znalazł on szerokie zastosowanie w urządzeniach mobilnych (np. telefonach komórkowych, iPhonie, iPodzie) oraz konsolach. W wyniku sporów patentowych między Microsoftem a Instytutu Frauenhofera powstał trzeci popularny format **WMA** (Windows Media Audio), bardzo zbliżony pod względem funkcjonalności do MP3. Niezależnie testy udowadniają, że przy niskim bitrate (do 96 kBps) jakość dźwięku jest lepsza niż MP3, zaś powyżej tej granicy standard MP3 daje lepsze efekty.

W odrębny sposób odtwarzane są tzw. media strumieniowe, czyli takie, które oferują dostęp w czasie rzeczywistym dostęp do medium ciągłego (analogicznie jak transmisja radiowa). Wyróżnia się przy tym dwa rodzaje dostępu: na żywo (ang. live), gdy transmitowane są dane, które od razu docierają do wszystkich użytkowników, np. transmisja na żywo lub na życzenie (ang. demand), gdzie serwer ma bazę zarejestrowanych danych, zaś użytkownik, wybiera sobie, który pokaz multimedialny od jakiego momentu chce oglądać. Do obsługi takich funkcjonalności stworzono specjalne formaty. Najpopularniejszym jest opracowany w 1995 r. przez firmę RealNetworks specyfikacja **RealAudio (RA, RM)**, który doskonale sprawdza się na słabych łączach internetowych. Wśród innych tego rodzaju warto wspomnieć o uniwersalnym formacie audio/wideo **QuickTime** rozwijaną przez Apple czy opracowanym przez Xiph.org kontenerze strumieniowym **Ogg**, który pozwala na wykorzystanie rozmaitych kodeków (większą popularnością cieszą się: uniwersalny kodek Vorbis oraz dedykowany dla mowy ogólnego Speex).

Trzecią grupą są formaty służące do rejestracji i odtwarzania dźwięków tworzonych sztucznie w środowisku komputerowym (w tym sztucznej mowy). Są one – mówiąc metaforyczne – dźwiękowym odpowiednikiem kolorów indeksowanych. Liderem wśród formatów tego typu jest stworzony w 1983 r. format **MIDI** skrót od (Musical Instrument Digital Interface) [MIDI Manufacturers Association – http://www.midi.org/]. Pliki tego rodzaju są bardzo małe i prócz kodowania dźwięku (próbki) zawie-

rają komendy sterujące, które pozwalają na współpracę z zewnętrznymi urządzeniami muzycznych.

4.5. Formaty wideo

Cyfrowa reprezentacja wideo (ruchomych obrazów) stanowi połączenie technologii reprezentacji grafiki rastrowej z problemami transmisji oraz technologią audio, która może być kodowana wspólnie lub na niezależnych ścieżkach. Wśród wielu pojęć i parametrów charakterystycznych dla tego obszaru można wskazać: przepływność (bitrate), pełniącą analogiczną funkcję jak dla dźwięku [por. 4.4], rozdzielczość i problemy związane z kolorem – jak dla grafiki rastrowej [4.2], lecz przede wszystkim parametr kluczowy dla filmów – czyli przepływność (framerate, frame frequency), która jest miarą prędkości i płynności wyświetlania ruchomych obrazów wyrażoną w klatkach na sekundę (FPS – frames per second). Wartość przepływności w telewizji i nagraniach filmowych jest stała i zależy od sposobu kodowania. Zbyt niska wartość FPS powoduje u widza wrażenie „skakania" obrazu, z kolei wysoki zwiększa wrażenie płynności wyświetlania. Z badań nad postrzeganiem wynika, że w przypadku ludzkiego oka w zupełności wystarczy wartość 24 FPS (urządzenia HD), 25 FPS (odtwarzacze DVD) lub 30 FPS (urządzenia mobilne). Warto jeszcze doprecyzować rozdzielczość (liczoną w liniach, która w przypadku wideo określają poszczególne standardy: VGA (640x480), PAL (768x576), XGA (1024x768) oraz HDTV w dwu wariantach (HDRedy 1280x720) lub (FullHD 1920x1080). Wynikają stąd znaczne różnice zapotrzebowania na dane, proste obliczenia pozwalają stwierdzić, że najmniejszy monitor VGA wyświetla jednorazowo 0,3 Mpix, zaś największy FullHD – prawie 2 Mpix. Nawet jeśli większość jest nadawana z przeplotem (na zmianę wyświetlane są linie parzyste i nieparzyste), co redukuje ilość danych o połowę – to i tak wideo w porównaniu z innymi mediami generuje gigantyczne zapotrzebowanie na transfer. W praktyce istnieją dwa wyjścia, aby ograniczyć bitrate przy zachowaniu akceptowalnej jakości: zastosować agresywne kompresję stratną (kodeki) oraz wykorzystać formaty wspierające skalowalność [ekranu]. Warto dodać, że formaty wideo, podobnie jak audio składają się zazwyczaj z kontenera multimedialny (istnieje ich ok. 25, np. MPEG-4 Part 14, AVI, ASF, 3GP, QuickTime, VOB, Ogg) [Comparison of container form. – http://en.wikipedia.org/wiki/Comparison_of_container_formats], w którym zastosowany jest odpowiedni kodek, osobny dla obrazu (powstało ich ok. 35, np. H.264, MPEG-2, MPEG-4 Part 2, XVid) [Comparison of video codecs – http://en.wikipedia.org/wiki/Comparison_of_video_codecs] i osobny dla dźwięku (np. MP3, AC3, ACC).

Na obecnym poziomie rozwoju technologii najpopularniejszy jest grupa formatów z rodziny MPEG, na której rozwojem czuwa Moving Picture Experts Group [Moving Picture Experts Group – http://mpeg.chiariglione.org/] – grupa robocza ISO/IEC zajmująca się rozwojem standardów kodowania audio i wideo, a także zatwierdzona przez ISO grupa powszechnie stosowanych formatów zapisu danych zawierających obraz i dźwięk. Aktualnie w praktyce telewizji cyfrowej używa się dwu formatów z tej

grupy: MPEG-2, który opracowano w 1994 (norma ISO/IEC 13818-2:1996) oraz MPEG-4, który rozwija się od 1998 roku (norma ISO/IEC 14496 – 27 arkuszy). Specyfikacja **MPEG-2** obsługuje wszelkie rozdzielczości (352x288 do 1920x1152), pozwala na kodowanie obrazu zarówno z przeplotem, jak i bez niego, obsługuje skalowanie i jest zoptymalizowana dla przepływności 4 Mbps. Jakkolwiek jest on wciąż bardzo popularny z wolna jednak traci użytkowników na rzecz standardu **MPG-4 [AVC]** (dokładnie MPG-4 Part 10 z kodekiem H.264), zwanego też AVC (Advanced Video Coding). Głównymi zaletami AVC jest koncepcja obiektowa, zorientowanie na media strumieniowe, skalowalność (176x144 do 4096x2304), a nade wszystko doskonały współczynnik kompresji uzyskany dzięki użyciu wielu innowacji oraz niezależność od medium i interakcja z użytkownikiem. Aktualnie istnieją dwaj główni konkurenci AVC: format **WMV** opracowany przez Microsoft, przyjęty przez amerykańskie stowarzyszenie SMPTE pod nazwą VC-1, oraz **AVS** opracowany przez konsorcjum stworzone przez rząd chiński. Wśród wielu innych formatów o funkcjonalności zbliżonej do H.264/AVC warto wspomnieć o popularnych kontenerach: **MOV** (znanej jako QuickTime firmy Apple), wspomnianym już wcześniej formacie **Ogg** oraz specyfikacji **FLV** (Flash Video), rozwijanej prze Adobe. Ostatnia specyfikacja stała się szczególnie popularna za sprawą używania jej w globalnych serwisach wideo: YouTube, Reuters.com, Yahoo! Video czy MySpace.

4.6. Inne formaty

Wśród znacznej grupy innych formatów warto osobno podkreślić rozwijany przez Adobe format **SWF** (Shockwave Flash Object) [SWF File format specification (v. 10) – pdf/swf_file_format_spec_v10.pdf], który w sieci jest dominującą formą prezentacji animacji i interaktywnych apletów. Funkcjonalność i rozbudowę animacji fleszowych umożliwia ponadto udostępniony w 2004 r. język programowania **ActionScript** [AAS – http://www.adobe.com/devnet/actionscript/references.html], który pozwala programować zdarzenia. Dzięki właściwościom języków skryptowych można łatwo tworzyć na stronach www rozmaite formy prezentacji dowolnego typu danych, np. grafiki rastrowej, dla której zaprogramowane są mechanizmy nawigacji, powiększania i efekty przejścia [przykład: http://www.paulvanroekel.nl/picasa/flashgrid/demo/index.html]. Niektóre rodzaje danych ze względu na posiadane parametry lub potrzeby użytkowników posługują się dedykowanymi formatami. Przykładem jest serwowanie map – czyli obiektów o znacznych rozmiarach. Popularnym rozwiązaniem w tym zakresie okazało się podzielenie grafiki rastrowej na wiele części (nawet tysięcy) i ich ponowne wirtualne składanie w zależności od wybranej skali powiększenia. Na takim mechanizmie opiera się m.in. serwis GoogleMaps oraz popularny format **Zoomify** [Zoomify – http://www.zoomify.com/]. Innym przykładem są nuty. Dedykowany format **Sibelius** [Silebius – http://www.sibelius.com/], który odtwarza jednocześnie zapis i podkład muzyczny.

Wiele dedykowanych formatów stworzono także z myślą o urządzeniach mobilnych. Jednak jedynie kilka zdobyło ugruntowaną pozycję i jest prezentowanych w sie-

ci dystrybucji e-booków, np. Powel's Books. Stosuje się albo formaty natywne: Microsoft Reader (**LIT**); MobiPocket (**PRC**), iPad (**iTunes**), Adobe Digital Editions (**PDF/DRM**) albo formaty otwarte: **EPUB**, **PDF** lub **Flash**. Uniwersalnym narzędziem umożliwiającym odtwarzanie, konwersję pierwszej grupy formatów jest aplikacja Calibre [http://calibre-ebook.com/], która działa na wielu platformach i ma wbudowane liczna narzędzia do zarządzania kolekcjami. Duży potencjał rozwojowy ma ostatnia grupa (formaty otwarte), gdyż pozwala serwować dowolny typ dokumentu (np. PDF, TXT) oraz precyzyjnie zarządzać DRM. Rozwiązanie takie stosują popularne m.in. portale wymiany plików, np. ISSUU [http://issuu.com/], Sciribd [http://www.scribd.com/] czy DocStock [Docstoc – http://www.docstoc.com/] oraz wybrane księgarnie internetowe.

Bibliografia

1. Bojar, Bożenna red. (2002). *Słownik encyklopedyczny informacji, języków i systemów informacyjno-wyszukiwawczych*. Warszawa: Wydawnictwo SBP, 363 s.
2. Caplan, Priscilla (2009). *Understanding PREMIS*. Washington: Library of Congress, 26 s. [dok. elektr.] http://www.loc.gov/standards/premis/understanding-premis.pdf [odczyt: 20.04.2011].
3. *Data Dictionary for Preservation Metadata. Data Dictionary section from PREMIS*. Version 2.1 (2011) [dok. elekt.] http://www.loc.gov/standards/premis/v2/premis-dd-2-1.pdf [20.04.2011].
4. *Decyzja Nr 25 Naczelnego Dyrektora Archiwów Państwowych z dnia 30.12.2005 r. w sprawie standardów sporządzania cyfrowych kopii materiałów archiwalnych w archiwach państwowych* (2005) [dok. elekt.] http://www.archiwa.gov.pl/images/stories/file/pdf/dec_25_2005.pdf [odczyt: 20.04.2011].
5. *Informacja i dokumentacja – zarządzanie dokumentami*. Cz. 1: *Zasady ogólne PN-ISO 15489-1* (2006). Warszawa: Polski Komitet Normalizacyjny.
6. *Informacja i dokumentacja – zarządzanie dokumentami*. Cz. 2: *Wytyczne Raport Techniczny PKN-ISO/TR 15489-2* (2010). Warszawa: Polski Komitet Normalizacyjny.
7. *Internet Media Type registration, consistency of use TAG Finding 3 June 2002* (Revised 4 September 2002) [dok. elektr.] http://www.w3.org/2001/tag/2002/0129-mime [odczyt: 20.04.2011].
8. Jaworski, Radosław (2009). *Multimedia i grafika komputerowa*. Warszawa: Wydawnictwa Szkolne i Pedagogiczne, 328 s.
9. *JPEG image compression FAQ*. Part 1/2 (1999) [dok. elektr.] http://www.faqs.org/faqs/jpeg-faq/part1/ [2011.04.22].
10. Kamiński, Bogdan (2005). *Przetwarzanie informacji*. Cz. 1: *Przetwarzanie tekstu, przetwarzanie obrazu*. Warszawa: Wydawnictwa Szkolne i Pedagogiczne, 272 s.
11. Matejko, Andrzej; Januszewski, Mirosław (2006). *E-paczka, czyli wysyłam dokumenty do urzędu* [dok. elektr.] http://www.podpiselektroniczny.pl/epaczka.htm [odczyt: 20.04.2011].
12. *MIME Media Types* (2007) [dok. elektr.] http://www.iana.org/assignments/media-types/index.html [odczyt: 20.04.2011].

13. Owczarz-Dadan, Anna (2006). *Tworzenie filmów w Windows XP*. Gliwice: Wydawnictwo Helion, 240 s.
14. Owczarz-Dadan, Anna (2010). *Photoshop CS4 PL*. Gliwice: Wydawnictwo Helion, 112 s.
15. *PostScrip language reference*. Third edition. (1999). Reading: Addison-Wesley Publishing Company, 912 s. [dok. elektr.] http://partners.adobe.com/public/developer/en/ps/PLRM.pdf [odczyt: 20.04.2011].
16. *Rozporządzenie Rady Ministrów z dnia 11 października 2005 r. w sprawie minimalnych wymagań dla systemów teleinformatycznych* (2005). „Dziennik Ustaw" nr 212, poz. 1766 – załącznik 2.
17. Rudny, Tomasz (2010). *Multimedia i grafika komputerowa*. Gliwice: Wydawnictwo Helion, 248 s.
18. Tamborski, Tomasz (2004). *Przetwarzanie informacji*. Cz. 2: *Przetwarzanie dźwięku, przetwarzanie strumieniowe*. Warszawa: Wydawnictwa Szkolne i Pedagogiczne, 128 s.
19. Ulrich, Katherine (2008). *Flash CS3 Professional PL*. Gliwice: Wydawnictwo Helion, 578 s.
20. Zakrzewski, Paweł (2007). *Adobe Premiere Elements*. Gliwice: Wydawnictwo Helion, 432 s.
21. *Zarządzanie dokumentacją. Format pliku dokumentu elektronicznego do długoterminowego zabezpieczenia*. Część 1, *Stosowanie PDF 1.4* (PDF/A-1) PN-ISO 19005-1 (2007). Warszawa : Polski Komitet Normalizacyjny.
22. Zimek, Roland (2005). *ABC grafiki komputerowej*. Gliwice: Wydawnictwo Helion, 240 s.

DjVu

Bartłomiej Siedlarz

W tekście omówiono często stosowany w bibliotekach cyfrowych format hybrydowy DjVu, dedykowany dla publikacji zdigitalizowanych.

Wstęp

Najbardziej istotnym z punktu widzenia użytkownika celem digitalizacji jest udostępnienie dokumentów w sieci globalnej. Konwersja dzieł drukowanych do hipertekstu jest jedną z możliwości reprezentacji treści. Hipertekst ma kilka zalet: jest podstawowym sposobem zapisu treści w sieci, jest w pełni przeszukiwalny, a pliki wynikowe mają niewielki rozmiar. Konwersja do hipertekstu ma jednak dwie bardzo poważne wady: po pierwsze – jest bardzo droga i czasochłonna; po drugie – trudno w ten sposób odzwierciedlić pierwotny układ graficzny i strukturalny publikacji [LeCun; Bottou; Haffner; Howard 1998]. Wspomniane problemy próbowano rozwiązać przez użycie bitmap do zapisu stron dokumentu w celu przeniesienia oryginalnego układu typograficznego, aczkolwiek powszechnie używane formaty zapisu grafiki rastrowej nie umożliwiały osadzania warstwy tekstowej oraz tworzenia użytecznych dokumentów wielostronicowych. Jako rozwiązanie problemu wielostronicowości oraz osadzania tekstu zaproponowano użycie opracowanego przez Adobe formatu PDF jako kontenera dla uzyskanych w procesie skanowania obrazów. Użycie PDF nie rozwiązało jednak wszystkich problemów, gdyż PDF, jako format zaprojektowany głównie do wykorzystania w poligrafii, nie zapewniał możliwości tworzenia dokumentów zoptymalizowanych do dystrybucji w Internecie. Konieczne było opracowanie nowego formatu, który nadawałby się do zapisu dokumentów zdigitalizowanych w sposób umożliwiający odzwierciedlenie struktury pierwotnej publikacji, układu graficznego stron, integracji przeszukiwalnego tekstu, a zarazem oferowałby stopień kompresji akceptowalny w dystrybucji sieciowej.

W 1996 r. w zespół z AT&T Laboratories kierowany przez Yanna LeCuna, Léona Bottou, Patricka Haffnera i Paula G. Howarda opracował format zapisu dokumentów

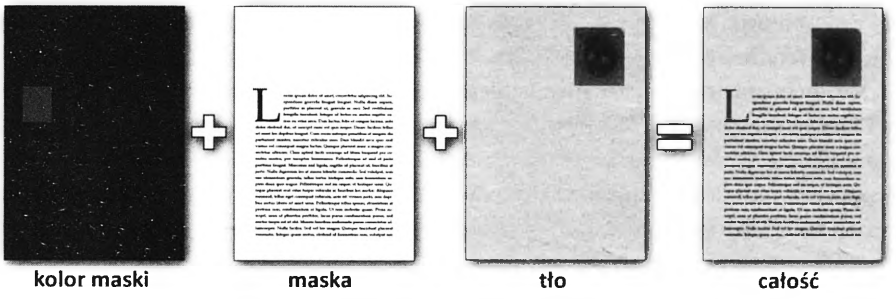
cyfrowych DjVu (nazwa pochodzi od fr. *déjà vu* – już widziane). Produkcją oprogramowania konwertującego, dekodującego oraz narzędziowego, a także komercjalizacją formatu zajęła się firma LizardTech, obecnie tę rolę przejęła Caminova. Pierwotni twórcy stworzyli oprogramowanie o otwartym kodzie źródłowym do tworzenia DjVu – DjVuLibre.

DjVu jest formatem hybrydowym stworzonym na podstawie popularnego Interchange File Format. Jest to format dedykowany dla publikacji, które powstały przez digitalizację dokumentów drukowanych i mają być udaostępniane zdalnie za pośrednictwem sieci globalnej – jest zoptymalizowany do przechowywania grafiki rastrowej z dołączonym tekstem. Główną zaletą formatu DjVu jest układ warstwowy, zgodny z opisanym w zaleceniach ITU-T modelem obrazowania, określonym jako *Mixed Raster Content* [ITU-T Recommendation 2005]. Według opisu formatu w specyfikacji, obraz wejściowy zostaje rozdzielony na warstwę tła, maskę i warstwę pierwszego planu, do których zostaje zastosowana zoptymalizowana dla każdego typu kompresja [Lizardtech DjVu Reference v3 2005].

DjVu jest formatem chętnie wykorzystywanym przez biblioteki cyfrowe. Większość bibliotek Federacji Bibliotek Cyfrowych ma w swoich zasobach takie dokumenty. DjVu jest także jednym z głównych formatów zapisu dokumentów pierwotnie drukowanych w Internet Archive, wykorzystywany jest także w Wikisource oraz przez firmy komercyjne (np. Samsung oraz LG oferują dokumentacje swoich urządzeń także w tym formacie). DjVu nadaje się także do zapisu dokumentów pierwotnie cyfrowych – jednym z pierwszych przykładów zastosowania w tym właśnie celu jest archiwum materiałów konferencji Neural Information Processing Systems (nips.djvuzone.org). Więcej przykładów zastosowań DjVu znaleźć można na wykazach zamieszczonych na stronach: djvu.org oraz djvu.pl.

Dokumenty DjVu

Zgodnie z założeniami modelu *Mixed Raster Content* obrazy rastrowe zostają rozdzielone na warstwy, przetworzone i zakodowane w sposób optymalny dla ich zawartości. MRC został opracowany w celu zwiększenia skuteczności transmisji i archiwizacji obrazów rastrowych [ITU-T Recommendation 2005]. Obraz DjVu zawiera trzy warstwy: pierwszy plan, tło oraz maskę. Podstawową warstwą jest maska, która jest obrazem bitonalnym o pełnej rozdzielczości, jej rolą jest oddzielenie pierwszego planu od tła. Jeżeli potraktujemy maskę jako obraz zawierający czarne i białe piksele, to w przypadku wyświetlania dokumentu DjVu przeglądarka podstawia obraz tła zamiast białych pikseli, a obraz lub kolor pierwszego planu zamiast pikseli czarnych. Bitonalna maska jest kompresowana przy użyciu algorytmu JB2, który umożliwia bardzo wydajną kompresję w przypadku występowania w obrazie licznych podobnych kształtów. Warstwa pierwszego planu oraz tło są zwykle obrazami kolorowymi o zmniejszonej rozdzielczości, zakodowanymi przy użyciu kompresji IW44, która jest wydajna w przypadku obrazów o dużej liczbie kolorów.



Rys. 1. Warstwy obrazu DjVu

Nie zawsze wszystkie warstwy muszą być obecne w dokumencie. Sama warstwa maski może być funkcjonalnym dokumentem DjVu – jeżeli informacje o tle oraz kolorach nie są dostępne, przeglądarka standardowo wyświetla białe tło i czarny tekst (jest to tzw. tryb bitonalny); podobnie warstwa tła może funkcjonować samodzielnie (tzw. tryb *photo*). w normalnym trybie (określanym jako *kolor*) dokument musi mieć wszystkie warstwy.

Dokumenty DjVu mogą być jednostronicowe lub wielostronicowe, przy czym każda strona dokumentu wielostronicowego jest w pełni funkcjonalnym samodzielnym dokumentem [Lizardtech DjVu Reference v3 2005]. Wielostronicowe dokumenty DjVu zapisać można w postaci scalonej lub rozdzielonej. Dokument w postaci scalonej jest jednym plikiem, który zawiera obrazy wszystkich stron, indeks, miniatury oraz inne dane pomocnicze. Postać rozdzielona oznacza z reguły folder, w którym wyżej wymienione części składowe są odrębnymi plikami. Możliwa jest łatwa konwersja pomiędzy wariantami zapisu nawet bez użycia zaawansowanych narzędzi – niektóre przeglądarki umożliwiają zapis pobranych dokumentów w obu postaciach. Rozdzielona postać DjVu ułatwia korzystanie z dokumentów udostępnionych w Internecie, gdyż do przejrzenia fragmentów dokumentu nie jest konieczne pobranie całości. Postać scalona ułatwia zarządzanie dokumentami tam, gdzie nie występuje znaczne ograniczenie prędkości odczytu.

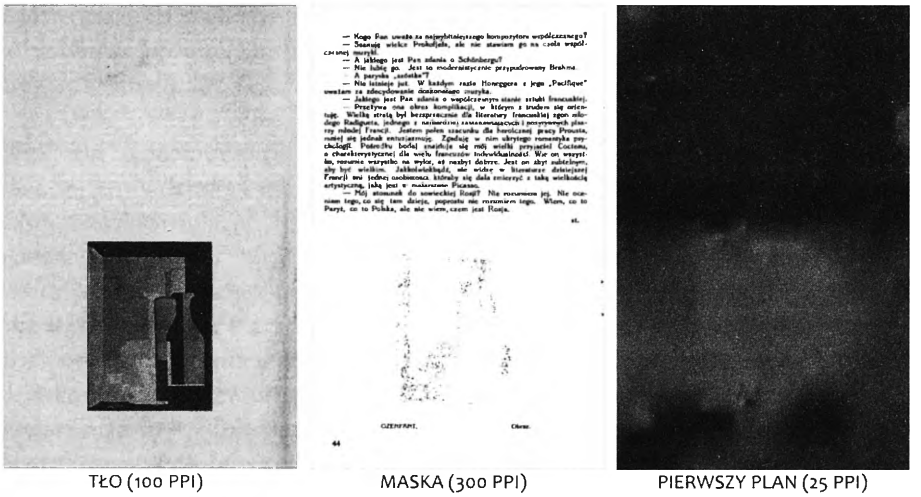
DjVu jest formatem progresywnym: zarys strony wyświetlany jest bardzo szybko, a jakość obrazu stopniowo się zwiększa wraz z wczytywaniem kolejnych bitów [What is DjVu 2011]. Progresywność w połączeniu z rozdzieloną postacią zapisu umożliwia sprawne korzystanie z dokumentu w środowisku sieciowym, gdyż pobranie dokumentu w całości nie jest konieczne do przejrzenia jego fragmentów.

DjVu używa obrazu bitonalnego w celu odzwierciedlenia treści dokumentu w odpowiedniej rozdzielczości, a obrazów kolorowych w celu zachowania cech drugorzędnych obiektu źródłowego oraz grafiki, zawierającej przejścia tonalne. Redukcja rozdzielczości warstw kolorowych nie wpływa znacznie na obniżenie czytelności dokumentu, pozwala natomiast radykalnie zmniejszyć rozmiar wynikowego pliku. Efektem zarówno redukcji rozdzielczości jak i kompresji stratnej jest odrzucenie zbęd-

nych danych i zachowanie tylko tych, które wystarczają do przeniesienia pożądanych cech dokumentu. w przypadku tekstu można odrzucić nawet 90% danych zawartych w obrazie źródłowym.

Obrazy bitonalne są generalnie mniejsze niż obrazy kolorowe, jednak nie pozwalają na szczegółowe odzwierciedlenie wszystkich cech obiektu źródłowego. Obraz strony publikacji drukowanej zawierający tekst można skutecznie skompresować stratnie poprzez redukcję podobnych kształtów do wspólnego prototypu, na podstawie którego będą one odtwarzane przy dekodowaniu. Zakodowanie jednego znaku oraz drobnych zmian w poszczególnych jego wariantach zajmuje mniej miejsca niż zakodowanie każdego znaku indywidualnie.

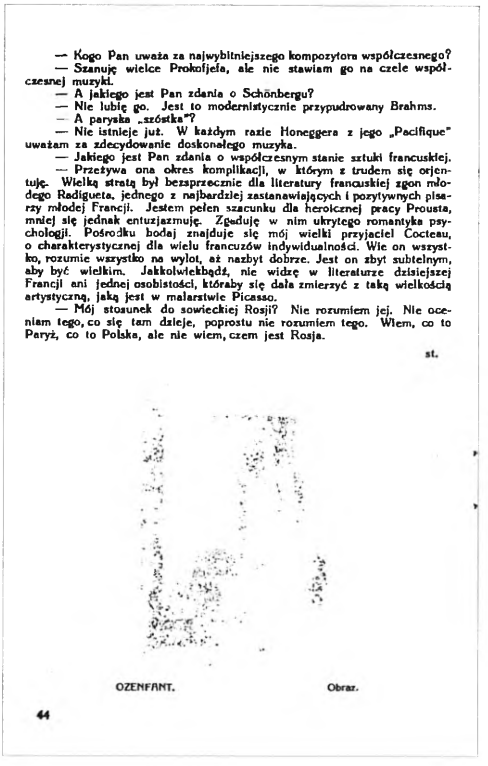
Obrazy warstwy tła zwykle redukuje się tak, aby rozdzielczość była zgodna z rozdzielczością przeciętnego ekranu, czyli 100 PPI. w szczególnych przypadkach rozdzielczość warstwy tła może być większa jeżeli strona zawiera ilustracje, lub mniejsza, jeżeli pragnie się jeszcze bardziej zoptymalizować rozmiar pliku.



TŁO (100 PPI) MASKA (300 PPI) PIERWSZY PLAN (25 PPI)

Rys. 2. Rzeczywisty wygląd warstw obrazu DjVu
(obrazy powiększone do równych rozmiarów)

Rysunek 3 ukazuje proporcjonalną wielkość obrazów poszczególnych warstw. Znaczny zysk na rozmiarze pliku wynika z redukcji rozdzielczości warstwy tła oraz warstwy pierwszego planu w stosunku do warstwy maski. Wykorzystanie segmentacji obrazu pozwala na zachowanie szczegółów (głównie tekstu) w wysokiej rozdzielczości, podczas gdy elementy mniej istotne (kolor tekstu) lub niewymagające dużej dokładności (grafika) mogą zostać zredukowane. Obraz wynikowy po złożeniu zapewnia wystarczającą czytelność.



Rys. 3. Rzeczywisty wygląd warstw obrazu DjVu
(obrazy przedstawione proporcjonalnie)

Struktura plików DjVu

Pliki DjVu są zgodne z opracowanym w 1985 r. przez Electronic Arts standardem Interchange File Format [EA IFF 85 1985]. IFF85 opisuje strukturę kontenerów, które zawierać mogą części składowe dokumentu DjVu. Kontenery te zwane są *kawałkami*, a ich rola w dokumencie oznaczana jest przez czteroznakowy identyfikator.

Przykładowa struktura jednostronicowego obrazu DjVu wygląda następująco: kawałek `INFO` zawiera informacje o wielkości obrazu, rozdzielczości, wartości współczynnika gamma oraz wersji; *`Sjbz`* zawiera bitonalną maskę; *`FG44`* to obraz pierwszego planu; *`BG44`* to obraz tła; *`TXTz`* zawiera wynik rozpoznania tekstu, czyli znaki lub wyrazy oraz ich koordynaty. Wszystkie powyższe kawałki zawarte są w zbiorczym kawałku, oznaczonym jako *`FORM:DJVU`*. Stworzenie wielostronicowego dokumentu polega na upakowaniu wszystkich składowych *`FORM:DJVU`* wraz z ich wykazem, oznaczonym jako *DIRM* do zbiorczego kontenera *`FORM:DJVM`*. Na rys. 4 przedstawiono strukturę prostego dokumentu, który nie zawiera danych współdzielonych:

```
FORM:DJVM [669935]
  DIRM [95]         Document directory (bundled, 4 files 4 pages)
  FORM:DJVU [207485] {DIGCZAS001283_R005_1858_005_0001.djvu} [P1]
    INFO [10]         DjVu 2360x3317, v25, 300 dpi, gamma=2.2
    Sjbz [67729]      JB2 bilevel data
    FG44 [39980]      IW4 data #1, 100 slices, v1.2 (color), 295x415
    BG44 [1937]       IW4 data #1, 70 slices, v1.2 (color), 1180x1659
    BG44 [12042]      IW4 data #2, 13 slices
    BG44 [7577]       IW4 data #3, 5 slices
    BG44 [67099]      IW4 data #4, 10 slices
    TXTz [11039]      Hidden text (text, etc.)
  FORM:DJVU [218082] {DIGCZAS001283_R005_1858_005_0002.djvu} [P2]
    INFO [10]         DjVu 2360x3317, v25, 300 dpi, gamma=2.2
    Sjbz [76534]      JB2 bilevel data
    FG44 [43805]      IW4 data #1, 100 slices, v1.2 (color), 295x415
    BG44 [2016]       IW4 data #1, 70 slices, v1.2 (color), 1180x1659
    BG44 [12930]      IW4 data #2, 13 slices
    BG44 [8702]       IW4 data #3, 5 slices
    BG44 [63473]      IW4 data #4, 10 slices
    TXTz [10542]      Hidden text (text, etc.)
  FORM:DJVU [123904] {DIGCZAS001283_R005_1858_005_0003.djvu} [P3]
    INFO [10]         DjVu 2360x3317, v25, 300 dpi, gamma=2.2
    Sjbz [53596]      JB2 bilevel data
    FG44 [11894]      IW4 data #1, 100 slices, v1.2 (color), 197x277
    BG44 [1978]       IW4 data #1, 74 slices, v1.2 (color), 787x1106
    BG44 [5709]       IW4 data #2, 13 slices
    BG44 [6668]       IW4 data #3, 5 slices
    BG44 [33984]      IW4 data #4, 10 slices
    TXTz [9996]       Hidden text (text, etc.)
  FORM:DJVU [120323] {DIGCZAS001283_R005_1858_005_0004.djvu} [P4]
    INFO [10]         DjVu 2360x3317, v25, 300 dpi, gamma=2.2
    Sjbz [54275]      JB2 bilevel data
    FG44 [10775]      IW4 data #1, 100 slices, v1.2 (color), 197x277
    BG44 [1610]       IW4 data #1, 74 slices, v1.2 (color), 787x1106
    BG44 [4227]       IW4 data #2, 13 slices
    BG44 [5738]       IW4 data #3, 5 slices
    BG44 [33994]      IW4 data #4, 10 slices
    TXTz [9623]       Hidden text (text, etc.)
```

Rys. 4. Struktura czterostronicowego dokumentu DjVu.

DjVu umożliwia także współdzielenie plików składowych pomiędzy stronami, głównie są to słowniki kształtów oraz adnotacje wspólne dla całego dokumentu. Współdzielenie słownika kształtów pozwala na zmniejszenie rozmiaru pliku, szczególnie w przypadku konwersji dokumentów pierwotnie cyfrowych. Wszystkie możliwe typy kawałków zamieszczone zostały na poniższym wykazie [Lizardtech DjVu Reference v3 2005]:

FORM	Kawałek zbiorczy, umożliwia osadzanie innych kawałków.
FORM:DJVM	Dokument wielostronicowy.
FORM:DJVU	Dokument jednostronicowy.
FORM:DJVI	Plik współdzielony.
FORM:THUM	Kontener dla osadzonych miniatur.
DIRM	Lista stron dokumentu wielostronicowego.
NAVM	Informacje o zakładkach.
ANTa, ANTz	Adnotacje, hyperlinki, ustawienia indywidualne dokumentu.
TXTa, TXTz	Tekst oraz koordynaty.
Djbz	Współdzielony słownik kształtów.
Sjbz	Bitonalny obraz maski JB2.
FG44	Obraz warstwy pierwszego planu IW44.
BG44	Obraz warstwy tła IW44.
TH44	Miniatura strony.
WMRM	Dane potrzebne do zdjęcia znaku wodnego.
FGBz	Opis koloru poszczególnych kształtów.
INFO	Informacje o stronie.
INCL	Odnośnik do pliku współdzielonego.
BGjp	Obraz warstwy pierwszego planu JPEG.
FGjp	Obraz warstwy tła JPEG.
Smmr	Bitonalny obraz maski G4.

DjVuLibre zawiera narzędzia, dzięki którym możliwe jest rozbijanie pojedynczych obrazów DjVu na poszczególne kawałki oraz ich ponowne składanie. DjVu może zawierać obrazy zakodowane przy użyciu kompresji JB2, IW44, G4, JPEG, JPEG2000 oraz dane tekstowe i adnotacje, skompresowane przy użyciu BZZ [Lizardtech DjVu Reference v3 2005]. Każdy z wymienionych powyżej typów obrazów może funkcjonować jako samodzielny jednostronicowy dokument DjVu jeżeli zawiera odpowiedni nagłówek. Istnieje możliwość dodania do kontenera DjVu także innych danych, aczkolwiek nie są one odczytywalne przez przeglądarki.

Przeglądarki i wsparcie

Oficjalnym typem MIME dla DjVu jest od 2002 r. image/vnd.djvu [MIME Media Type 1999], poprzednio stosowano image/x-djvu oraz image/x.djvu. Dokument DjVu w postaci rozdzielonej jest strukturalnie podobny do plików HTML, dlatego może być serwowany w dokładnie taki sam sposób, jednakże żadna z obecnych przeglądarek internetowych nie obsługuje DjVu bezpośrednio. Dokumenty odczytywane

są przy użyciu dedykowanych wtyczek, dzięki którym jest możliwe jest także wyświetlenie DjVu osadzonego w HTML (DjVuViewer, DjView). Istnieją także przeglądarki DjVu, które umożliwiają jedynie odtwarzanie zasobów dostępnych lokalnie (np. WinDjView), zawierają natomiast inne, bardziej zaawansowane funkcje, których nie mają wtyczki (np. tworzenie przez użytkownika własnych zakładek i adnotacji).

Najpopularniejsze przeglądarki DjVu:



DjVuViewer

Obecnie rozwijana przez firmę Caminova. Dostępna w wersjach dla systemów Windows, MacOS oraz BeOS. Może działać niezależnie lub jako wtyczka do przeglądarek internetowych. Umożliwia konwersję dokumentów DjVu z postaci scalonej do rozdzielonej i odwrotnie.
http://www.caminova.jp



DjView

Jest częścią pakietu DjVuLibre. Dostępna w wersjach dla systemów Linux, Windows, MacOS oraz BeOS. Jest oprogramowaniem OpenSource. Może działać niezależnie lub jako wtyczka do przeglądarek internetowych. Umożliwia konwersję dokumentów DjVu z postaci scalonej do rozdzielonej i odwrotnie.
http://djvu.sourceforge.net/djview4.html

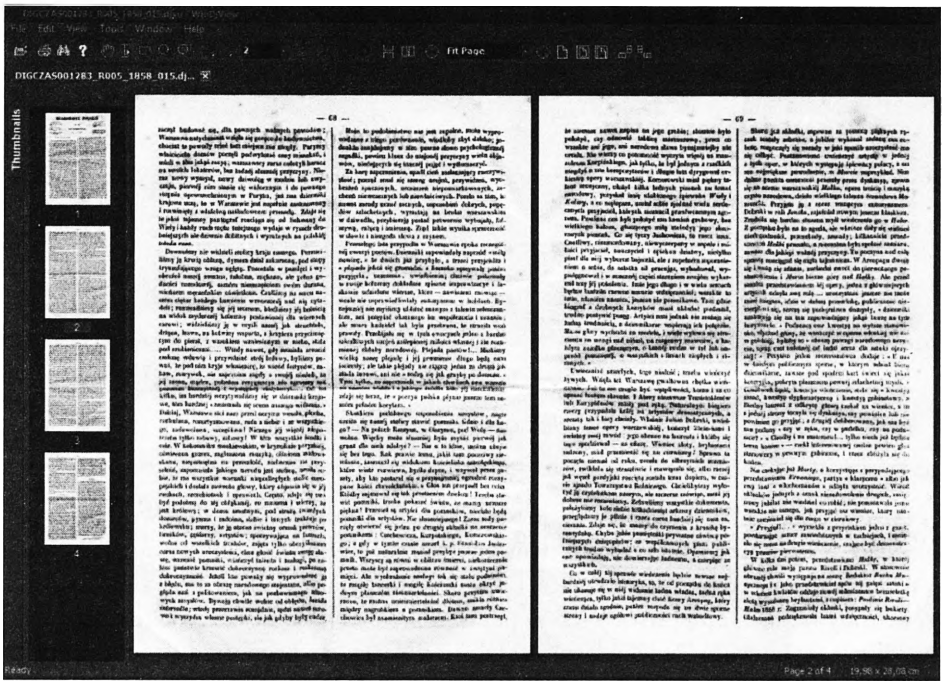


WinDjView
MacDjView

Dostępna dla Windows oraz MacOS (jako MacDjView). Zoptymalizowana do szybkiego wyświetlania dokumentów. Nie może działać jako wtyczka do przeglądarek internetowych ani konwertować dokumentów, umożliwia jednak tworzenie zakładek oraz adnotacji przez użytkownika. Istnieje także wersja z renderowaniem podpikselowym. Jest oprogramowaniem OpenSource.
http://windjview.sourceforge.net

Przeglądarki dla platform mobilnych: SdjVu (Symbian), XDjvu (iPhone), VuDroid (Android), PocketDjVu (PocketPC), GrxView Pro (Palm OS), SmartDjVu (Pocket PC/Windows CE), Handy DjVu (Symbian). Format DjVu jest także wspierany przez czytniki eboków: Sony Librie, eGriver, BeBook, Hanlin, Kogan Ebook Reader, Boox, PocketBook, OnyxBook, Pixelar. Przeglądarki graficzne IrfanView oraz ACDSee oraz multiprzeglądarki Okular, Evince oraz STDU Viewer także odczytują DjVu. Istnieje także przeglądarka dla środowiska Java – Java DjVu.

Bardzo istotną cechą przeglądarek DjVu jest umożliwienie odzwierciedlenia układu publikacji źródłowej, głównie przez zastosowanie widoku dwóch stron z wyodrębnieniem strony tytułowej. Taki sposób wyświetlania zapewnia odbiór bardzo zbliżony do odbioru tradycyjnego kodeksu, niestety w żadnej z obecnie dostępnych przeglądarek nie jest trybem domyślnym.



Rys. 5. Układ dwustronicowy w WinDjview.

Przeglądarka umożliwia także wyświetlanie poszczególnych warstw dokumentu DjVu. Jeżeli maska dokumentu kolorowego jest w pełni funkcjonalnym obrazem bitonalnym, który zawiera wszystkie istotne elementy graficzne strony, można zwiększyć czytelność dokumentu przez wyłączenie tła i kolorowania maski. z dobrze przygotowanej maski dokumentu DjVu można uzyskać wydruk monochromatyczny o jakości kserokopii. Jeżeli dokument ma zintegrowaną warstwę tekstową, możliwe jest wyszukiwanie fragmentów tekstu, a także ich kopiowanie.

Tworzenie obrazów DjVu

W zależności od rodzaju materiału oraz przeznaczenia docelowych plików można zastosować kilka metod rozdziału elementów składowych strony między warstwy. Teoretycznie obiekt drukowany składa się z papierowego tła oraz z obszarów wypełnionych farbą drukarską. Bazując na tym założeniu można zastosować restrykcyjną segmentację i wszystkie elementy graficzne strony przydzielić do maski.



Rys. 6. Przykład restrykcyjnej segmentacji.

Metodą taką można konwertować np. schematy, proste grafiki liniowe, strony zawierające wyłącznie tekst, a także rastrowane grafiki, aczkolwiek obrazy muszą być pozyskane w rozdzielczości pozwalającej odzwierciedlić ziarno rastra. Przy zastosowaniu tej metody maska każdej strony jest w pełni funkcjonalnym obrazem bitonalnym, co umożliwia wykorzystanie jej właśnie w takiej roli (np. do bezpośredniego przeglądania lub wydruku czarno-białego) – jest to niewątpliwie zaletą. Wadą natomiast jest dość duży rozmiar pliku wynikowego jeżeli dokument zawiera ilustracje rastrowane. Nie jest możliwe efektywne zastosowanie tej metody w przypadku obrazów o gładkich przejściach tonalnych i dużej liczbie kolorów.

Drugim sposobem jest przydzielanie ilustracji do warstwy tła. Metoda ta sprawdza się dobrze w przypadku, gdy dokument zawiera ilustracje o gładkich przejściach tonalnych, a także gdy ilustracje rastrowane zostały zeskanowane z rozdzielczością niewystarczającą do odzwierciedlenia ziarna rastra (por. rysunek 2). Zaletą tego sposobu tworzenia dokumentów jest zysk na rozmiarze pliku wynikowego, natomiast maska nie zawiera żadnych danych o ilustracjach, lub zawiera szczątkowe ich fragmenty, nie może więc samoistnie być funkcjonalnym dokumentem.

Trzecią metodą jest stworzenie obrazu wynikowego, zawierającego wyłącznie jedną warstwę. Może to być zarówno obraz bitonalny jak i kolorowy – w takich przypadkach nie wykonuje się segmentacji obrazu.

Możliwe jest także złożenie obrazu strony z wcześniej przygotowanych obrazów poszczególnych warstw.

Oprogramowanie konwertujące

DocumentExpress http://www.caminova.jp

Opracowany przez Lizardtech, a obecnie rozwijany przez firmę Caminova pakiet do tworzenia dokumentów DjVu. Występuje w dwóch wersjach:

Professional

- **Editor:** umożliwia bezpośrednie skanowanie oraz konwersję obrazów do DjVu, a także wykonywanie OCR oraz dodawanie adnotacji.
- **VirtualPrinter:** pozwala eksportować do DjVu dokumenty utworzone przy pomocy dowolnej aplikacji.
- **Configuration Manager**: edytor profilów konwersji.

Enterprise

- **Workflow Manager:** program usprawniający wsadowe przetwarzanie dużej ilości dokumentów.
- **Configuration Manager:** edytor profilów konwersji.
- **Programy wywoływane z linii komend**:

documenttodjvu	Konwertuje bitmapy oraz pliki PDF lub PS o DjVu z wykonaniem segmentacji obrazu.
phototodjvu	Konwertuje bitmapy w trybie photo bez wykonywania segmentacji.
djvudecode	Konwertuje dokumentu DjVu do innych formatów.
djvujoin	Tworzy rozdzielone wielostronicowe dokumenty DjVu. Wykonuje rozpoznanie tekstu.
djvubundle	Tworzy scalone wielostronicowe dokumenty DjVu. Wykonuje rozpoznanie tekstu.
djvuparsexml	Dodaje adnotacje, hyperlinki, metadane oraz tekst.
djvutoxml	Eksportuje adnotacje, hyperlinki, metadane oraz tekst.
djvutotext	Eksportuje wyłącznie tekst.
watermarkdjvu	Dodaje lub usuwa znak wodny.

DjVu Editor sprawdza się jedynie w przypadku konwersji niewielkiej ilości krótkich dokumentów. Masową konwersję umożliwia wersja Enterprise. Programy wywoływane z linii komend można łatwo wykorzystać jako część większego zindywidualizowanego systemu w zależności od potrzeb.

DjVuLibre http://djvu.sourceforge.net

Jest to pakiet OpenSource zapoczątkowany przez pierwotnych twórców formatu. Zawiera podstawowe oprogramowanie dekodujące oraz kodujące, nie posiada jednak segmentera ani zintegrowanego silnika OCR. Specjalne narzędzia DjVuLibre umożliwiają manipulację dokumentami DjVu na poziomie elementów składowych poszczególnych stron, co nie jest możliwe w DocumentExpress.

Zawartość pakietu DjVuLibre:

bzz Kompresor/dekompresor ogólnego przeznaczenia. Pozwala na kompresję/dekompresję adnotacji, tekstu oraz innych danych.

c44 Konwertuje bitmapy w trybie photo bez wykonywania segmentacji.

cjb2 Konwertuje obrazy bitonalne.

cpaldjvu Konwertuje obrazy o niewielkiej liczbie kolorów.

csepdjvu Umożliwia tworzenie DjVu z przygotowanych wcześniej warstw.

ddjvu Dekoduje obrazy DjVu.

djvm Tworzy scalone wielostronicowe dokumenty DjVu.

djvmcvt Przekształca dokumenty DjVu z postaci scalonej do rozdzielonej i odwrotnie.

djvudump Wyświetla strukturę dokumentu.

djvuextract Eksportuje kawałki z obrazów DjVu.

djvumake Tworzy obrazu DjVu z przygotowanych wcześniej kawałków.

djvups Konwertuje dokumenty DjVu do formatu PostScript.

djvused Pozwala na import oraz eksport adnotacji, warstwy tekstowej oraz generuje miniatury.

djvuserve Program CGI, konwertuje scalone dokumenty do postaci rozdzielonej w locie. Przydatny przy dystrybucji w sieci.

djvutxt Eksportuje warstwę tekstową.

djvutoxml Narzędzia XML dla DjVuLibre. Umożliwiają edycję metadanych, odnośników oraz warstwy tekstowej.

djvuxmlparser

miniDjVu http://minidjvu.sourceforge.net

Odmiana enkodera cjb2 z DjVuLibre zoptymalizowana do stratnej kompresji obrazów bitonalnych, pozwala na uzyskanie mniejszych plików przy jednoczesnej minimalizacji przekłamań wynikających ze stratnej kompresji.

ABBYY FineReader http://finereader.abbyy.com

Od wersji 11. FineReader umożliwia eksport do DjVu, jednak w stosunku do DocumentExpress możliwości konfiguracji są znacznie ograniczone.

Podsumowanie

Pomimo znacznego wykorzystania w przedsięwzięciach digitalizacyjnych DjVu nadal jest formatem mało popularnym w skali globalnej. Zawartość dokumentów DjVu wciąż nie jest indeksowana przez główne wyszukiwarki, co powoduje spore ograniczenie dostępności takich dokumentów. W zakresie tworzenia dokumentów przeznaczenie DjVu nie do końca pokrywa się z potrzebami przeciętnego użytkownika, a brak łatwych w obsłudze

i niedrogich narzędzi oraz stopień trudności procesu przygotowania dokumentu dobrego jakościowo, mogą odstraszyć nawet chętnych amatorów. Rozwiązaniem powyższych problemów może być zwiększenie zakresu zastosowania DjVu, promocja formatu oraz pojawienie się większej ilości nowych narzędzi do wygodnego tworzenia dokumentów.

Zalety

- Bardzo wydajna kompresja pozwala na zachowanie dobrej jakości przy niewielkim rozmiarze pliku.
- Pozwala na odzwierciedlenie wyglądu dokumentu oryginalnego.
- Możliwość tworzenia dokumentów wielostronicowych.
- Może zawierać przeszukiwalną warstwę tekstową.
- Dzięki zastosowaniu rozdzielonej postaci zapisu DjVu świetnie nadaje się do prezentacji dokumentów w Internecie.
- Może zawierać metadane XMP.
- Układ warstwowy pozwala na zawarcie funkcjonalnego dokumentu bitonalnego w dokumencie kolorowym.
- Dekodowanie nie wymaga dużej mocy obliczeniowej oraz znacznych ilości pamięci, co umożliwia płynne powiększanie i przewijanie.
- Może zawierać adnotacje oraz hyperlinki.
- Przeglądarki umożliwiają wyświetlanie dokumentu w układzie dwustronicowym z wyodrębnieniem strony tytułowej, co pozwala na odzwierciedlenie układu książki.
- Obrazy DjVu mogą być osadzane w HTML.
- Można linkować do dowolnej strony dokumentu wielostronicowego, a także określać sposób wyświetlania bezpośrednio z odsyłacza.
- Dokument jest wyświetlany tak samo niezależnie od platformy sprzętowej oraz użytej przeglądarki.

Wady

- Złożony proces tworzenia dokumentów wysokiej jakości.
- Oprogramowanie wyposażone w graficzny interfejs użytkownika oraz zawierające segmenter jest dosyć drogie.
- Automatyczna segmentacja obrazu nie jest idealna.
- Brak narzędzi umożliwiających ręczne wsparcie segmentacji.
- Brak narzędzi umożliwiających wygodną korektę automatycznie rozpoznanego tekstu.
- Wciąż mała popularność.
- Brak zintegrowanych narzędzi eksportu w większości edytorów tekstu i programów DTP.
- Brak bezpośredniego wsparcia przez przeglądarki internetowe (wymaga wtyczki).
- Niewielka ilość narzędzi do tworzenia dokumentów.
- Niedostateczne wsparcie w popularnych programach do rozpoznawania tekstu.
- Nie jest indeksowany przez wyszukiwarki.

Przydatne linki

Oficjalna strona DjVu
http://djvu.org

Strona projektu DjVuLibre
http://djvu.sourceforge.net

Strona formy Caminova, producenta DocumentExpress
http://www.caminova.jp

Strona firmy GB Soft, dystrybutora DocumentExpress w Polsce
http://djvu.pl

Kopia nieistniejącej już strony PlanetDjVu
http://www.djvu-soft.narod.ru/planetdjvu

Strony domowe twórców formatu
http://leon.bottou.org
http://yann.lecun.com
http://www2.research.att.com/~haffner

Usługi konwersji online
http://any2djvu.djvuzone.org
http://www.docspal.com
http://www.djvu-pdf.com

Bibliografia

1. *„EA IFF 85" Standard for Interchange Format Files* (1985). [dok. elektr.] http://www.martinreddy.net/gfx/2d/IFF.txt [odczyt: 12.12.2011].
2. *ITU-T Recommendation T.44: Mixed Raster Content* (2005). [dok. elektr.] http://www.itu.int/rec/T-REC-T.44-200501-I/en [odczyt: 12.12.2011].
3. LeCun, Yann; Bottou, Léon; Haffner, Patrick; Howard, Paul G. (1998). *DjVu: a compression method for distributing scanned documents in color over the Internet, Color 6, IST, 1998.* [dok. elektr.] http://yann.lecun.com/exdb/publis/djvu/lecun-98c.djvu [odczyt: 12.12.2011].
4. *Lizardtech DjVu Reference v3* (2005). [dok. elektr.] http://djvu.org/docs/DjVu3Spec.djvu [odczyt: 12.12.2011].
5. *MIME Media Type: image/vnd.djvu* (1999). http://www.iana.org/assignments/media-types/image/vnd-djvu [odczyt: 12.12.2011].
6. *What is DjVu, and what are the secrets behind DjVu's superior performance?* (2011) [dok. elektr.] http://djvu.org/resources/whatisdjvu.php [odczyt: 12.12.2011].

Organizacja i logistyka digitalizacji

Tomasz Kalota, Marcin Szala

W tekście zaprezentowano procesy digitalizacji, ich organizację i logistykę. Terminem 'digitalizacja' określany jest proces ucyfrowienia obrazu fizycznego obiektu na potrzeby przetwarzania, udostępniania i archiwizacji za pomocą urządzeń teleinformatycznych. Jednak powszechnie w środowisku bibliotekarskim poprzez „digitalizację" rozumie się zbiór wszystkich procesów, które towarzyszą ucyfrowianiu dokumentów (szeroko rozumianych). Oprócz samego ucyfrowienia są to czynności towarzyszące, które przyczyniają się do stworzenia dokumentacji cyfrowej obiektu począwszy od typowania, prac konserwatorskich, po opracowanie metadanych, obróbkę cyfrową, publikację w sieci oraz archiwizację przygotowanych danych cyfrowych.

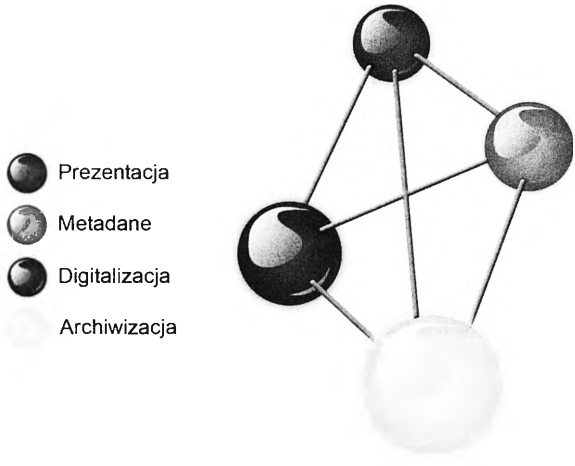
Organizując procesy digitalizacji ważne jest umiejscowienie ich jako elementów większego przedsięwzięcia, jakim jest polski system rozproszonych bibliotek cyfrowych. Funkcjonowanie tego systemu obrazują modele, przedstawione w dalszym toku rozważań.

Spróbujmy sobie wyobrazić model podstawowej komórki organizacyjnej odpowiedzialnej za stworzenie, opracowanie, udostępnianie i archiwizowanie zasobu cyfrowego. Taka podstawowa komórka składa się z czterech elementów, które na rysunku 1 reprezentowane są w ich wzajemnych relacjach.

Elementami tymi są:

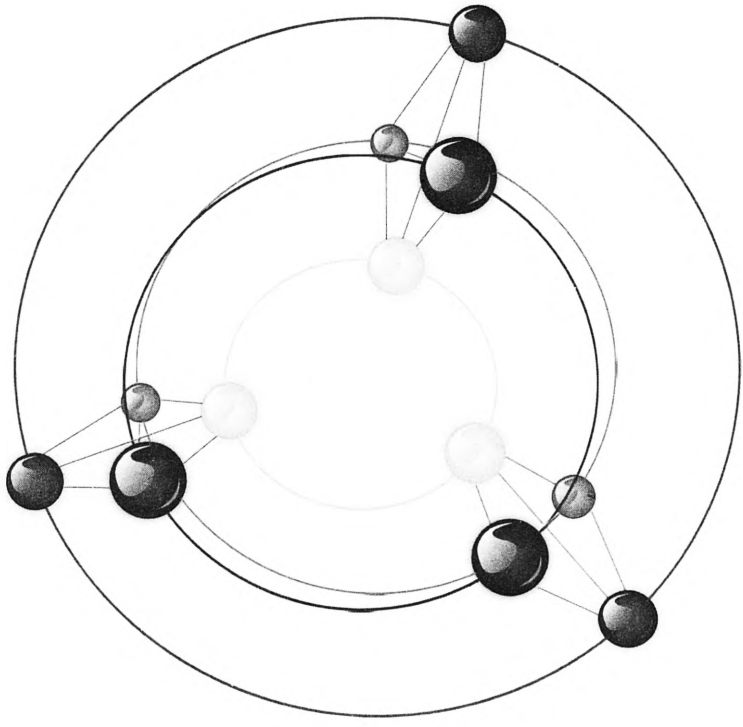
1. Prezentacja – ogół metod i praktyk służących udostępnianiu i prezentowaniu zasobów cyfrowych użytkownikom. Na prezentację składają się: infrastruktura techniczna, oprogramowanie, sprzęt, interfejsy użytkowników, mechanizmy wymiany danych, obsługa techniczna i merytoryczna.

2. Archiwizacja – ogół metod i praktyk służących zabezpieczaniu i długoterminowemu przechowywaniu wytwarzanych zasobów cyfrowych. Na archiwizację składają się: infrastruktura techniczna, oprogramowanie, sprzęt, interfejsy użytkowników, mechanizmy wymiany danych, mechanizmy zabezpieczania i migracji danych, obsługa techniczna.
3. Metadane (opracowanie) – ogół metod i praktyk służących profesjonalnemu opracowaniu zasobów cyfrowych, tak aby użytkownik mógł z nich bez problemu korzystać przy pomocy mechanizmów prezentacyjnych. Na opracowanie składają się: merytoryczne przygotowanie metadanych zgodnie z obowiązującymi standardami, szkolenia w zakresie opracowania zasobów cyfrowych, integracja ze źródłami i systemami wspomagającymi prezentowanie i przeszukiwanie zasobów cyfrowych, analiza potrzeb użytkowników oraz wykorzystania zasobów cyfrowych.
4. Digitalizacja – ogół metod i praktyk służących wytwarzaniu zasobów cyfrowych. Na digitalizację składają się: infrastruktura techniczna, oprogramowanie, sprzęt, stosowanie standardów technicznych i jakościowych, monitorowanie rynku i wdrażanie nowych rozwiązań technologicznych.



Rys. 1. Elementy procesu digitalizacji.

Elementy te tworzą podstawowe komórki organizacyjne wchodzące w relacje, które zachodzą pomiędzy poszczególnymi składowymi elementami pojedynczych komórek. Tworząc dla nich wizualizacje można zaprezentować orbity, przedstawiające jednocześnie płaszczyzny współpracy określone zakresami działań specyficznych dla poszczególnych elementów. Jest to oczywiście model abstrakcyjny, ale może być pomocny w określeniu oraz umiejscowieniu obszarów działań związanych z organizacją i logistyką digitalizacji.



Rys. 2. Płaszczyzny współpracy.

Planując przedsięwzięcia digitalizacyjne należy zwrócić uwagę na istotny problem organizacyjny. Z punktu widzenia administrowania projektem digitalizacji, najkorzystniejszym rozwiązaniem byłoby zdefiniowanie w sposób bardzo konkretny zakresu wykonywanych prac. Upraszczałoby to planowanie zadań i tworzenie kosztorysów. Digitalizacja jest jednak procesem bardziej złożonym i nie zawsze da się ją uprościć do kilku operacji powtarzalnych w każdym przypadku. Na koszty oraz czas realizacji projektów digitalizacji mają wpływ różne czynniki, między innymi: rodzaj, jakość i stan zachowania materiałów przeznaczonych do skanowania, rodzaj posiadanego sprzętu, możliwości kadrowe, różne oczekiwania dotyczące prezentacji zasobów cyfrowych, czy też metody obróbki i przetwarzania danych. To wszystko wpływa na złożoność procesów digitalizacji i konieczność projektowania dedykowanych linii technologicznych dla różnych rodzajów zbiorów.

Przed rozpoczęciem planowania digitalizacji trzeba sobie odpowiedzieć na kilka pytań, które mają wpływ na decyzje inwestycyjne i organizacyjne. Po pierwsze: czy tworzyć własną pracownię digitalizacji, czy skorzystać z outsourcingu i zlecić część lub wszystkie zadania instytucjom, bądź firmom zewnętrznym, które zajmują się świadczeniem tego typu usług. Przed podjęciem prac związanych z tworzeniem zasobu cyfrowego wskazane jest więc przeanalizowanie różnych wariantów oraz

modeli współpracy wewnątrz własnej instytucji oraz z instytucjami zewnętrznymi. Przy podejmowaniu tej decyzji pomocna może być rada znanego amerykańskiego przemysłowca Henry'ego Forda: *Jeśli jest coś, czego nie potrafimy zrobić wydajniej, taniej i lepiej niż konkurenci, nie ma sensu, żebyśmy to robili i powinniśmy zatrudnić do wykonania tej pracy kogoś, kto zrobi to lepiej niż my* [http://pl.wikiquote.org/wiki/Henry_Ford].

Innym rozwiązaniem jest Społeczna Pracownia Digitalizacji (dalej: SPD) – model opracowany i wdrożony przez twórców Śląskiej Biblioteki Cyfrowej [http://www.sbc.org.pl/]. Wiele ciekawych informacji na temat SPD można znaleźć na blogu Biblioteka 2.0 w artykule *Społeczna Pracownia Digitalizacji ŚBC po roku działalności* [Śpiechowicz; Strokosz 2008] oraz w prezentacji Agnieszki Koszowskiej i Remigiusza Lisa *Biblioteka cyfrowa w wersji 2.0?: wybór możliwości* [2008], a także w rozdziale tej książki autorstwa Remigiusza Lisa.

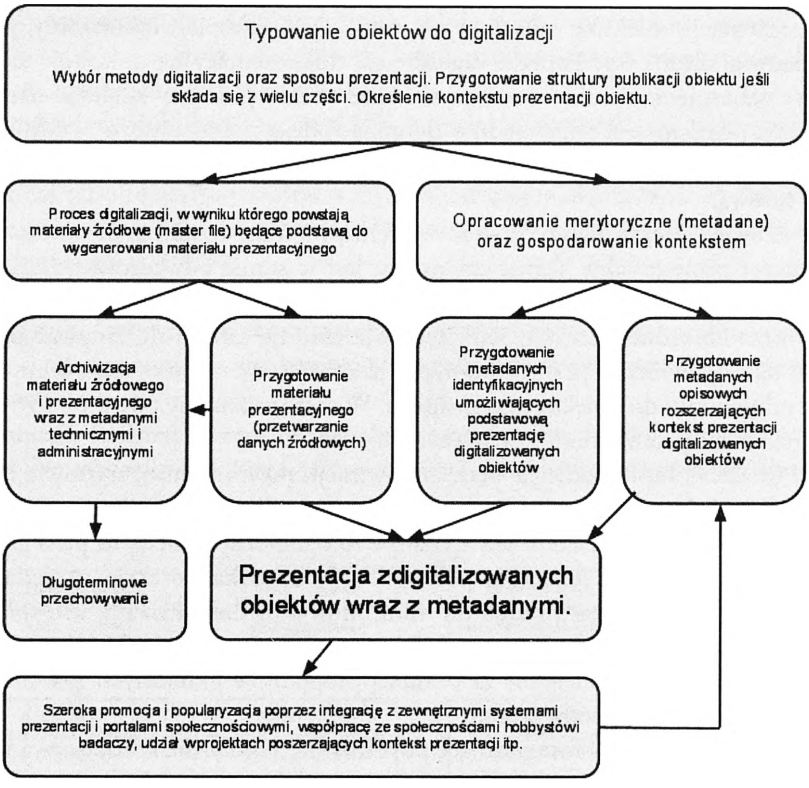
W przypadku podjęcia decyzji o uruchomieniu własnej pracowni digitalizacji trzeba dokonać kolejnego wyboru pomiędzy uniwersalnym stanowiskiem „All-in-One" (wszystko w jednym), a linią technologiczną lub skorzystać z obydwu rozwiązań.

Główne cechy charakteryzujące uniwersalne stanowiska „All-in-One" to:
- uniwersalność i funkcjonalność w szerokim zakresie;
- łatwość obsługi indywidualnych zleceń;
- możliwość realizacji zleceń wykraczających poza ramy planowanej digitalizacji;
- niska efektywność przy realizacji masowej digitalizacji;
- brak niezależnej analizy błędów powstałych podczas digitalizacji.

Natomiast główne cechy charakteryzujące linię technologiczną:
- dedykowana do konkretnych zadań i projektów;
- modułowa budowa, umożliwiająca tworzenie różnych konfiguracji procesu digitalizacji, w zależności od rodzaju obiektów i wielkości zasobu;
- wysoka efektywność przy realizacji masowej digitalizacji.

Wybór rozwiązania uzależniony jest w głównej mierze od skali przedsięwzięcia, jakie chcemy realizować. Stanowiska „All-in-One" sprawdzają się w przypadku digitalizacji niewielkich i bardzo zróżnicowanych zbiorów. Nie nadają się natomiast do realizacji dużych projektów digitalizacji, w których bardziej efektywne są linie technologiczne konfigurowane pod kątem digitalizacji, konkretnych typów obiektów, np. zbiorów kartograficznych, mikroform, rękopisów, książek drukowanych współcześnie itd.

Na rys. 3 zaprezentowano schemat blokowy, który w uproszczony sposób ilustruje ogólne zakresy prac wykonywane przy masowej digitalizacji i udostępnianiu zasobów cyfrowych oraz pokazuje ich przepływy w obrębie „podstawowej komórki".



Rys. 3. Schemat blokowy prac digitalizacyjnych.

Kolejną decyzją, którą należy podjąć i która ma wpływ na konfiguracje linii technologicznych jest wybór odpowiedniego modelu biblioteki cyfrowej dla digitalizowanych obiektów. Od tej decyzji uzależnione są kolejne, które dotyczą parametrów technicznych, sprzętu, oprogramowania, logistyki itd.

Czym są modele bibliotek cyfrowych? Aby odpowiedzieć na to pytanie posłużymy się cytatem z forum dyskusyjnego EBIB [http://ebib.oss.wroc.pl/phpBB/]. Remigiusz Lis napisał tam, iż: *(...) Podstawą określenia modeli bibliotek cyfrowych może być, jak sądzę, funkcja takiego zasobu, a więc funkcja informacyjna oraz dokumentacyjna. Przyjęcie takiego podziału pozwala na ustalenie „logik" obu rodzajów zasobu cyfrowego: określenie ich cech, sposobu pozyskania, przetwarzania, konstrukcji i prezentacji. (...)* [http://ebib.oss.wroc.pl/phpBB/viewtopic.php?t=636].

Biblioteki cyfrowe możemy budować na dwa sposoby – dokumentacyjny oraz informacyjny.

Model dokumentacyjny tworzony jest w celu zachowania wiernego wyglądu oryginału wraz ze wszystkimi jego szczegółami. Dedykowany jest głównie dla archiwizacji i prezentacji obiektów zabytkowych, takich jak: rękopisy, inkunabuły, grafiki, zbiory kartograficzne itp. Poprzez digitalizację dokumentacyjną dokonuje się również tzw. zabezpieczania zbiorów. Pozyskiwane obrazy powinny spełniać określone wymagania jakościowe z odpowiednio dobraną rodzielczością obrazu w odniesieniu do rzeczywistych wymiarów obiektu oraz możliwie najszerszą przestrzenią kolorów i głębią bitową.

Dla danego obiektu należy dobrać właściwą metodę digitalizacji, ponieważ nie każdy nawet profesjonalny skaner czy aparat jest w stanie odwzorować rzeczywisty obraz obiektu. Proces skanowania powinien być pozbawiony automatycznej korekcji ostrości oraz kolorów. Korekcji, jeśli jest to konieczne można dokonywać podczas kontroli na odpowiednio przygotowanym i skalibrowanym stanowisku komputerowym w odniesieniu do obiektu oryginalnego. Wnoszone zmiany powinny być protokołowane wraz z parametrami stosowanych filtrów.

Pliki (master), które podlegać będą archiwizacji, powinny mieć możliwie najbardziej źródłowy charakter, co uzależnione jest od możliwości urządzenia skanującego bądź fotografującego. Zatem dla aparatów fotograficznych będą to pliki „RAW", natomiast w przypadku skanerów pliki „TIFF". Wszelka korekcja o charakterze stratnym powinna być ograniczona do minimum, a błędny skan np. krzywy obraz należy skorygować na drodze ponownej digitalizacji danego ujęcia. Należy pamiętać, że większość działań na bitmapie za pomocą programów graficznych ma charakter stratny, a procesy te są nieodwracalne.

Urządzenia skanujące i fotografujące powinny być regularnie konserwowane i kalibrowane w celu zachowania możliwie najlepszej zdolności pozyskiwania obrazu cyfrowego.

Przykładem publikacji typowej dla modelu dokumentacyjnego może być rękopis udostępniany w Bibliotece Cyfrowej Uniwersytetu Wrocławskiego *Topographia oder Prodromus Delineati Silesiae Ducatus [...]* [http://www.bibliotekacyfrowa.pl/publication/7114].

Model informacyjny tworzony jest głównie w celu zachowania i udostępniania warstwy informacyjnej. Pliki źródłowe (master) oraz pliki prezentacyjne charakteryzują się stosunkowo niedużą objętością oraz wyodrębnioną warstwą tekstową umożliwiającą wyszukiwanie informacji.

Model ten jest właściwy dla archiwizacji i prezentacji publikacji książek, czasopism, publikacji 'born digital'.

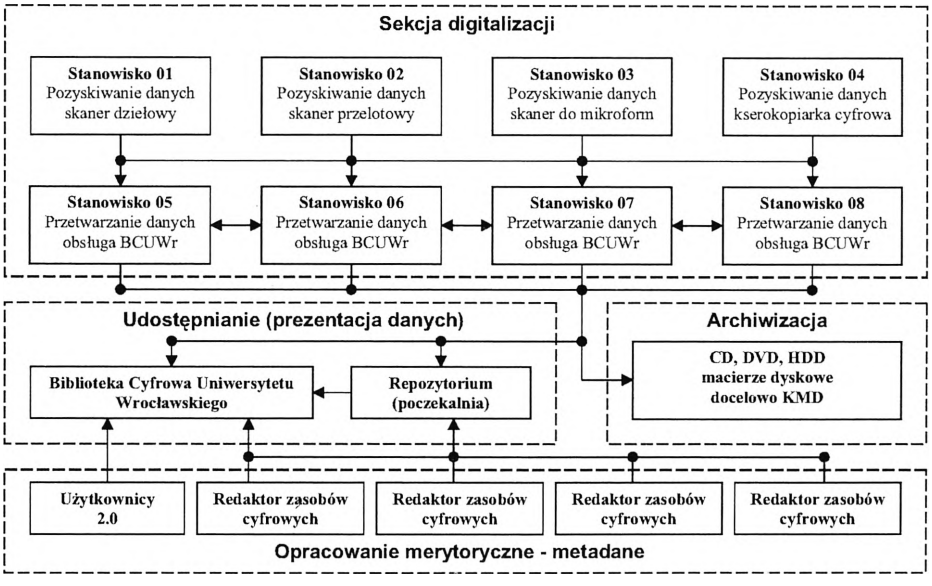
Jakość rozpoznanego tekstu jest w tym modelu kluczowym elementem, ponieważ od tego zależy skuteczność wyszukiwania informacji oraz możliwość korzystania z publikacji cyfrowej przy pomocy wszelkiego rodzaju urządzeń mobilnych takich, jak czytniki ebooków czy smartfony. Jest to możliwe dzięki zastosowaniu technologii OCR, która umożliwia automatyczne rozpoznanie tekstu w plikach graficznych. Jednak w wielu przypadkach automatyczne rozwiązania nie sprawdzają się i konieczne jest ręczne ingerowanie w proces, co w znaczny sposób wydłuża czas produkcji

publikacji cyfrowych. Pojawia się więc dylemat, czy digitalizować szybko, ale z ograniczeniami czy wolniej, ale bardziej profesjonalnie. Mając na uwadze dynamiczny rozwój technologii urządzeń mobilnych należy bardzo poważnie rozważyć, czy jest sens ponosić niższe koszty digitalizacji, której produkty za moment mogą okazać się mało praktyczne w użytkowaniu.

Przykładem publikacji typowej dla modelu informacyjnego może być książka znajdująca się w Prawniczej Bibliotece Cyfrowej [http://www.bibliotekacyfrowa.pl/dlibra/collectiondescription?dirids=53], wchodzącej w skład BCUWr *Gospodarowanie za pomocą planowania : analiza krytyczna narodów* [http://www.bibliotekacyfrowa.pl/publication/10554].

Nawiązując do zaproponowanego wcześniej modelu graficznego można skonstruować szczegółowy schemat blokowy obrazujący strukturę organizacyjną komórki zajmującej się tworzeniem, opracowaniem, udostępnianiem i archiwizowaniem zasobu cyfrowego.

Poniższy schemat prezentuje wdrożony model linii technologicznej na przykładzie Sekcji Digitalizacji i Mediów Cyfrowych Biblioteki Uniwersyteckiej we Wrocławiu.



Rys. 4. Model linii technologicznej.

Na schemacie można wyodrębnić cztery sekcje, skupiające zadania zilustrowane wcześniej na rys. 1 i 2. Sekcja digitalizacji składa się ze stanowisk do skanowania, którym towarzyszą stanowiska do przetwarzania danych. Aby efektywnie eksploatować urządzenia do digitalizacji, które często są bardzo drogie, nie powinno się obciążać osób i sprzętu wyznaczonych do obsługi tych urządzeń dodatkowymi zadaniami polegającymi na przetwarzaniu danych. W takich przypadkach najlepiej sprawdza się

tworzenie tandemów (skanowanie – przetwarzanie) ukierunkowanych na digitalizację konkretnych typów materiałów np. mikroform, zbiorów specjalnych, zbiorów współczesnych, dokumentów luźnych itd. Dla każdego z typów konfigurujemy tandemy, które połączone w sieć mogą być łatwo modyfikowane i adaptowane do różnych zadań.

W sekcji odpowiedzialnej za merytoryczne opracowanie danych cyfrowych wykonywane są wszelkie prace związane z tworzeniem metadanych, analizą potrzeb czytelników, integracją z katalogami i systemami umożliwiającymi przeszukiwanie zbiorów i precyzyjne docieranie do poszukiwanych informacji.

Udostępnianie i archiwizacja to sekcje typowo techniczne, które można określić jako zaplecze narzędziowe służące do prezentowania wyników pracy bibliotekarzy odpowiedzialnych za profesjonalne opracowanie danych cyfrowych oraz techników IT odpowiedzialnych za profesjonalne wyprodukowanie zasobów cyfrowych.

Udostępnianie zbiorów wymaga zastosowania rozwiązań informatycznych, dzięki którym prezentowane w bibliotekach treści będą mogły być dostępne w różnych kontekstach i dla jak największej liczby czytelników. Najpopularniejszym oprogramowaniem w Polsce umożliwiającym efektywne udostępnianie zbiorów na bardzo szeroką skalę jest system dLibra, który wykorzystuje powszechnie uznane standardy i protokoły takie, jak: RSS, RDF, MARC, DublinCore czy OAI-PMH.

Długoterminowa archiwizacja danych cyfrowych w bibliotekach jest dosyć sporym wyzwaniem logistycznym ponieważ poza zapewnieniem infrastruktury technicznej, która podlega ciągłemu rozwojowi konieczne jest też zapewnienie monitorowania starzenia się danych oraz migracji do nowych formatów. Być może w rozwiązaniu tego problemu pomogą prace nad uruchomieniem usług powszechnej archiwizacji polegających na udostępnieniu w skali kraju funkcjonalności zdalnej archiwizacji i backupu danych jako wartości dodanej do ogólnopolskiej akademickiej sieci naukowej PIONIER – http://www.platon.pionier.net.pl/online/archiwizacja.php.

Schemat blokowy ilustruje również sieć powiązań pomiędzy wymienionymi wcześniej sekcjami. Te powiązania są odwzorowaniem przepływu danych oraz informacji w sieciowej strukturze. Cały system organizacyjny składający się z poszczególnych sekcji i powiązań między nimi możemy porównać do lokalnej gospodarki sieciowej, która na zasadzie modułowej budowy może być łączona z innymi.

Aby taka pojedyncza gospodarka sieciowa mogła sprawnie funkcjonować konieczne jest dokumentowanie wyników realizowanych w niej zadań oraz monitorowanie efektywności poszczególnych stanowisk w celu zapewnienia im zsynchronizowanej współpracy. W tym celu warto skorzystać z narzędzi dostępnych online.

W wyżej wspomnianej instytucji zanim wdrożono własny system monitorowania digitalizacji korzystano z usług dostarczanych przez firmę Google. Przy pomocy usługi *Dokumenty Google* [http://www.google.com/google-d-s/intl/pl/tour1.html] możliwe jest zaprojektowanie narzędzi ułatwiających komunikację oraz monitorowanie procesów digitalizacji[1].

[1] Przykładowe szablony dokumentów można pobrać ze strony: http://www.tomasz.kalota.pl/monitorowanie-procesow-digitalizacji-przy-pomocy-dokumentow-google/

W pracy Sekcji Digitalizacji i Mediów Cyfrowych BUWr dokumenty Google wykorzystywane były do następujących celów:

1. Inwentarz zdigitalizowanych obiektów – arkusz kalkulacyjny, zbiera informacje takie, jak: sygnatury digitalizowanych dzieł, objętość oraz ilość plików, inicjały osób realizujących kolejne etapy pracy, data archiwizacji. Prowadzenie takiego inwentarza pozwala na szybkie dotarcie do informacji o zdigitalizowanych obiektach oraz osobach odpowiedzialnych za realizację poszczególnych zadań, co jest bardzo przydatne przy kontroli jakości pracy.

2. Monitoring efektywności digitalizacji – arkusz kalkulacyjny umożliwiający monitorowanie efektywności pracy osób zaangażowanych w realizację poszczególnych zadań. Składa się z 12 zakładek odpowiadających kolejnym miesiącom w roku. Każdą zakładkę tworzą komórki, w których technicy IT wpisują ilości wykonanych plików. Dane są automatycznie sumowane w każdym miesiącu, a na koniec roku robiony jest zbiorczy raport.

3. Monitoring kompletności – dokument, w którym osoby zaangażowane w prace nad digitalizacją i opracowaniem obszernych, wielotomowych zbiorów notują wszystkie spostrzeżenia dotyczące kompletności, błędów, braków itd. Podczas digitalizacji każdy obiekt jest bardzo dokładnie przeglądany i jest to jedyna okazja, aby zanotować wszelkie uwagi na jego temat, które po merytorycznym opracowaniu mogą zostać udostępnione czytelnikom.

Wspomniane dokumenty są edytowane online przez uprawnione osoby biorące udział w wytwarzaniu zasobów cyfrowych. Ponadto są one udostępnione do wglądu wszystkim osobom odpowiedzialnym za kwestie organizacyjne co pozwala na obserwowanie stanu prac w czasie rzeczywistym i nie wymaga sporządzania dodatkowych raportów.

Po latach zbierania informacji za pomocą arkuszy określono, jakie dane i parametry konieczne są do zarządzania procesem digitalizacji. W związku z tym zaprojektowana została baza danych, która wyeliminowała papierową dokumentację procesów digitalizacji, stała się inwentarzem zdigitalizowanych obiektów, a jednocześnie wspomaga i kontroluje przepływ pracy, dzięki właściwym statusom przydzielanym w poszczególnych fazach prac.

Nawiązując do wspomnianego na początku modelu graficznego, wyobraźmy sobie polski system rozproszonych bibliotek cyfrowych jako związek chemiczny, w którym pomiędzy poszczególnymi komórkami, obrazującymi rozproszone biblioteki cyfrowe, występują wiązania odpowiedzialne za ich scalenie. Wiązanie chemiczne, które według klasycznej definicji jest trwałym połączeniem dwóch atomów, pozwolimy sobie porównać do „faktycznego partnerstwa". Przy wiązaniach chemicznych mamy jeszcze do czynienia z dwoma istotnymi dla tego porównania zjawiskami. Jednym z nich jest kataliza – przyspieszanie reakcji chemicznych oraz inhibicja – spowalnianie tych reakcji. Owe porównania mają podkreślić znaczenie i wagę partnerstwa w organizowaniu digitalizacji, oraz udostępnianiu dziedzictwa kulturowego. Taka współpraca jest jednym z najistotniejszych elementów skutecznej logistyki.

W Polsce można już zaobserwować zaczątki takiej współpracy pomiędzy rożnymi instytucjami, czego efektem jest rozwiązanie kilku, wspólnych dla wszystkich bibliotek cyfrowych problemów. Jak się jednak okazuje tych wspólnych problemów do rozwiązania jest jeszcze bardzo wiele, a najważniejsze z nich i jednocześnie najtrudniejsze, jak np. długoterminowe przechowywanie, problemy związane z prawem autorskim, optymalizacja kosztów digitalizacji, standardy digitalizacji itp. wymagają ciągłego rozwoju i wykraczają poza możliwości pojedynczych instytucji. Bardzo ważne wydaje się więc wspieranie wspomnianego partnerstwa np. poprzez pielęgnowanie istniejących i poszukiwanie nowych katalizatorów potęgujących współdziałanie oraz unikanie inhibitorów, które hamują i utrudniają współpracę.

Bibliografia

1. Śpiechowicz, Dawid; Strokosz Ernest (2008). *Społeczna Pracownia Digitalizacji ŚBC po roku działalności* [dok. elektr.] *Biblioteka 2.0 : Blog społeczności czytelników i bibliotekarzy cyfrowych* http://blog.biblioteka20.pl/?p=77 [odczyt: 12.12.2011].
2. Koszowska, Agnieszka; Lis, Remigiusz (2008). *Biblioteka cyfrowa w wersji 2.0? : wybór możliwości* [dok. elektr.] *Małopolska Biblioteka Cyfrowa* http://mbc.malopolska.pl/publication/14693 [odczyt: 12.12.2011].

Część 4

Przykładowe polskie biblioteki cyfrowe (case studies)

Federacja Bibliotek Cyfrowych – studium przypadku

Cezary Mazurek, Marcin Werla

W tekście przedstawiono Federację Bibliotek Cyfrowych poprzez jej historię, zadania oraz stan aktualny. Omówiono również problematykę agregacji metadanych z rozproszonych systemów informacyjnych i zasady działania protokołu OAI-PMH, będącego technologiczną podstawą Federacji.

1. Wstęp

Dostępna publicznie od 22 czerwca 2007 r. Federacja Bibliotek Cyfrowych (FBC) jest bardzo ważnym składnikiem infrastruktury polskich bibliotek cyfrowych, ułatwiającym dotarcie do ich zasobów oraz wspomagającym na poziomie krajowym koordynację procesów digitalizacji. Z punktu widzenia użytkownika końcowego, zaprojektowana i zrealizowana przez Poznańskie Centrum Superkomputerowo-Sieciowe Federacja to przede wszystkim portal WWW [http://fbc.pionier.net.pl/] pozwalający na przeszukiwanie w jednym miejscu zasobów wielu instytucji pamięci przechowujących swoje zbiory w bibliotekach cyfrowych podłączonych do FBC. Z punktu widzenia współtwórców bibliotek cyfrowych, FBC to przede wszystkim narzędzie promocji udostępnianych online kolekcji oraz środowisko współpracy i wymiany informacji. Jest to też regularnie aktualizowana baza informacji na temat bibliotek cyfrowych w Polsce.

Celem niniejszego rozdziału jest zaprezentowanie Federacji Bibliotek Cyfrowych oraz różnych aspektów jej funkcjonowania, utrzymania i rozwoju. Punktem wyjścia do niniejszego studium przypadku będzie omówienie procesu agregacji informacji i związanych z nim problemów z tzw. interoperacyjnością, czyli wielopoziomową zgodnością pomiędzy różnymi systemami informatycznymi, która jest niezbędna do współpracy i wymiany informacji pomiędzy tymi systemami. Na tym tle, przedstawiony zostanie podstawowy dla realizacji procesu agregacji, mechanizm okresowego pobierania meta-

danych z wykorzystaniem protokołu OAI-PMH, który jest techniczną podstawą funkcjonowania FBC i wielu innych podobnych serwisów na całym świecie. W dalszej części rozdziału omówiona zostanie geneza FBC, wyjaśniająca co było podstawą uruchomienia agregatora treści bibliotek cyfrowych jakim jest FBC i co wpłynęło na jej obecny kształt. Kolejna część niniejszego rozdziału poświęcona będzie omówieniu zasad funkcjonowania FBC zarówno na poziomie organizacyjnym jak i technicznym. Rozdział kończy się odniesieniem do współpracy międzynarodowej, w którą zaangażowana jest Federacja oraz nakreśleniem kierunków jej dalszego rozwoju.

2. Wprowadzenie do problematyki agregacji informacji

Praktycznie każda osoba korzystająca z Internetu zetknęła się – świadomie bądź nie – z agregacją informacji. Natura Internetu jest rozproszona. Nie ma jednego, oficjalnego i centralnego punktu dostępu – zamiast tego mamy sieć połączonych ze sobą ponad 100 milionów serwerów[1], z którymi nieustannie komunikują się komputery użytkowników poszukujących interesujących dla siebie informacji. Można sobie wyobrazić jak trudne byłoby to zadanie gdyby rozproszona informacja nie była nigdzie agregowana. Z pomocą przychodzą tutaj wyszukiwarki internetowe, z których najpopularniejsza obecnie to oczywiście Google. Misją tej firmy jest *uporządkowanie światowych zasobów informacji, tak aby stały się powszechnie dostępne i użyteczne* [http://www.google.com/about/corporate/company/]. Podstawowym sposobem działania wyszukiwarek internetowych ogólnego przeznaczenia jest gromadzenie (agregowanie) informacji na temat zawartości jak największej liczby stron WWW i umożliwienie przeszukiwania tych informacji w jednym miejscu. Podstawowym problemem natomiast jest ogromne zróżnicowanie agregowanych stron WWW – strony te, nawet jeżeli dotyczą tej samej tematyki, mogą być opracowywane w różnych technologiach, cechuje je zróżnicowany rozkład tekstu i funkcji, a umieszczone na nich informacje udostępniane są w różnych językach czy formatach danych. Dlatego też nawet dla firmy z tak ogromnym potencjałem technologicznym i finansowym jak Google, rola internetowego przewodnika jest nie lada wyzwaniem. Każda inna wyszukiwarka, stosująca własne rozwiązania dla agregacji i przeszukiwania informacji, może okazać się poza tym skuteczniejsza w odniesieniu do szczególnych potrzeb użytkownika. Z tego powodu profesjonalni brokerzy informacji w swojej pracy nigdy nie ograniczają się wyłącznie do jednej wyszukiwarki.

Kolejnym przykładem usług bazujących na agregacji informacji są różnego rodzaju serwisy porównujące ceny określonych produktów, takie jak na przykład popularne w Polsce Ceneo.pl oraz Nokaut.pl, czy jeden z pierwszych tego typu serwisów na świecie – StreetPrices.com. Serwisy te od omówionych powyżej ogólnych wyszukiwarek odróżniają dwa zasadnicze elementy: mniejszy zakres i większe ustrukturyzowanie informacji. Nawet w przypadku ogólnych porównywarek cen, zawsze mamy do czynienia z informacjami o pewnych materialnych bądź niematerialnych produktach odznaczających się

[1] Comiesięczne statystyki na temat liczby serwerów oraz nazw domenowych w Internecie są publikowane przez firmę Netcraft pod adresem: http://news.netcraft.com/archives/category/web-server-survey/.

jakimś zestawem cech. Część z tych cech jest zależna od rodzaju produktu, a część (np. nazwa, producent, sprzedawca, cena) występuje zawsze. Celem takich serwisów jest ułatwienie użytkownikowi odnalezienia produktu spełniającego zadane kryteria, jednym z których jest zazwyczaj najniższa cena. To, co jest podstawową trudnością w konstruowaniu serwisów porównujących ceny, a zarazem czyni te serwisy potrzebnymi, jest fakt iż dany produkt może być oferowany przez więcej niż jednego sprzedawcę, a ponadto każdy sprzedawca w swoim sklepie internetowym może opisywać ten produkt w trochę inny sposób – np. inną terminologią czy zestawem informacji o innej strukturze. Stąd też niezbędna jest agregacja i ujednolicenie danych do poziomu pozwalającego na operowanie na jednym spójnym zbiorze danych.

Przykładem może być tutaj opis tego samego modelu telefonu komórkowego publikowany przez dwa różne sklepy internetowe (por. Tabela 1). W opisie po lewej stronie aspekty komunikacyjne telefonu rozbito na kilka atrybutów produktu (*Bluetooth, IrDA, W-Lan, GPRS, EDGE, UMTS*), a w opisie po prawej stronie te same informacje zawarto częściowo w innych polach (Łącza/Komunikacja, Zakresy częstotliwości pracy, Internet). W każdym z opisów mamy też pola o identycznej etykiecie i wartości (np. pola *Odtwarzacz MP3* oraz *Dyktafon*), bardzo podobnych etykietach i identycznej zawartości (*Radio FM:* tak; *Radio:* tak) oraz bardzo podobnych etykietach i znacznie rozbieżnej zawartości (*Wyświetlacz (rodzaj): TFT dotykowy; Wyświetlacz: 4,3", 16 mln kolorów...*). W tym momencie nasuwa się pytanie, w jaki sposób można stworzyć efektywną porównywarkę cen (czyli inaczej patrząc agregator informacji o ofertach sklepów internetowych, z wartością dodaną w postaci usługi porównywania tych ofert) przy tak zróżnicowanym zestawie potencjalnie interesujących informacji pochodzących z wielu źródeł. Przykład ten ilustruje problemy z jakimi zmierzyć się musi twórca agregatora na przedstawionych poniżej, kolejnych etapach jego budowy.

Tabela 1. Oferta dot. tego samego modelu telefonu komórkowego w dwóch różnych sklepach internetowych (po lewej stronie cały opis techniczny z www.hesjan.pl, po prawej fragment opisu technicznego z www.euro.com.pl, dane z dnia 08.08.2011 r.).

Waga	**116 g**
Kolor	**Czarny**
Bluetooth	**Tak**
IrDA	**Nie**
W-Lan	**Tak**
Wbudowany aparat cyfrowy	**Tak**
Wyświetlacz (rodzaj)	**TFT dotykowy**
Odtwarzacz MP3	**Tak**
Obsługiwane karty pamięci Flash	**micro SD**
GPRS	**Tak**
EDGE	**Tak**
UMTS	**Tak**
Dyktafon	**Tak**
Tryb głośno mówiący	**Tak**
Połączenia Video	**Tak**
Odtwarzacz Video	**Tak**
Radio FM	**Tak**

Wyświetlacz	4.3", 16 mln kolorów SUPER AMOLED Plus, 480 x 800 pikseli
Łącza/Komunikacja	Bluetooth 3.0, WiFi, USB, słuchawkowe
Zakresy częstotliwości pracy	1800, 1900, 850, 900, HSDPA, UMTS, WCDMA 2100
Rodzaj obudowy	tradycyjna - ekran dotykowy
Kolor obudowy	czarny
System operacyjny	Android 2.3 Gingerbread
Pamięć wbudowana	16 GB
Procesor	Cortex A9 Dual Core 1200 MHz
Menu w języku polskim	✓ tak
Funkcje multimedialne	
Rozdzielczość aparatu	8 mln pikseli
Funkcje aparatu	autofocus, efekty: negatyw, sepia i inne, geotagging - znaczanie zdjęć i wideo danymi geograficznymi, samowyzwalacz, wykrywanie twarzy, wykrywanie uśmiechu
Nawigacja GPS	✓ tak
Odtwarzacz MP3	✓ tak
Radio	✓ tak
Gry	✓ tak
Rozmowy wideo	✓ tak
Funkcje telefonu	
Standardy wysyłania/odbierania wiadomości	e-mail, EMS, MMS, tekstowe SMS
Internet	JAVA - tak, WAP - tak, GPRS - tak, EDGE - tak
Tryb głośnomówiący	✓ tak
Wybieranie głosowe	✓ tak
Dyktafon	✓ tak

Pierwszym krokiem w kierunku zbudowania efektywnego agregatora informacji powinno być określenie, jakie funkcje ma on oferować użytkownikom końcowym (np. jakie powinny być możliwe kryteria wyszukiwania czy filtrowania informacji), a następnie zaprojektowanie uniwersalnych struktur dla agregowanych informacji, które są niezbędne aby te funkcje zaimplementować. Drugi krok to określenie źródeł danych, które będą agregowane oraz określenie sposobu przetwarzania informacji ze struktury charakterystycznej dla źródła danych do struktury charakterystycznej dla agregatora informacji. W trakcie opracowywania tego sposobu przetwarzania informacji może okazać się, że struktury po stronie agregatora i po stronie źródła danych są na tyle rozbieżne, iż konieczne jest zmodyfikowanie struktury zaprojektowanej dla agregatora. Może być też tak, że środowisko dostawców danych jest tak homogeniczne, że wybór struktury danych dla agregatora jest oczywisty. Ostatnim krokiem jest określenie, w jaki sposób będą pobierane dane i na jakim etapie będzie następowała ich transformacja. W wariancie optymalnym źródło danych może dostosować się do agregatora i udostępnić mu informacje o oczekiwanej strukturze przy pomocy ustalonego protokołu komunikacyjnego[2]. Wariant gorszy polega na tym, że agregator samodzielnie pobiera i transformuje informacje pochodzące wyłącznie z publicznie dostępnych stron WWW[3], które w dodatku mogą być nieprzystosowane do automatycznego indeksowania (i stanowić tym samym część tzw. głębokiego czy ukrytego Internetu). Proces agregowania informacji może być w pełni zautomatyzowany, ale może też bazować na pracy ludzkiej (np. korekcie błędów powstałych wskutek automatycznego przetwarzania danych) wykonywanej przez opłacone w tym celu osoby lub np. przez użytkowników danego serwisu internetowego działających na zasadzie crowdsourcingu[4]. W nawiązaniu do przytoczonego wcześniej przykładu schematów opisu telefonów komórkowych, można zapoznać się ze stroną http://www.ceneo.pl/Telefony_komorkowe[5], na której w dostępnym po lewej stronie menu „Opcje przeglądania" widać przykładową strukturę informacji opracowaną na poziomie agregatora w celu umożliwienia zaawansowanego filtrowania ofert sprzedaży. Jest ona odmienna od struktur każdego z przykładów z Tabeli 1, a mimo to obydwie te oferty są w serwisie Ceneo.pl ujęte.

Wszystkie omówione i zilustrowane powyżej problemy występują w bardzo zbliżonej formie również w przypadku agregatora takiego jak FBC, który funkcjonuje w domenie informacji dotyczących obiektów udostępnianych w bibliotekach cyfrowych. Kwestie te omówiono dokładniej w dalszych punktach tego rozdziału.

[2] Choć oczywiście może się okazać, że w konkretnym przypadku sama struktura informacji i protokół komunikacyjny to za mało i niezbędny jest wyższy poziom interoperacyjności [Nilsson 2009].

[3] Oczywiście działania tego typu powinny być realizowane z poszanowaniem praw autorskich i praw chroniących bazy danych, jednak omawianie tej problematyki wykracza poza zakres niniejszego rozdziału.

[4] Zainteresowani tą tematyką mogą zapoznać się z artykułem *Price comparison service* [2012], w szczególności z sekcją „Technology" tego artykułu.

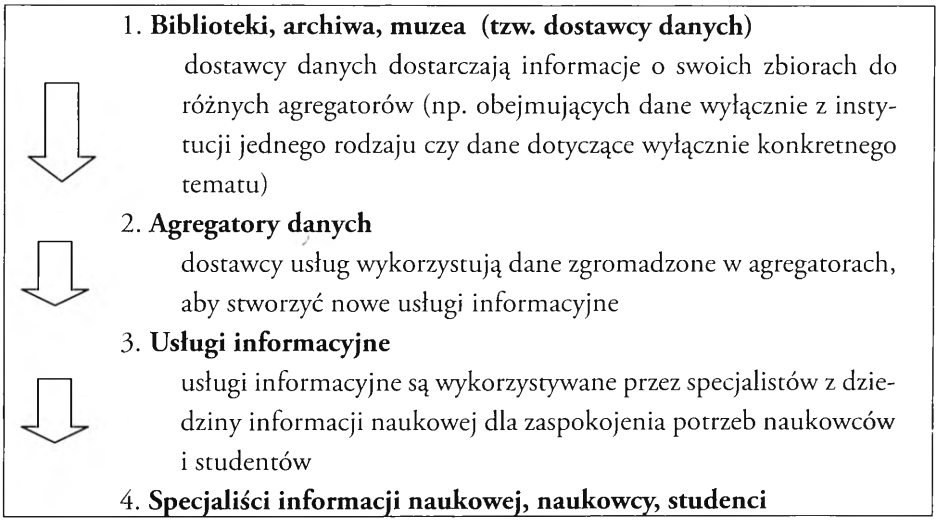
[5] Odczyt 8.08.2011 r.

3. Agregacja w świecie informacji naukowej – agregacja jako taktyka

Zespół Zadaniowy ds. Odkrywania Zasobów (ang. Resource Discovery Task Force) [strona Zespołu Zadaniowego dostępna jest pod adresem http://rdtf.jiscinvolve.org/] powołany przez JISC [Joint Information Systems Committee – http://www.jisc.ac.uk/] i RLUK [Research Libraries UK – http://www.rluk.ac.uk/] w połowie 2010 r. postawił sobie następującą wizję działania:

Naukowcy i studenci w Wielkiej Brytanii będą mieć łatwy, elastyczny i bieżący dostęp do treści i usług, dzięki zagregowanemu i zintegrowanemu systemowi szkieletowemu odkrywania i dostarczania zasobów, który jest wszechstronny, otwarty i możliwy do utrzymywania [JISC 2010].

W praktyce wizja ta oznacza, iż działania podejmowane przez wspomniany zespół zadaniowy mają skupiać się na ułatwianiu odkrywania zasobów zarówno przez ludzi, jaki i przez automaty, wykorzystując w tym celu szereg powiązanych ze sobą koncepcji technologicznych [Burnhill; Macdonald 2011]. Według członków tego zespołu jedną z takich koncepcji jest właśnie agregacja danych, nastawiona na wykorzystanie tych danych nie tylko przez ludzi, ale również przez automaty i rozumiana znacznie szerzej, niż tylko jako kopiowanie informacji z rozproszonych źródeł. Agregacja nie jest tutaj celem samym w sobie, ale taktyką służącą strategicznemu celowi poprawy jakości i dostępności danych oraz zwiększenia ich widoczności w sieci – zarówno w popularnych serwisach ogólnego przeznaczenia, jak i w serwisach profesjonalnych. W praktyce efektem docelowym może być przepływ informacji przedstawiony na rysunku 1.



Rys. 1. Przykładowy przepływ informacji od instytucji pamięci, przez agregatory danych i usługi informacyjne, aż do użytkowników końcowych – opracowane na podstawie (JISC, 2010).

Te sformułowania odzwierciedlają kierunki praktycznych działań podejmowane już wcześniej w Europie i w Polsce. Przykładami mogą tutaj być inicjatywy takie jak Europeana [http://europeana.eu/], DRIVER [http://search.driver.research-infrastructures.eu/], DART-Europe [http://www.dart-europe.eu], Scientific Commons [http://en.scientificcommons.org/] czy zaprezentowana w dalszej części tego rozdziału Federacja Bibliotek Cyfrowych. Mimo tego, iż każda z tych inicjatyw ma swój unikalny cel i każda została pierwotnie stworzona w innym środowisku, wszystkie one wykorzystują w agregacji tę samą podstawę technologiczną. Jest nią protokół OAI-PMH, który pokrótce omówiony jest w następnym punkcie.

4. Podstawy protokołu OAI-PMH

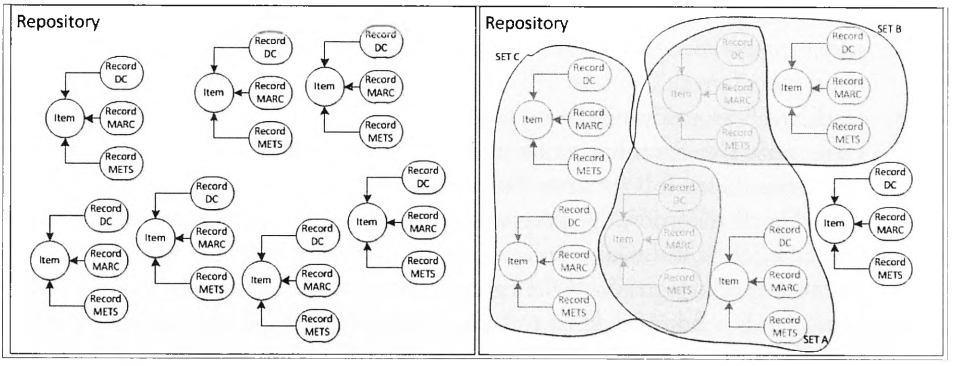
W przedstawionym na rysunku 1 schemacie przepływu danych, na poziomie usług informatycznych (sieciowych) można wyróżnić dwie strony tego przepływu – dostawców danych (ang. *data providers*) przesyłających dane do agregatorów oraz dostawców usług (ang. *service providers*), którzy wykorzystują dane z agregatorów do budowy bardziej zaawansowanych narzędzi. Model ten zawarty został również w specyfikacji protokołu OAI-PMH (Lagoze, 2008) – najpopularniejszego protokołu komunikacyjnego wykorzystywanego do okresowego pobierania danych z bibliotek cyfrowych[6]. Poniżej przedstawione zostaną dwa najistotniejsze elementy specyfikacji tego protokołu: model danych oraz model komunikacji. W ramach podstawowych informacji technicznych warto jeszcze dodać, że nośnikiem komunikatów protokołu OAI-PMH jest protokół HTTP, a pobierane dane są zapisywane w języku XML. Połączenie dwóch tak popularnych w Internecie technologii znacznie ułatwia implementację OAI-PMH. Osoby zainteresowane szczegółowymi informacjami oraz aspektami implementacyjnymi powinny zapoznać się ze specyfikacją protokołu.

4.1. Model danych

Specyfikacja OAI-PMH zakłada, że repozytorium *(ang. repository)* pełniące rolę dostawcy danych, poprzez interfejs zgodny z tym protokołem, udostępnia informacje na temat zasobów dowolnego rodzaju (ang. *item*). Nie muszą to być obiekty cyfrowe – równie dobrze może to być baza danych na temat obiektów fizycznych. Każdy taki obiekt powinien po stronie jego dostawcy posiadać swój opis (metadane). Opis ten może być wyrażony w jednym bądź wielu schematach (schematach metadanych). W ramach protokołu OAI-PMH opis konkretnego obiektu w konkretnym schemacie nazywa się rekordem *(ang. record)*, i składa się z nagłówka i właściwego opisu. Nagłówek rekordu zawiera podstawowe informacje takie jak identyfikator obiektu, czy data ostatniej modyfikacji opisu.

[6] Inne wykorzystywane do tego typu zastosowań protokoły komunikacyjne to m.in. RSS oraz ATOM.

Wg specyfikacji protokołu, każdy obiekt na temat którego można pobrać informacje przy pomocy protokołu OAI-PMH, musi posiadać metadane przynajmniej w schemacie Dublin Core Metadata Element Set (DCMI, 2010). Ponadto dostawca danych może udostępniać rekordy metadanych obiektu w dowolnych innych schematach (por. rysunek 2, część lewa). Innym opcjonalnym elementem protokołu jest obsługa usuniętych rekordów, polegająca na tym że dostawca danych może udostępniać dodatkowo nagłówki rekordów, które zostały usunięte wraz z datą usunięcia przesyłaną jako data ostatniej modyfikacji danego rekordu. Informacje tego typu znacznie ułatwiają okresową synchronizację danych pomiędzy dostawcą i agregatorem.



Rys. 2. Wizualizacja modelu danych protokołu OAI-PMH. Lewy schemat pokazuje repozytorium danych z obiektami posiadającymi metadane w trzech schematach (oznaczonych jako DC, MARC, METS). Prawy schemat pokazuje model zbiorów w repozytorium. Część obiektów należy do jednego bądź kilku zbiorów, jest też obiekt nienależący do żadnego zbioru.

Model danych OAI-PMH zawiera jeszcze jeden istotny, aczkolwiek opcjonalny komponent nazwany zbiorem *(ang. set)*. Zbiór w OAI-PMH to po prostu grupa obiektów. Każdy obiekt może (ale nie musi) należeć do jednego lub więcej zbiorów, a zbiory mogą być zorganizowane w hierarchię (por. rysunek 2, część prawa). Zbiory mają na celu umożliwienie agregatorom wykonywania tzw. operacji selektywnego pobierania metadanych, która sprowadza się w praktyce do pobrania metadanych obiektów należących wyłącznie do konkretnego zbioru.

4.2. Model komunikacji

Komunikacja w protokole OAI-PMH opiera się na sześciu rodzajach poleceń czy żądań (nazwanych w specyfikacji czasownikami – *ang. verbs*), jakie agregator może przesłać do dostawcy danych. W odpowiedzi na żądanie dostawca danych przesyła odpowiedź zawierającą stosowne informacje, bądź też komunikat o błędzie. Specyfikacja protokołu definiuje następujące rodzaje żądań:

- GetRecord – odpowiedź na to żądanie zawiera pojedynczy rekord, którego identyfikator został podany w żądaniu. Agregator powinien też podać w swoim żądaniu schemat metadanych rekordu.

- Identify – odpowiedź na to żądanie zawiera podstawowe informacje o dostawcy danych, takie jak nazwę, adres URL, adres e-mail administratora czy datę modyfikacji najstarszego rekordu.
- ListMetadataFormats – odpowiedź na to żądanie powinna zawierać informacje o formatach metadanych, jakie są obsługiwane przez dostawcę danych. Każdy format powinien posiadać unikalny identyfikator, który może być wykorzystywany przy innych żądaniach np. ListIdentifiers.
- ListSets – odpowiedź na to żądanie powinna zawierać listę zbiorów zdefiniowanych przez dostawcę danych. Każdy ze zbiorów powinien posiadać unikalny identyfikator, który może być wykorzystywany przy innych żądaniach, np. ListIdentifiers. Poza tym zbiór może być opisany przy pomocy dowolnego schematu metadanych.
- ListIdentifiers – odpowiedź na to żądanie zawiera listę nagłówków rekordów. Lista może być ograniczona na podstawie przesłanego wraz z żądaniem identyfikatora zbioru lub też przez podanie górnej i/lub dolnej granicy zakresu dat modyfikacji rekordów. W przypadku długiej listy wyników możliwe jest podzielenie jej na pobierane kolejno części, do których dostęp możliwy jest po podaniu specjalnego wystawianego przez dostawcę danych tokenu.
- ListRecords – odpowiedź na to żądanie zawiera listę pełnych rekordów metadanych. Analogicznie do ListIdentifiers, lista ta może być ograniczona na podstawie przesłanego wraz z żądaniem identyfikatora zbioru lub też przez podanie górnej i/lub dolnej granicy zakresu dat modyfikacji rekordów. Analogiczny jest również mechanizm stronicowania wyników.

Przykładowy scenariusz komunikacji pomiędzy agregatorem a dostawcą danych może wyglądać następująco:

- Pobranie przez agregator podstawowych informacji o dostawcy danych (*Identify*).
- Pobranie listy zdefiniowanych zbiorów w celu zidentyfikowania zbiorów zawierających dane potencjalnie interesujących rekordów (*ListSets*).
- Pobranie listy formatów metadanych obsługiwanych przez dostawcę danych (*ListMetadataFormats*).
- Pobranie listy identyfikatorów obiektów z określonego zbioru (*ListIdentifiers* z użyciem identyfikatora zbioru).
- Dla każdego obiektu z listy zwróconej w odpowiedzi na *ListIdentifiers* pobranie metadanych we wszystkich dostępnych formatach – poprzez wydanie żądania *GetMetadataRecord* z użyciem identyfikatora rekordu i identyfikatora formatu metadanych.

Praktyczne przetestowaniem zasad działania protokołu OAI-PMH można przeprowadzić np. przy pomocy narzędzie OAI Repository Explorer, które dostępne jest pod adresem http://re.cs.uct.ac.za/. Do działania tego narzędzia niezbędne jest podanie adresu interfejsu OAI-PMH. Dla przykładu, taki adres dla Wielkopolskiej Biblioteki Cyfrowej to http://www.wbc.poznan.pl/dlibra/oai-pmh-repository.xml.

Celem powyższego skróconego opisu protokołu OAI-PMH było przedstawienie realiów technicznych, w jakich odbywa się wymiana danych pomiędzy dostawcami

danych, a agregatorem. Dalsza część tekstu jest poświęcona Federacji Bibliotek Cyfrowych, która wykorzystuje ten protokół, zarówno do pobierania jak i do udostępniania metadanych.

5. Powstanie Federacji Bibliotek Cyfrowych

Dynamiczny rozwój bibliotek cyfrowych w Polsce, który rozpoczął się w pierwszych latach obecnego stulecia jest bardzo silnie powiązany z rozwojem oprogramowania dLibra (Mazurek 2005). Oprogramowanie to wspomaga budowę bibliotek cyfrowych i od 1999 r. jest rozwijane z myślą o maksymalnym ułatwieniu instytucjom kultury i nauki udostępniania ich zbiorów w Internecie. Zakres funkcjonalny tego oprogramowania jest szeroki (Dudczak, 2007), jednak w kontekście tematu agregacji danych najistotniejsze są funkcje udostępniania metadanych obiektów cyfrowych składowanych w dLibrze, w sposób umożliwiający ich automatyczne pobieranie i przetwarzanie. Od wersji 1.6 wydanej w kwietniu 2005 r. dLibra posiada interfejs OAI-PMH [Werla 2005]. Pierwotnie koncepcja wymiany informacji pomiędzy bibliotekami cyfrowymi funkcjonującymi w sieci PIONIER opracowana przez Zespół Bibliotek Cyfrowych PCSS zakładała, iż każda z tych bibliotek posiada możliwość udostępniania oraz pobierania danych przy pomocy protokołu OAI-PMH, a także że wszystkie biblioteki wymieniają się między sobą informacjami o zbiorach w sposób w pełni zautomatyzowany. Koncepcja ta została w praktyce zaimplementowana w lutym 2006 roku, wraz z wydaniem wersji 2.2 systemu dLibra.

Dynamiczny rozwój bibliotek cyfrowych owocował nieustannym i coraz szybszym wzrostem liczby dostępnych obiektów cyfrowych. Oznaczało to konieczność przetwarzania coraz większej ilości informacji przez każdą z bibliotek cyfrowych, które gromadziły komplet informacji o zasobach dostępnych w całej sieci bibliotek. Z drugiej strony, nieskomplikowany w swej naturze proces uruchomienia biblioteki cyfrowej w oparciu o oprogramowanie dLibra i łatwa dostępność tego oprogramowania spowodowały, że tego typu inicjatywy przestały być już domeną dużych instytucji posiadających silne zaplecze technologiczne. Przykładem tego może być uruchomiona w marcu 2007 r. Wejherowska Biblioteka Cyfrowa [http://biblioteka.wejherowo.pl/dlibra/], utrzymywana przez Powiatową i Miejską Bibliotekę Publiczną im. Aleksandra Majkowskiego w Wejherowie. Oczywiście tego typu instytucje posiadały infrastrukturę w zupełności wystarczającą do udostępniania ich własnych kolekcji. Niestety informacje na temat zasobów całej sieci bibliotek stały się z czasem zbyt duże do efektywnego przetwarzania w każdym węźle tej sieci.

Dlatego też już pod koniec 2005 r. zaczęto planować drugi etap kształtowania infrastruktury bibliotek cyfrowych w Polsce. Podstawą tego rozwoju była koncepcja utworzenia jednego agregatora metadanych dla polskich bibliotek cyfrowych, który służyłby za punkt informacyjny zarówno dla użytkowników końcowych, jaki i dla wszystkich współpracujących bibliotek, likwidując tym samym konieczność gromadzenia i przetwarzania tak dużego zbioru danych równolegle w każdej z bibliotek

[Werla 2005]. Prototypowa implementacja tej koncepcji zrealizowana została w latach 2006-2007 poprzez wydzielenie z systemu dLibra usługi odpowiadającej za agregację metadanych przy pomocy protokołu OAI-PMH i zbudowanie na podstawie takiej usługi dedykowanego portalu dostępowego do danych agregowanych z całej sieci bibliotek. Publiczna premiera portalu FBC odbyła się 22 czerwca 2007 r. na spotkaniu Rady Konsorcjum PIONIER [http://www.pionier.net.pl/online/pl/jednostki/]. Portal FBC w swojej pierwszej postaci dawał możliwość podstawowego i zaawansowanego przeszukiwania metadanych około 80 000 obiektów pochodzących z 15 bibliotek cyfrowych [Lewandowska 2007] oraz lokalizowania tych obiektów na podstawie automatycznie przydzielanego, unikalnego identyfikatora OAI, zgodnego ze specyfikacją opracowaną w ramach przewodnika dla implementacji protokołu OAI-PMH [Lagoze 2006]. Dodatkowo agregowane i udostępniane były również informacje na temat obiektów przeznaczonych do digitalizacji.

W ciągu kilku lat swojego rozwoju zakres funkcjonalny Federacji Bibliotek Cyfrowych rozszerzył się znacznie, pociągając za sobą również wzrost znaczenia FBC. Kluczowe okazały się tutaj również powstałe po utworzeniu FBC międzynarodowe inicjatywy takie jak Europeana czy DART-Europe, dla których FBC stało się naturalnym partnerem.

6. Zakres i zasady działania FBC

Misją Federacji Bibliotek Cyfrowych jest:

- ułatwienie wykorzystania zasobów polskich bibliotek cyfrowych i repozytoriów,
- zwiększenie widoczności i popularyzacja zasobów polskich bibliotek cyfrowych i repozytoriów w Internecie,
- udostępnienie użytkownikom Internetu i twórcom bibliotek cyfrowych nowych, zaawansowanych usług sieciowych opartych na zasobach polskich bibliotek cyfrowych i repozytoriów.

Ponadto, u podstaw funkcjonowania FBC leżą następujące podstawowe założenia organizacyjne:

- współpraca z FBC nie wymaga przenoszenia obiektów cyfrowych z bibliotek cyfrowych do FBC – udostępnianie obiektów użytkownikom końcowym odbywa się zawsze poprzez strony źródłowe biblioteki cyfrowej;
- nie ma opłat za korzystanie z FBC ani za dostarczanie danych do FBC;
- otwarte standardy są podstawą wymiany danych dla realizacji wszystkich funkcji, jakie FBC udostępnia – dotyczy to zarówno pobierania danych z bibliotek cyfrowych, jak i udostępniania danych przez FBC.

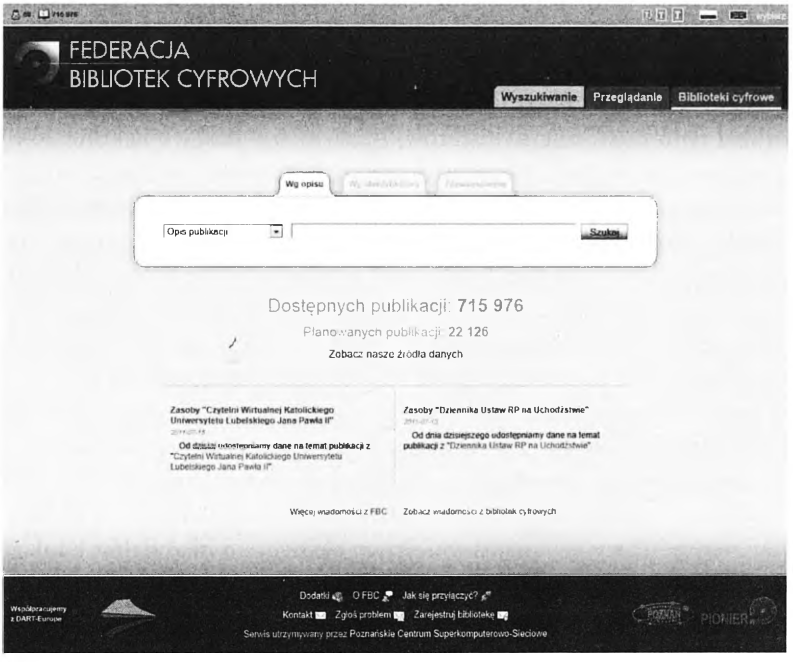
Pod koniec lipca 2011 r. Federacja Bibliotek Cyfrowych agregowała informacje z 70 bibliotek cyfrowych, a kolejne 14 bibliotek było już zarejestrowanych i przygotowywało się do przyłączenia. Łączna liczba obiektów cyfrowych udostępnianych przez te biblioteki była zbliżona do 715 tysięcy. Funkcje oferowane użytkownikom końcowym FBC, realizowane w dużej mierze na podstawie tych setek tysięcy opisów można

podzielić na trzy grupy, które mają też swoje odzwierciedlenie w interfejsie graficznym FBC (por. rysunek 3, zakładki w prawym górnym narożniku). Grupy te to:

- Funkcje przeszukiwania metadanych zagregowanych przez FBC,
- Funkcje przeglądania bazy metadanych FBC,
- Funkcje związane z bazą informacji na temat bibliotek cyfrowych.

Pierwsza z grup umożliwia proste oraz zaawansowane przeszukiwanie metadanych oraz lokalizowanie obiektów cyfrowych na podstawie ich identyfikatora i jest zbliżona do pierwotnych możliwości oferowanych przez FBC. Otrzymane wyniki można filtrować na podstawie źródeł danych oraz stanu obiektu (dostępny czy planowany do digitalizacji), a ponadto są one wzbogacone wynikami z Europeany [Dudczak 2011].

Grupa funkcji związanych z przeglądaniem bazy metadanych FBC przygotowana została z myślą o użytkownikach, którzy nie mają jasno sprecyzowanych kryteriów odnośnie do interesujących ich publikacji i raczej woleliby móc wybrać potencjalnie ciekawe obiekty ze wstępnie przygotowanych zestawów. Użytkownikom takim FBC pozwala na zapoznanie się z listą obiektów ostatnio udostępnionych w polskich bibliotekach cyfrowych (wraz z pełnym archiwum codziennych zmian w tych bibliotekach), dostępna jest również aktualizowana praktycznie w czasie rzeczywistym lista obiektów ostatnio przeglądanych przez użytkowników Federacji oraz generowana raz dziennie lista obiektów najpopularniejszych w ciągu ostatnich 30 dni. Ponadto dostępna jest lista obiektów rekomendowanych przez redaktorów poszczególnych bibliotek cyfrowych (każda z bibliotek może wskazywać jednocześnie trzy takie obiekty).



Rys. 3. Strona główna Federacji Bibliotek Cyfrowych [http://fbc.pionier.net.pl/, odczyt: 08.08.2011].

Trzecia, najbardziej zróżnicowana grupa funkcji dotyczy informacji na temat bibliotek cyfrowych. Na tę grupę składają się:

- Baza bibliotek cyfrowych funkcjonujących w Polsce – dla każdej z bibliotek zawiera informacje na temat nazwy, adresu WWW, adresu kontaktowego, stanu biblioteki cyfrowej (w bazie mogą znajdować się również biblioteki będące w fazie przygotowania do publicznego udostępnienia), opisu, listy instytucji współtworzących daną bibliotekę cyfrową (wraz z danymi kontaktowymi i lokalizacją geograficzną), listy polecanych obiektów oraz podstawowego zestawu informacji technicznych. Podstawowym źródłem danych jest tu personel biblioteki cyfrowej, który posiada uprawnienia edycyjne rekordu opisującego daną bibliotekę w bazie. Obsługa FBC pełni rolę moderatorów bazy. Informacje z bazy prezentowane są w trzech postaciach:
  - zbioru stron WWW zawierającego komplet dostępnych informacji;
  - pliku CSV zawierającego podstawowe dane liczbowe o bibliotekach, przygotowanego do wykorzystania w popularnych arkuszach kalkulacyjnych;
  - wizualizacji przestrzennej pokazującej lokalizację serwerów bibliotek cyfrowych oraz instytucji współtworzących te biblioteki, opartej na technologii Google Maps.
- Wiadomości z bibliotek cyfrowych – zawiera zagregowane wiadomości publikowane na stronach wszystkich przyłączonych bibliotek cyfrowych. Wiadomości te są pobierane z bibliotek poprzez kanały RSS i udostępniane użytkowników w postaci zbiorczej strony z wiadomościami oraz zbiorczego kanału RSS.
- Raporty i statystyki – zawierają informacje na temat łącznej liczby obiektów dostarczanych przez poszczególne źródła danych, rozkładu wartości wybranych elementów schematu metadanych w danych agregowanych w FBC (np. rozkład typu czy formatu obiektów) oraz zmiany liczby obiektów dostępnych i planowanych w czasie. Ponadto dostępny jest raport zawierający wyniki automatycznej analizy bazy FBC pod kątem występowania w zbiorach różnych bibliotek cyfrowych tych samych obiektów (tzw. duplikatów) oraz raport zawierający zestawienie wszystkich obiektów znajdujących się w planach digitalizacji pobieranych automatycznie z bibliotek cyfrowych. Funkcja wykrywania duplikatów jest również udostępniana w postaci publicznego interfejsu programistycznego (API), pozwalającego na sprawdzanie występowania duplikatów już na etapie opracowywania planów digitalizacji w poszczególnych instytucjach (Lewandowska, 2008), zapewniając w ten sposób w podstawowym zakresie zautomatyzowaną koordynację digitalizacji prowadzonej w rozproszonych ośrodkach. Funkcja tego typu jest dostępna w oprogramowaniu dLibra.
- Materiały edukacyjne – najistotniejszym elementem materiałów edukacyjnych dot. bibliotek cyfrowych są dwa kursy e-learningowe opracowane przez PCSS, dotyczące tworzenia bibliotek cyfrowych przez małe instytucje pamięci oraz współpracy z Europeaną. Kursy zostały pierwotnie opracowane przez PCSS w roku 2010 w języku angielskim (w ramach projektu Access IT współfinansowanego z programu Culture) i wdrożone w Serbii, Grecji i Turcji, gdzie z po-

wodzeniem skorzystało z nich ponad 1 000 osób (Dudczak, 2010). Tłumaczenie na język polski i adaptacja kursów do polskich realiów wykonano przy współpracy z Międzynarodowym Centrum Zarządzania Informacją w Toruniu. Kursy te od lipca 2011 r. są dostępne bezpłatnie pod adresem http://fbc.pionier.net.pl/elearning i składają się łącznie z 39 modułów. Ponadto w ramach materiałów edukacyjnych przy FBC utrzymywany jest również spis szkoleń, konferencji i publikacji związanych z polskimi bibliotekami cyfrowymi.

Oferowanie użytkownikom tak szerokiego wachlarza możliwości oraz stała współpraca z kilkudziesięcioma bibliotekami cyfrowymi wymagają systematycznego i precyzyjnego działania, zarówno na poziomie organizacyjnym, jak i technicznym. Te dwa aspekty zostaną omówione w następnych podpunktach.

6.1. Działanie na poziomie organizacyjnym

Pierwszym krokiem dla instytucji, które chcą rozpocząć współpracę z FBC powinno być zarejestrowanie własnej biblioteki cyfrowej w bazie bibliotek cyfrowych FBC. Jeżeli biblioteka cyfrowa jest publicznie dostępna w sieci, to po wprowadzeniu do bazy adresu WWW strony głównej biblioteki oraz adresu interfejsu OAI-PMH administrator FBC weryfikuje merytoryczną i techniczną zgodność zgłoszonego serwisu z wymaganiami stawianymi przez FBC. Wymagania te są szczegółowo omówione na stronie dokumentacji FBC zatytułowanej *Informacje dla prowadzących repozytoria* [http://fbc.pionier.net.pl/owoc/about?id=repo-info]. Po pozytywnej weryfikacji wymagań następuje pierwsze pobranie danych przez FBC. Od tego momentu FBC każdej nocy pobiera dane na temat zmian (nowych/dodanych/usuniętych rekordów), jakie zaszły w danej bibliotece cyfrowej w ciągu minionej doby. Proces pobierania i przetwarzania danych jest omówiony bardziej szczegółowo w następnym podpunkcie.

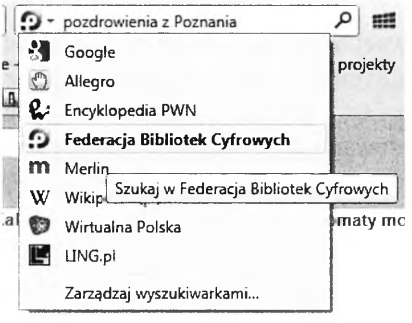
6.2. Działanie na poziomie technicznym

Realizowany okresowo proces pobierania i przetwarzania metadanych przez FBC każdorazowo wygląda w ten sam sposób. W pierwszej kolejności metadane o dostępnych i planowanych publikacjach pobierane są przy pomocy protokołu OAI-PMH. Obecnie pobierane są wyłącznie metadane w formacie Dublin Core Metadata Element Set (DCMES), gdyż jak wspomniano wcześniej jest to schemat występujący w każdym poprawnie zaimplementowanym interfejsie OAI-PMH, a więc udostępniany przez każdą bibliotekę cyfrową współpracującą z FBC. Każdy z pobranych rekordów jest sprawdzany pod kątem poprawności. Obecnie wymagane jest, aby każdy rekord zawierał tytuł i określenie typu obiektu oraz odnośnik do jego wersji on-line (w przypadku obiektów z planów digitalizacji wymagany jest odnośnik do strony informacyjnej o obiekcie). Zaakceptowane rekordy są składowane w bazie danych FBC i indeksowane na potrzeby przeszukiwania tej bazy.

Dodatkowo dla każdego z nowych bądź zmodyfikowanych rekordów, przy pomocy protokołu HTTP, pobierana jest miniatura graficzna reprezentująca opisywany

obiekt. Jeżeli miniatura taka nie jest wskazana bezpośrednio w rekordzie metadanych, to w przypadku publikacji w formacie DjVu, PDF oraz popularnych formatach graficznych (JPG, BMP, PNG) FBC próbuje wygenerować miniaturę automatycznie na podstawie treści obiektu. Co noc pobierane są również informacje o usuniętych rekordach. Rekordy te są oznaczane jako usunięte również w bazie danych FBC oraz usuwane są z indeksów wyszukiwawczych.

Baza danych oraz indeks wyszukiwawczy FBC są podstawą do realizacji opisanych wcześniej funkcji przeznaczonych dla użytkowników końcowych. Poza tym dane te są udostępniane przy pomocy otwartych protokołów innym usługom sieciowym. Dwa podstawowe sposoby takiego dostępu do danych to protokoły OpenSearch [http://www.opensearch.org/] oraz OAI-PMH. Pierwszy z tych protokołów jest wykorzystywany między innymi do integracji wyszukiwania FBC z przeglądarkami WWW (por. rysunek 4) oraz z Katalogiem Rozproszonym Bibliotek Polskich (KaRo) [http://karo.umk.pl/]. Protokół OpenSearch jest w KaRo wykorzystywany równolegle z protokołem Z39.50, który służy do przeszukiwania bibliotecznych systemów katalogowych.



Rys. 4. Interfejs OpenSearch FBC jest umieszczony w predefiniowanym zestawie wyszukiwarek w polu wyszukiwania w prawym górnym narożniku okna przeglądarki Firefox.

Dostęp do danych FBC przy pomocy protokołu OAI-PMH jest z kolei wykorzystywany przez międzynarodowe agregatory metadanych, omówione szerzej w następnym punkcie, które okresowo pobierają z FBC całość lub pewien podzbiór danych. Na potrzeby takiej współpracy interfejs OAI-PMH zaimplementowany w FBC został wzbogacony o szereg zaawansowanych możliwości przetwarzania i udostępniania danych składowanych w bazie FBC. Należą do nich:

- wybór źródeł danych – ze względu na wymagania zarówno dostawców jak i odbiorców danych FBC, konieczne jest precyzyjne określenie, z jakich źródeł danych powinien korzystać dany serwis pobierający dane z FBC (dla przykładu Biblioteka Narodowa przekazuje dane do portalu Europeana poprzez serwis The European Library i nie zezwala na przekazywanie tych samych danych do Europeany poprzez FBC);
- filtrowanie rekordów metadanych – wymagania niektórych dostawców danych wymuszają filtrowanie rekordów metadanych na większym poziomie szczegó-

łowości niż tylko określenie źródła danych (np. DART-Europe oczekuje wyłącznie informacji na temat publicznie dostępnych prac doktorskich) – w takiej sytuacji do określania zakresu przesyłanych danych wykorzystywane są specjalnie skonstruowane zapytania wyszukiwawcze (Mazurek, 2008);

- mapowanie danych pomiędzy schematami – niektóre serwisy agregujące oczekują, że dane udostępniane przez FBC będą zgodne z ich wewnętrznym schematem danych, w związku z tym w razie potrzeby FBC musi być w stanie przetworzyć metadane ze schematu DCMES na schemat odbiorcy danych (na przykład mapowanie ze schematu DCMES do Europeana Semantic Elements);
- standaryzacja danych – niektóre serwisy agregujące oczekuję, że dane udostępniane przez FBC będą zgodne z ich wytycznymi odnośnie do kodowania specyficznych elementów danych, np. wartości pola Język DCMES mogą być kodowane w standardzie ISO 639-2 – w takiej sytuacji np. jedna wartość „polski, niemiecki" jest w trakcie udostępniania przekształcana w dwie odrębne wartości: „pol" oraz „ger";
- wzbogacanie danych – schematy metadanych wymagane przez niektórych dostawców danych mogą być bogatsze niż schemat metadanych FBC, w związku z tym tam gdzie jest to potrzebne i możliwe, dane z bazy FBC służą do automatycznego wypełnienia wartości dodatkowych pól schematu metadanych (np. na podstawie wartości pola „Typ" ze schematu DCMES określane jest czy dany obiekt należy do jednej z czterech klas obiektów wykorzystywanych przez Europeanę: obraz, tekst, audio, wideo);

Taka skala elastyczności interfejsu OAI-PMH Federacji pozwala na współpracę z wieloma różnymi serwisami zewnętrznymi. Jest to sytuacja wygodna zarówno dla tych serwisów, jak i dla dostawców danych. Zewnętrzne serwisy mają tylko jeden punkt kontaktowy, z którego mogą pobrać komplet danych z polskich bibliotek cyfrowych, a dostawcy danych skupiają się na wymaganiach stawianych przez FBC i zazwyczaj nie muszą się dostosowywać do wymagań poszczególnych usług zewnętrznych zainteresowanych publikowanymi przez tych dostawców danymi. W następnym punkcie przedstawiono krótki przegląd międzynarodowych serwisów agregujących informacje również z FBC.

7. Współpraca międzynarodowa

Federacja Bibliotek Cyfrowych, będąc bazą metadanych opisujących obiekty cyfrowe dostępne w polskich bibliotekach cyfrowych, jest naturalnym partnerem do współpracy dla wszystkich międzynarodowych agregatorów – usług, które do realizacji swojego celu tego typu danych potrzebują. Podstawową korzyścią płynącą z takiej współpracy jest znaczne zmniejszenie komunikacyjnej złożoności przedsięwzięcia agregacji metadanych na dużą skalę – zarówno na poziomie komunikacji międzyludzkiej, jak i komunikacji między maszynami.

Po nawiązaniu kontaktu pomiędzy obsługą takiego międzynarodowego agregatora i obsługą FBC następuje zazwyczaj wstępne zdefiniowanie zakresu obiektów, których metadanymi zainteresowany jest taki agregator oraz cel pobierania danych. Następnie określane są kryteria techniczne na poziomie transferu danych oraz kryteria odnośnie do schematu czy stopnia standaryzacji danych. Jeżeli FBC jest w stanie spełnić wszystkie wymagania samodzielnie, to przyłączone do FBC biblioteki są tylko informowane o nawiązaniu nowej współpracy. Jeżeli wymagane jest jakieś wsparcie ze strony dostawców danych FBC, to administrator FBC zajmuje się koordynacją tych działań w skali kraju i informuje partnera zagranicznego, gdy niezbędne prace są już wykonane.

Obecnie Federacja Bibliotek Cyfrowych współpracuje z następującymi partnerami zagranicznymi:

- Europeana [http://europeana.eu] – przekazywanie danych na temat wszystkich publicznie dostępnych obiektów od 11 grudnia 2009 r. – wyznaczonym przez Komisję Europejską celem Europeany jest ułatwienie dostępu do dziedzictwa kulturowego i naukowego Europy w Internecie. Ma ona prezentować bogactwo dziedzictwa europejskiego, łączyć wielokulturowość i wielojęzyczność z nowoczesną technologią i współczesnymi modelami biznesowymi. Europeana jest jedną z flagowych inicjatyw Komisji Europejskiej finansowaną poprzez szereg międzynarodowych projektów. FBC jest największym polskim dostawcą metadanych dla Europeany[7].
- DART-Europe [http://www.dart-europe.eu] – przekazywanie danych na temat publicznie dostępnych elektronicznych rozpraw doktorskich i habilitacyjnych (ang. Electronic Theses and Dissertations, ETD) od 12 stycznia 2010 r. – celem projektu jest ułatwienie dostępu do prac dyplomowych i dysertacji powstających w Europie. Cel ten realizowany jest poprzez agregowanie informacji o takich pracach dostępnych online, a następnie udostępnianie tych informacji w portalu umożliwiającym ich przeszukiwanie i przeglądanie. Projekt ten jest wspierany przez LIBER. FBC jest agregatorem metadanych dla DART-Europe, a Poznańskie Centrum Superkomputerowo-Sieciowe jest oficjalnym partnerem DART-Europe.
- ViFaOst [http://www.vifaost.de] – przekazywanie danych na temat wszystkich obiektów od 14 maja 2010 r. – ViFaOst jest interdyscyplinarnym portalem, który oferuje społeczności akademickiej szeroki zakres informacji na temat Europy Wschodniej, umożliwiając dostęp do specjalistycznych baz informacji naukowej na temat historii, języka, literatury, polityki i kultury wschodnich, środkowo-wschodnich i południowo-wschodnich krajów i regionów europejskich. Dostęp do treści i usług jest bezpłatny.

[7] Listę obiektów widocznych w Europeanie na podstawie danych przekazanych poprzez FBC można zobaczyć pod adresem http://www.europeana.eu/portal/brief-doc.html?query=*%3A*&qf=PROVIDER:Federacja+Bibliotek+Cyfrowych

Ponadto wstępnie nawiązana została współpraca pomiędzy FBC i serwisem WorldCat, dzięki której zbiory bibliotek cyfrowych przyłączonych do FBC są również dostępne w tym serwisie pod adresem http://oaister.worldcat.org/.

8. Podsumowanie i kierunki dalszych prac

Celem niniejszego rozdziału było zaprezentowanie na przykładzie funkcjonującej w Polsce Federacji Bibliotek Cyfrowych koncepcji agregacji metadanych, jako ważnej techniki (a w szerszym kontekście również taktyki) umożliwiającej szerokie propagowanie informacji na temat zbiorów rozproszonych bibliotek cyfrowych i wykorzystywanie tych informacji przy opracowywaniu nowych usług informacji naukowej. Przedstawiono podstawy technologiczne agregacji metadanych oraz główne wyzwania związane z praktyczną implementacją tego typu usług. Na tym tle, omówiona została Federacja Bibliotek Cyfrowych – usługa sieciowa, która gromadzi informacje o kilkuset tysiącach obiektów z kilkudziesięciu różnych bibliotek cyfrowych.

Dalszy rozwój Federacji Bibliotek Cyfrowych związany będzie z podstawową funkcją tego serwisu, czyli agregacją metadanych. Zasygnalizowane na przykładzie porównywania opisów produktów, problemy z interoperacyjnością na poziomie struktury metadanych są obecnie w FBC rozwiązane w bardzo podstawowy sposób – poprzez wymuszenie zgodności schematu źródła metadanych z DCMES. W 2010 r. na podstawie analizy schematów metadanych funkcjonujących w polskich bibliotekach cyfrowych [Potęga 2010] oraz wytycznych związanych z najpopularniejszymi formatami metadanych wykorzystywanymi w międzynarodowych agregatorach metadanych opracowana została specyfikacja schematu metadanych PLMET [Werla 2011], który ma zostać wprowadzony jako wewnętrzny schemat danych w FBC i będzie zalecany do wdrożenia w polskich bibliotekach cyfrowych z FBC współpracujących.

Ponadto prace badawcze realizowane przez PCSS w ramach projektu SYNAT[8] umożliwią w niedalekiej przyszłości pobieranie metadanych w schematach możliwie najbogatszych semantycznie, a nie wyłącznie w postaci DCMES. Takie bogate metadane będą następnie przekształcane do postaci ontologii i przetwarzane przy pomocy technologii sieci semantycznej, co pozwoli na skonstruowanie na ich podstawie bazy wiedzy i udostępnienie w efekcie nowego zaawansowanego narzędzia informacji naukowej pozwalającego użytkownikom eksplorować dane pobrane i zintegrowane z szeregu rozproszonych i heterogenicznych systemów informatycznych.

[8] Projekt SYNAT [http://www.synat.pl/], finansowany przez Narodowe Centrum Badań i Rozwoju (nr umowy: SP/I/1/77065/10), ma na celu realizację zadania badawczego pod tytułem „Utworzenie uniwersalnej, otwartej, repozytoryjnej platformy hostingowej i komunikacyjnej dla sieciowych zasobów wiedzy dla nauki, edukacji i otwartego społeczeństwa wiedzy", które jest realizowane w ramach Programu Badań Strategicznych „Interdyscyplinarny system interaktywnej informacji naukowej i naukowo technicznej". Koordynatorem projektu jest Uniwersytet Warszawski – ICM, a Poznańskie Centrum Superkomputerowo-Sieciowe jest jednym z głównych partnerów projektu.

Bibliografia

1. Burnhill, Peter; Macdonald, Stuart (2011). *'aggregation as a tactic' – to support discovery* [dok. elektr.] http://www.slideshare.net/edinadocumentationofficer/aggregation-as-tactic [odczyt: 12.11.2011].
2. DCMI – Dublin Core Metadata Initiative (2010). *Dublin Core Metadata Element Set, Version 1.1.* [dok. elektr.] http://dublincore.org/documents/dces/ [odczyt: 12.11.2011].
3. Dudczak, Adam; Werla, Marcin (2010). *Projekt Access IT – intensyfikacja obiegu kultury poprzez wymianę umiejętności z zakresu technologii informacyjnych.* Prezentacja z IX Konferencji z cyklu: Automatyzacja bibliotek publicznych wygłoszona w Warszawie w dniu 26.11.2010 r. [dok. elektr.] http://www.koszykowa.pl/bpw/images/stories/imprezy/imprezy2010/20101125_konferencja_automatyzacji/dudczak_adam-werla-marcin.pdf [odczyt: 12.11.2011].
4. Dudczak, Adam; Heliński, Marcin; Mazurek, Cezary; Parkoła, Tomasz; Werla, Marcin (2007). *dLibra – platforma do budowy bibliotek cyfrowych.* Poznań: Wydawnictwo Politechniki Poznańskiej, s. 458-467 [II Krajowa Konferencja Naukowa "Technologie przetwarzania danych", Poznań, 24-26 wrzesień, 2007].
5. Dudczak, Adam; Lewandowska, Agnieszka; Werla, Marcin (2011). *Europeana API – Example of use in Polish digital libraries.* eChallenges e-2011 Conference Proceedings, Paul Cunningham and Miriam Cunningham (Eds), IIMC International Information Management Corporation Ltd 2011, [dok. elektr] http://lib.psnc.pl/publication/408/content [odczyt: 12.11.2011].
6. JISC (2010). *One to many, many to one: the resource discovery taskforce vision* [dok. elektr.] http://ie-repository.jisc.ac.uk/475/1/JISC%26RLUK_VISION_FINAL.pdf [odczyt: 12.11.2011].
7. Lagoze, Carl; Van de Sompel, Herbert (2006). *Implementation guidelines for the Open Archives Initiative Protocol for Metadata Harvesting – specification and XML Schema for the OAI Identifier Format* [dok. elektr.] http://www.openarchives.org/OAI/2.0/guidelines-oai-identifier.htm [odczyt: 12.11.2011].
8. Lagoze, Carl; Van de Sompel, Herbert (2008). *The Open Archives Initiative Protocol for Metadata Harvesting Version 2.0* [dok. elektr.] http://www.openarchives.org/OAI/openarchivesprotocol.html [odczyt: 12.11.2011].
9. Lewandowska, Agnieszka; Mazurek, Cezary; Stroiński, Maciej; Werla, Marcin; Węglarz, Jan (2008). *Federacja Bibliotek Cyfrowych w sieci PIONIER*, Gdańsk: Centrum Informatyczne TASK, s. 216-221 [INFOBAZY 2008 Systemy-Aplikacje-Usługi 15-17 października 2008r., Sopot].
10. Lewandowska, Agnieszka; Mazurek, Cezary; Werla, Marcin (2007). *Federacja Bibliotek Cyfrowych w sieci PIONIER – dostęp do otwartych bibliotek cyfrowych i repozytoriów.* Warszawa: SBP KWE,, [dok. elektr] http://eprints.rclis.org/handle/10760/10967 [IV Ogólnopolska Konferencja EBIB Internet w bibliotekach – Open Access, Toruń, 7-8 grudzień, 2007].
11. Mazurek, Cezary ; Werla, Marcin (2008). *Extending OAI-PMH Protocol with Dynamic Sets Definitions Using CQL Language.* IADIS International Conference Information Systems 2008, Algarve, Portugal, 9-11 kwietnia, 2008 [dok. elektr.] http://dl.psnc.pl/biblioteka/publication/176/content [odczyt: 12.11.2011].
12. Mazurek, Cezary; Stroiński, Maciej; Werla, Marcin (2005). *Wdrażanie regionalnych bibliotek cyfrowych w sieci PIONIER w oparciu o środowisko dLibra.* Gdańsk: Centrum Infor-

matyczne TASK, s. 58-64 [INFOBAZY 2005 – Bazy Danych dla Nauki, Gdańsk, 25-27 wrzesień, 2005].

13. Nilsson, Mikael; Baker, Thomas; Johnston, Pete (2009). *Interoperability Levels for Dublin Core Metadata* [dok. elektr.] http://dublincore.org/documents/interoperability-levels/ [odczyt: 12.11.2011].
14. Potęga, Joanna; Wróbel, Agnieszka (2010). *The Dublin Core Metadata Element Set, Ver. 1.1 a potrzeby i oczekiwania bibliotekarzy cyfrowych – analiza przypadków.* [W:] Mazurek, Cezary; Stroiński, Maciej; Węglarz, Jan red. (2010). *Polskie Biblioteki Cyfrowe 2009 : materiały z konferencji zorganizowanej dnia 9 grudnia 2009 r. przez Bibliotekę Kórnicką PAN, Poznańską Fundację Bibliotek Naukowych, Poznańskie Centrum Superkomputerowo-Sieciowe.* Poznań: Ośrodek Wydawnictw Naukowych, s. 71-78.
15. *Price comparison service* (2012). [dok. elektr.] *Wikipedia* http://en.wikipedia.org/wiki/Price_comparison_service#Technology [odczyt: 23.01.2012].
16. Werla, Marcin (2005). *Rozproszone biblioteki cyfrowe. Prezentacja wygłoszona w czasie drugich Warsztatów „Biblioteki cyfrowe" w listopadzie 2005 r. w Poznaniu* [dok. elektr.] http://dl.psnc.pl/biblioteka/publication/90/content [odczyt: 12.11.2011].
17. Werla, Marcin (2011). *Dokumentacja schematu metadanych PLMET* [dok. elektr.] http://dl.psnc.pl/community/display/FBCMETGUIDE [odczyt: 12.11.2011].

Wielkopolska Biblioteka Cyfrowa – analiza przypadku

Ewa Wichlińska

W pracy zaprezentowano historię powstawania Wielkopolskiej Biblioteki Cyfrowej – pierwszej ***regionalnej*** biblioteki cyfrowej w kraju. Przedstawiono początki konstruowania projektu: zasady funkcjonowania, instytucje współuczestniczące. Omówiono zasoby WBC: kolekcje główne, kolekcje wydzielone, wystawy okolicznościowe. Zaprezentowano również statystyki WBC w formie graficznej.

*

W styczniu 1999 r. w Dziale Usług Sieciowych Poznańskiego Centrum Superkomputerowo-Sieciowego PCSS[1], rozpoczęły się prace nad systemem do budowy bibliotek cyfrowych dLibra [http://www.pionier.net.pl/magazine/pl].

Pierwszym etapem było powstanie Polskiego Internetu Optycznego – Zaawansowane Aplikacje, Usługi i Technologie dla Społeczeństwa Informacyjnego (program Pionier), w którym zdefiniowano potrzebę budowy systemu sieciowych usług bibliotek cyfrowych, umożliwiających rozwój aplikacji. Jest rok 2000 (wrzesień).

Decyzją rektorów uczelni wyższych miasta Poznania i dyrektorów poznańskich bibliotek akademickich i naukowych, w kwietniu 2001 r. miało miejsce uruchomienie pierwszego projektu utworzenia środowiskowej biblioteki cyfrowej. Projekt koordynuje Poznańska Fundacja Bibliotek Naukowych (PFBN)[2].

[1] Poznańskie Centrum Superkomputerowo-Sieciowe (PCSS) – poznańska instytucja informatyczna afiliowana przy Instytucie Chemii Bioorganicznej PAN. Jest jednym z polskich Centrów Komputerów Dużej Mocy. PCSS zarządza siecią PIONIER (Polski Internet Optyczny), która łączy wszystkie jednostki naukowe (uczelnie, instytuty badawcze) za pomocą sieci miejskich. Jednym z głównych zadań operacyjnych PCSS jest także dostarczanie infrastruktury obliczeniowej oraz zaawansowanych usług związanych z przetwarzaniem danych czy udostępnianiem systemów archiwizacji o dużych pojemnościach. Siedziba PCSS mieści się w Ośrodku Nauki PAN przy ul. Wieniawskiego 17/19 w Poznaniu.

[2] Poznańska Fundacja Bibliotek Naukowych utworzona została w 1996 r. z inicjatywy rektorów państwowych poznańskich szkół wyższych. 6 maja 1996 r. Fundację zarejestrowano w Sądzie Okręgowym Miasta Stołecznego Warszawy. Fundacją zarządza bezpośrednio Zarząd Fundacji, a nadzór sprawuje Rada Fundacji i Kolegium Rektorów Miasta Poznania. Pierwsze lata istnienia Fundacji związane były

Październik 2002 r. to data inauguracji Wielkopolskiej Biblioteki Cyfrowej [http://www.wbc.poznan.pl], powstałej dzięki współpracy Poznańskiej Fundacji Bibliotek Naukowych i Poznańskiego Centrum Superkomputerowo-Sieciowego z bibliotekami w Wielkopolsce. W projekcie uczestniczą biblioteki naukowe i publiczne Poznania, Archiwum Państwowe w Poznaniu oraz niektóre biblioteki regionalne jak np. Biblioteka Publiczna im. Stefana Rowińskiego w Ostrowie Wielkopolskim. WBC udostępniła dotychczas 146 832 obiekty cyfrowe i jak dotychczas jest największą biblioteka cyfrową w kraju.

Rok 2002 to również rok, w którym swoje zbiory na platformie cyfrowej WBC rozpoczyna umieszczać Biblioteka Kórnicka PAN [http://www.bkpan.poznan.pl], która jest biblioteką wiodącą – liczba udostępnionych publikacji: 78 515. Biblioteka Uniwersytecka UAM [http://www.lib.amu.edu.pl] do projektu przystępuje we wrześniu 2004 r., umieszczając dotychczas 51 132 obiekty. Pozostałe biblioteki uczestniczące w projekcie powoli, acz sukcesywnie umieszczają swoje zbiory na platformie[3]. Ważną pozycję w projekcie odgrywa Archiwum Państwowe w Poznaniu udostępniając użytkownikom online 4607 obiektów archiwalnych.

Coraz większą rolę odgrywa Biblioteka Akademii Wychowania Fizycznego w Poznaniu [http://biblioteka.awf.poznan.pl] mogąca się poszczycić najczęściej przeglądanymi skryptami i podręcznikami.

*

Pierwsze warsztaty „Biblioteki Cyfrowe", współorganizowane przez PFBN i PCSS odbyły się w listopadzie 2004 r i zgromadziły ok. 60 uczestników.

Obiekty cyfrowe pochodzące z zasobów polskich bibliotek pojawiają się w światowych wyszukiwarkach (OAI-PMH) w 2005 r.

Poznańskie Centrum Superkomputerowo-Sieciowe w 2007 r. uruchamia Federację Bibliotek Cyfrowych (FBC)[4]. Polskie Biblioteki Cyfrowe do końca listopada 2011 r. udostępniły ***816 658*** obiektów.

WBC podobnie jak większość bibliotek cyfrowych w Polsce działa, wykorzystując system dLibra stworzony i cały czas udoskonalany przez Poznańskie Centrum Superkomputerowo-Sieciowe. Zasoby sprzętowo-programowe stanowią integralną część bibliotek uczestniczących w projekcie. Nadzór nad oprogramowaniem używanym przez WBC sprawuje PCSS.

---

z komputeryzacją poznańskich bibliotek naukowych. Zadaniem Fundacji był wybór i zakup oprogramowania zintegrowanego systemu bibliotecznego, organizacja przetargu na zakup sprzętu, koordynacja procesu wdrażania systemu, organizacja szkoleń dla bibliotekarzy, wreszcie nawiązywanie współpracy z innymi ośrodkami akademickimi w Polsce eksploatującymi podobny system biblioteczny.

[3] Od 1999 r. Fundacja koordynuje w skali ogólnokrajowej działalność konsorcjum bibliotek korzystających z zasobów czasopism elektronicznych udostępnianych w ramach programu EIFL.

[4] Federacja Bibliotek Cyfrowych (FBC) jest kolejnym etapem budowy sieci rozproszonych bibliotek cyfrowych i repozytoriów w Polsce. Nazwa serwisu FBC odzwierciedla jego charakter – serwis ten jest zbiorem zaawansowanych usług sieciowych opartych na zasobach cyfrowych dostępnych w polskich bibliotekach cyfrowych i repozytoriach uruchomionych w sieci PIONIER. Zasoby te współtworzone są przez wiele instytucji naukowych i publicznych, takich jak wyższe uczelnie, biblioteki, archiwa, muzea czy ośrodki badawcze. Serwis FBC utrzymywany jest przez Poznańskie Centrum Superkomputerowo-Sieciowe afiliowane przy Instytucie Chemii Bioorganicznej PAN, a za jego rozwój odpowiada Zespół Bibliotek Cyfrowych PCSS.

Gromadzenie zasobów WBC koordynuje PFBN, jednakże każda z bibliotek wprowadzających swoje zbiory do WBC autonomicznie podejmuje decyzje o kryteriach umieszczania zbiorów na platformie cyfrowej.

Publikacje archiwizowane w WBC należą do domeny publicznej lub udostępniane są w sieci lokalnej. Każdy użytkownik ma prawo swobodnego dostępu do materiałów przechowywanych w WBC:

- kopia dokumentu może być wykorzystywana zgodnie z prawem: wykonywana na własny użytek; wykonywana dla celów edukacyjnych lub badawczych z powołaniem się na źródło,
- kopię należy opatrywać dokładnym adresem bibliograficznym, wykorzystując w tym celu link do zacytowanego obiektu lub identyfikator zasobu WBC: http://www.wbc.poznan.pl/publication/226944.

Katalog komputerowy Horizon Biblioteki Uniwersyteckiej w Poznaniu połączony jest z Wielkopolską Biblioteką Cyfrową. Publikacje mające swój opis w Horizonie, a jednocześnie znajdujące się na platformie cyfrowej, posiadają link http://www.wbc.poznan.pl/publication/47164 umożliwiający przejście do przeglądania pełnych tekstów. Należy pamiętać także o tym, że podstawową zasadą przyświecającą funkcjonującym w Polsce bibliotekom cyfrowym jest otwarty dostęp do publikacji (*open access*).

Wielkopolska Biblioteka Cyfrowa cieszy się od momentu powstania, ogromnym zainteresowaniem, o czym świadczą nie tylko liczby odwiedzin na jej stronach internetowych, ale też liczne prośby, uwagi i zapytania dotyczące wprowadzania konkretnych dzieł w ramach projektu, napływające z całego świata. WBC posiada największy zasób czasopism (ok. 60% obiektów cyfrowych w WBC stanowią czasopisma i gazety to w praktyce przekłada się na ok. 10860 numerów) spośród wszystkich bibliotek cyfrowych w kraju. Zawiera bogaty zbiór podręczników akademickich, także starodruków i rękopisów [zob. aneks].

Na szczególną uwagę zasługują kolekcje, zarówno stałe jak i wydzielone czy wystawy okolicznościowe. Zanim zostaną omówione, chciałabym zwrócić uwagę na kolekcję wydzieloną dotyczącą Wojska Polskiego, a w niej na *Dziennik Personalny* zawierający informację o wszelkich ruchach kadrowych w armii w okresie międzywojennym (awanse, przesunięcia, przeniesienia, mianowania, odkomenderowania, zwolnienia ze służby, zgony degradacje, zezwolenia na zmiany nazwisk). Kolekcja zawiera również interesujący zbiór innych pozycji dotyczących wojskowości – m.in. listy oficerów, roczniki oficerskie, spisy poległych i zaginionych, listy odznaczonych.

Wielkopolska Biblioteka Cyfrowa umożliwia poprzez Internet dostępu do czterech typów podstawowych **kolekcji**:

- ***Materiały dydaktyczne*** (zasób obejmuje skrypty, podręczniki i monografie naukowe wydane lokalnie). Liczba publikacji w kolekcji: 2190.
- ***Materiały regionalne*** (zasób obejmuje dokumenty dotyczące Poznania i Wielkopolski. Obok monografii historycznych, zawiera liczne dokumenty archiwalne (akty lokacyjne, przywileje, dekrety), ulotki reklamowe poznańskich firm, katalogi wystaw, statuty poznańskich stowarzyszeń, ulotki wybor-

cze itd. Najstarsze materiały tu zgromadzone pochodzą z wieku XIII). Liczba publikacji w kolekcji: 102714.

- ***Dziedzictwo kulturowe*** (zasób ten obejmuje najcenniejsze zabytki piśmiennictwa znajdujące się w zbiorach bibliotek poznańskich. Poza tym w jego skład wchodzi również pewna liczba dzieł historycznych i z zakresu literatury pięknej wydanych głównie w wieku XIX (m.in. Bibljoteka Pisarzów Polskich) Liczba publikacji w kolekcji: 52448. Liczba publikacji w kolekcji i we wszystkich podkolekcjach: 52449.
- ***Muzykalia*** (zasób obejmuje głównie nuty ze zbiorów Akademii Muzycznej w Poznaniu). Liczba publikacji w kolekcji: 62.

Kolekcje wydzielone WBC to:

- Albumy z fotografiami i pocztówkami
- Akta Braci Czeskich
- Archiwalia miast wielkopolskich (akta miast, materiały źródłowe do dziejów)
- Czasopisma i gazety (zdygitalizowanych ponad 322 tytułów: ok. 10860 numerów; czasopisma i gazety stanowią ok. 60% obiektów cyfrowych w WBC)
- Bibliografie online
- Jordaniana
- Księgi adresowe i listy nazwisk
- Rozprawy doktorskie i habilitacyjne (nauki medyczne i ekonomiczne)
- Wojsko Polskie
- Powstanie styczniowe w zbiorach Wielkopolskiej Biblioteki Cyfrowej (145 rocznica wybuchu powstania styczniowego)
- Telegramy Kościuszkowskie

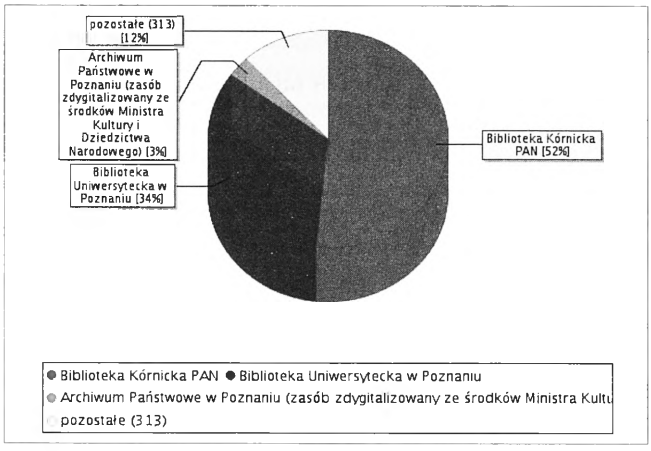
Wystawy Okolicznościowe:

- 85 lat Uniwersytetu Ekonomicznego w Poznaniu (podstawy prawne Uczelni, spisy wykładów oraz składy osobowe, organizacje studenckie, absolwenci)
- Kolekcja pamiętników BUUAM. Druga wojna światowa (1939-1945)
- Powstanie Wielkopolskie 1918-1919
- 90 lat Politechniki Poznańskiej
- 90 lat Uniwersytetu im. Adama Mickiewicza w Poznaniu
- Polonia Maior-fontes.

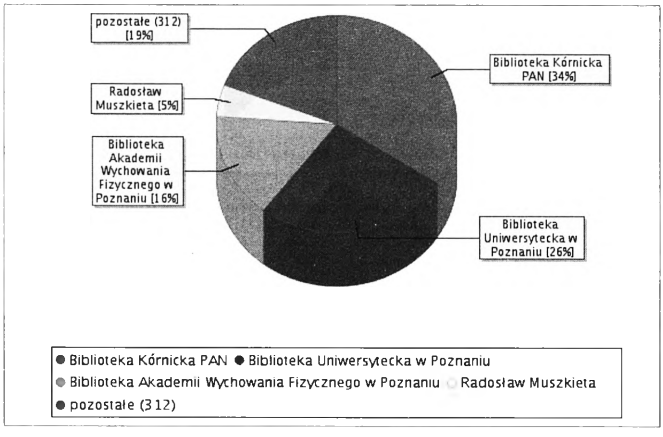
Jak widać Wielkopolska Biblioteka Cyfrowa rozwija się w sposób dynamiczny, powiększając swoje zasoby oraz tworząc coraz to nowsze kolekcje i wystawy. Cieszy się też popularnością wśród internautów. Należy w tym miejscu wyrazić nadzieję, iż wraz z upływem czasu będzie ona sukcesywnie powiększana o nowe publikacje i rozbudowywana o nowe funkcje, aby spełniać coraz bardziej sprecyzowane potrzeby, które tworzą się u czytelników, w trakcie ich interakcji z biblioteką cyfrową.

Aneks

Wybrane statystyki, które w WBC zbierane są od 10.06.2004 r., zaprezentowane zostały w aneksie, aby potwierdzić tezę o dynamicznym rozwoju biblioteki cyfrowej.



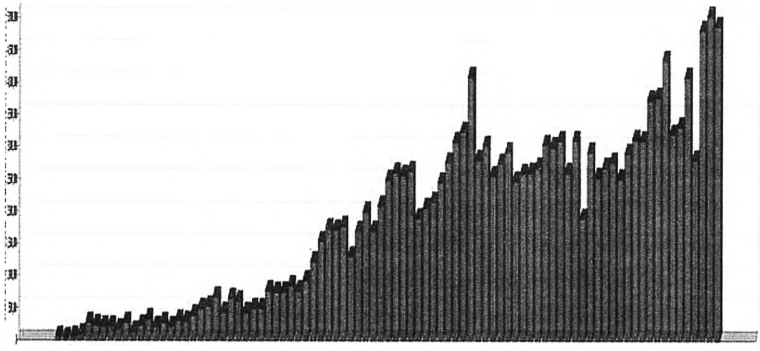
Rys. 1. Udostępnione publikacje: zestawienie wg. liczby udostępnionych publikacji.



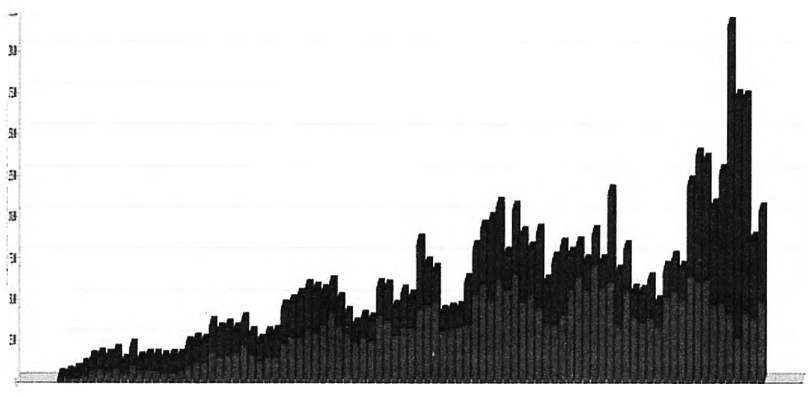
Rys. 2. Wyświetlone publikacje: zestawienie wg. liczby wyświetlonych publikacji.



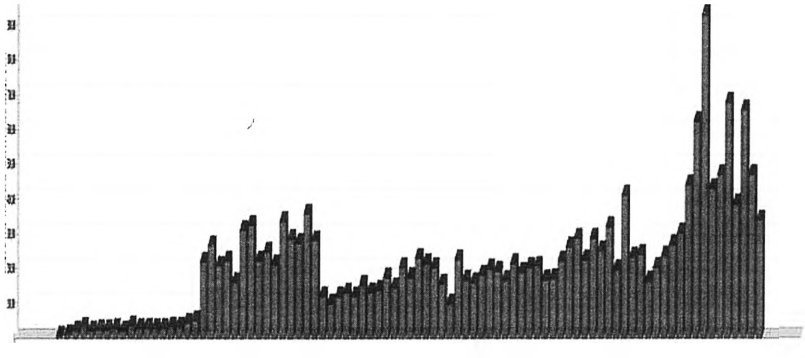
Rys. 3. Liczba publikacji: łączna liczba publikacji w bibliotece (czerwiec 2004 – listopad 2011).



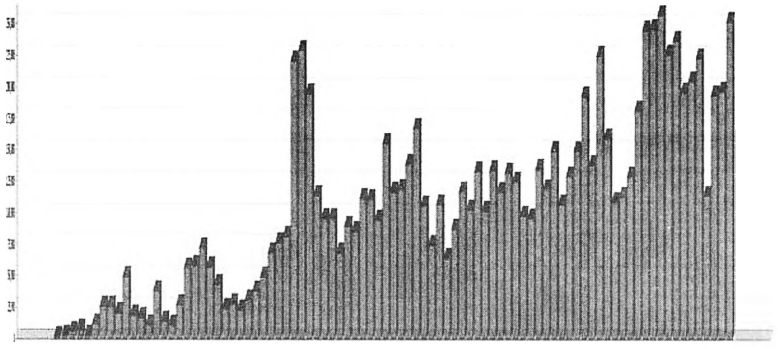
Rys. 4. Wygenerowane strony WWW: łączna liczba wygenerowanych stron WWW w ciągu miesiąca (czerwiec 2004 – październik 2011).



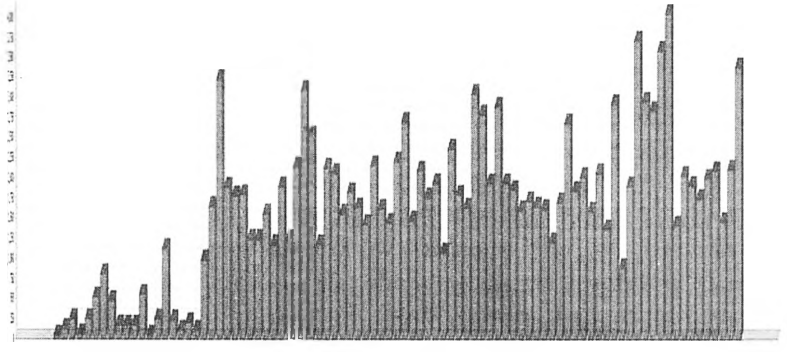
Rys. 5. Wyszukiwanie informacji.



Rys. 6. Odwiedziny: Łączna liczba odwiedzających w ciągu miesiąca (czerwiec 2004 – październik 2011).



Rys. 7. Wyświetlone publikacje: łączna liczba wyświetlonych publikacji w ciągu miesiąca (czerwiec 2004 – październik 2011).



Rys. 8. Nowe publikacje: łączna liczba nowych publikacji w ciągu miesiąca (czerwiec 2004 – październik 2011).

Cyfrowa Biblioteka Narodowa Polona – (cBN Polona)

Joanna Potęga

W pracy przedstawiono Cyfrową Bibliotekę Narodową *Polona* poprzez jej historię, funkcjonalność oraz zasoby. Na początek zaznaczyć należy, iż realizacja zaawansowanych i dużych projektów digitalizacyjnych w bibliotekach zagranicznych (głównie amerykańskich oraz zachodnioeuropejskich) na szeroką skalę rozwinęła się już w początkach lat 90. XX w. W 1994 r. w Internecie pojawił się projekt *American Memory* (porzedzony czteroletnim okresem przygotowawczym), a w 1997 r. Biblioteka Narodowa Francji udostępniła własną bibliotekę cyfrową *Galica* [Kolasa 2010]. W Polsce digitalizacja zbiorów bibliotecznych rozpoczęła się z ponad 10-letnim opóźnieniem. Pod koniec lat 90. XX wieku [Kowalska 2007] zaczęły pojawiać się pojedyncze inicjatywy poszczególnych bibliotek, nierzadko realizowane przy pomocy komercyjnych firm zewnętrznych [Kowalska 2006]. W tym okresie działania te przypominały bardziej nieśmiałe eksperymenty i ograniczone były do digitalizacji wytypowanych (zazwyczaj najcenniejszych) obiektów. Bardzo często zdigitalizowane materiały, nagrywane na płyty CD lub DVD, dostępne były bardzo ograniczonej liczbie użytkowników – w systemie sprzedaży lub prezentowane jedynie w czytelniach bibliotek [Stachowska-Musiał 2002].

W 2002 r. utworzone zostały pierwsze w Polsce biblioteki cyfrowe: Wielkopolska Biblioteka Cyfrowa oraz Polska Biblioteka Internetowa – przedsięwzięcia o odmiennych koncepcjach i sposobie wykonania, ale realizujących te same cele – przedstawiania online zdigitalizowanych książek, czasopism i innych zasobów bibliotecznych (w przypadku WBC również zasobów archiwalnych oraz muzealnych).

Uruchomienie w październiku 2006 r. przez Bibliotekę Narodową własnej platformy do prezentacji zasobów cyfrowych, Cyfrowej Biblioteki Narodowej Polona [http://www.polona.pl], było naturalną konsekwencją prowadzonej od 2003 r. planowej digitalizacji zasobów BN. Przyjęty w 2003 r. program digitalizacji opiera się na celach wspólnych dla większości książnic narodowych – zapewnienie ochrony gromadzo-

nym zbiorom poprzez wykonanie i udostępnienie, zamiast oryginału, cyfrowej kopii. Przede wszystkim skanowaniu poddawane są zbiory będące pod szczególną ochroną ze względu na swoją wartość historyczną czy kulturową (cenne rękopisy, inkunabuły, stare druki, materiały kartograficzne, unikatowe zasoby będące świadectwem tworzenia się państwowości czy dziejowych losów naszego narodu). Digitalizacji poddawane są też materiały zachowane w bardzo złym stanie, które realnie zagrożone są rozpadem (zwłaszcza zbiory wydane na tzw. „kwaśnym" papierze). Kolejną kategorią zasobów kierowanych do digitalizacji są egzemplarze archiwalne, nie posiadające drugiego egzemplarza użytkowego (udostępnianego czytelnikom). W 2010 roku zakres digitalizacji poszerzony został o współczesne publikacje naukowe (książki i czasopisma) w związku z realizacją przez bibliotekę projektu Cyfrowa Biblioteka Publikacji Naukowych ACADEMICA. Głównym celem tego projektu jest utworzenie cyfrowej bazy 170 tysięcy publikacji naukowych oraz ich udostępnienie za pomocą systemu wypożyczeń międzybibliotecznych plików elektronicznych [Cyfrowa Biblioteka Publikacji Naukowych ACADEMICA, http://www.bn.org.pl/programy-i-uslugi/cyfrowa-biblioteka-publikacji-naukowych-academica].

Opracowany program digitalizacji zbiorów Biblioteki Narodowej zakłada:

- zapewnienie autentyczności i integralności tworzonych zasobów cyfrowych,
- bazowanie na własnym sprzęcie do skanowania zbiorów,
- opracowanie standardów przetwarzania i archiwizacji zasobu cyfrowego,
- prowadzenie prac nad znaczną automatyzacją procesów tworzenia zasobu cyfrowego.

Ustalono, że każdy oryginalny obiekt pochodzący ze zbiorów biblioteki i kierowany do digitalizacji, powinien posiadać opis bibliograficzny oraz opis zasobu w komputerowym katalogu BN INNOPAC lub przynajmniej w wewnętrznych, samodzielnych bazach katalogowych MAK.

Przystępując do realizacji programu opracowano standardy techniczne procesu digitalizacji. Przyjęto, że kopie cyfrowe archiwizowane będą w pliku graficznym TIFF, a do celów reprograficznych oraz do udostępniania zasobów cyfrowych wytwarzany będzie niższej jakości plik graficzny JPEG. Do digitalizacji kierowane są przede wszystkim zbiory nie posiadające swojej kopii mikrofilmowej. W przypadku zbiorów specjalnych (rękopisy, stare druki, materiały ikonograficzne oraz kartograficzne) oraz zasobów o postaci kodeksowej – książki, dokumenty życia społecznego – wydanych do 1945 r. rozdzielczość skanowania ustalono od 300 do 600 dpi (rozdzielczość urządzenia do digitalizacji). Ze względu na specyfikę tych zbiorów dopuszcza się możliwość stosowania innych oraz różnych dla jednego obiektu rozdzielczości na zlecenie kierującego do digitalizacji (opiekuna zbiorów) lub na podstawie indywidualnej oceny na każdym etapie ścieżki digitalizacji (w konsultacji z opiekunem obiektu). W zależności od typu obiektu, stanu ich zachowania oraz innych istotnych cech fizycznych materiały skanowane są pojedynczo lub dwie strony na skanie (głównie materiały kodeksowe). Skanowanie wykonywane jest z dokładnym zachowaniem koloru.

W przypadku materiałów o postaci kodeksowej (książki, dokumenty życia społecznego, czasopisma) opublikowanych po 1945 r. rozdzielczość skanowania określo-

na została na 300 dpi (w uzasadnionych przypadkach 600 dpi). Ze względu na to, że jest to digitalizacja z przeznaczeniem do procesu OCR, zasoby te skanowane sa pojedynczo (jedna strona na skan) w skali szarości (okładki, istotne materiały ilustracyjne, a także całe obiekty o walorach artystycznych skanowane są w trybie kolor).

Kopie cyfrowe z mikrofilmów cyfrowe wykonywane są również w rozdzielczości od 300 do 600 dpi (w zależności od jakości taśmy), ale w odcieniach szarości lub trybie czarno-białym.

Wszystkie prace digitalizacyjne wykonywane są w Zakładzie Reprografii i Digitalizacji Zbiorów Biblioteki Narodowej. Utworzona tam w 2004 r. Pracownia Digitalizacji Zbiorów wyposażona jest w nowoczesne skanery do oryginałów: DigiBook 6002 RGB, Zeuschel OS 10000, trzy skanery Zeutschel 12000, jeden skaner Zeutschel 14000 (zakupione w 2008 r.) oraz dwa skanery Canon MS 800 do mikrofilmów [stan na 2011 r.]. Biblioteka Narodowa posiada również dwie pracownie fotografii cyfrowej (w gmachu głównym oraz w Pałacu Rzeczypospolitej) wyposażone każda w profesjonalny cyfrowy aparat fotograficzny Hasselblad H3D-39 z matrycą o wymiarach 48 x 36 mm i rozdzielczości 39 mln pikseli (5412 x 7212). W pracowniach tych fotografowaniu poddawane są głównie obiekty wymagające szczególnego bezpieczeństwa przy digitalizacji (obiekty skarbcowe BN), materiały o dużych formatach (obiekty kartograficzne), jak również o innych nietypowych cechach fizycznych – ciekawych zdobieniach, kolorystyce czy formie, czyli wszystkie te wymagające uzyskania wysokiej jakości odwzorowania cyfrowego.

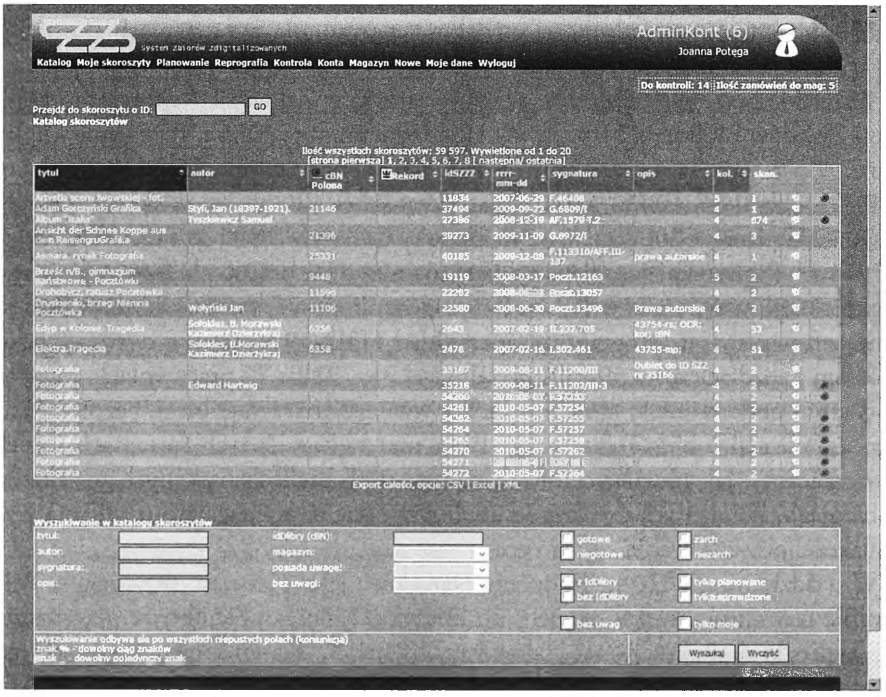
W październiku 2010 r. Biblioteka Narodowa jako pierwsza instytucja w Polsce zakupiła trzy automatyczne skanery do książek (Qidenus). Skanery te przeznaczone są głównie do digitalizacji materiałów współczesnych, skanowanych w ramach realizowanego przez BN programu Cyfrowa Biblioteka Publikacji Naukowych ACADEMICA. Zakup skanerów został sfinansowany ze środków Europejskiego Funduszu Rozwoju Regionalnego (Program Operacyjny Innowacyjna Gospodarka).

Oprócz profesjonalnego sprzętu do cyfryzacji zbiorów analogowych istotnym elementem w procesie digitalizcji jest dbałość o bezpieczeństwo materiałów bibliotecznych. W związku z tym, wszystkie pomieszczenia pracowni i stanowiska pracy zostały odpowiednio przygotowane, m.in. zabezpieczono je przed wpływem czynników zewnętrznych i zainstalowano urządzenia zapewniające stabilne, odpowiednie warunki klimatyczne (utrzymywana jest stała temperatura 16-18°C oraz wilgotność względna na poziomie 40-50%). Wprowadzono również nadzór konserwatorski przy skanowaniu obiektów zabytkowych.

Ważnym elementem całego procesu digitalizacji jest zapewnienie długotrwałej i bezpiecznej archiwizacji wytworzonych materiałów cyfrowych. W Bibliotece Narodowej pliki użytkowe (jpg), składowane są na macierzach dyskowych dedykowanych serwerów zlokalizowanych w serwerowni Zakładu Technologii Informatycznych BN. Z kolei pliki podstawowe (tiff) archiwizowane są odrębnie z wykorzystaniem biblioteki taśmowej.

Dla wspierania prowadzonych w BN procesów skanowania i przetwarzania plików opracowano System Zbiorów Zdigitalizowanych, którego celem jest komputerowe wspomaganie całej ścieżki, jaką przechodzi dokument, od momentu selekcji aż do zarchiwizowania jego cyfrowej kopii. SZZ umożliwia automatyczną kompresję for-

matów (z tiff na jpg) oraz rozdzielczości plików. Innymi ważnymi funkcjonalnociami tego systemu są: możliwość prezentacji cyfrowych zasobów (wykorzystana w cBN Polona) oraz wykonywanie reprodukcji na zamówienie. SZZ utworzony został z wykorzystaniem rozwiązań open source i oparty jest na platformie Java.



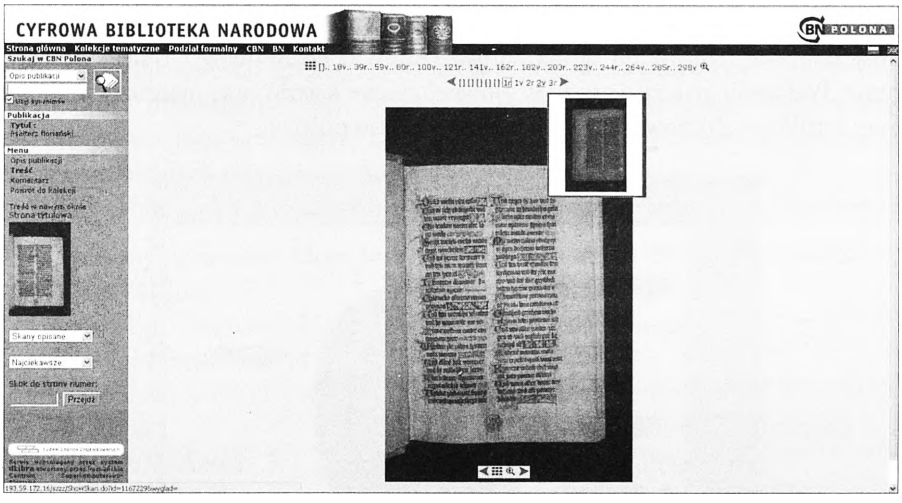
Rys. 1. Ekran Systemu Zbiorów Zdigitalizowanych BN.

Jak już wspomniano, Cyfrowa Biblioteka Narodowa Polona została uruchomiona w październiku 2006 r., a w jej tworzenie jest zaangażowanych wiele komórek biblioteki. Zarówno zakłady sprawujące bezpośrednio opiekę nad zbiorami BN, jak i wspominany wcześniej Zakład Technologii Informatycznych oraz Zakład Reprografii i Digitalizacji Zbiorów. Nad spójnością wszystkich procesów, działań, a przede wszystkim realizacji polityki digitalizacji BN, w początkowych miesiącach funkcjonowania cBN Polona czuwali wyznaczeni bibliotekarze (Biliotekarz Cyfrowy Zbiorów Nowszych oraz Bibliotekarz Cyfrowy Zbiorów Specjalnych), a od stycznia 2007 r. zespół nowoutworzonego Zakładu Cyfrowej Biblioteki Narodowej Polona.

Podstawowym celem cBN Polona jest zapewnienie szerokiego, powszechnego i bezpłatnego dostępu do zasobów cyfrowych, powstających w wyniku digitalizacji zbiorów własnych Biblioteki Narodowej. Polona ma za zadanie prezentować przede wszystkim dziedzictwo kulturowe Polski – najważniejsze dzieła kształtujące i określające polską tożsamość narodową. Istotną rolą tej biblioteki cyfrowej jest również umożliwianie dostępu do zasobów wspomagających czy umożliwiających badania naukowe, działalność kulturalną, oświatową i edukacyjną oraz pracę twórczą czy za-

wodową. Tak szeroko zdefiniowane cele tej biblioteki cyfrowej wskazują, że jest ona przeznaczona dla bardzo dużego kręgu odbiorców. Podobnie, polityka kształtowania zasobu cBN jest bardzo ogólna i uniwersalna. Główne kryteria, według których powstają zbiory Polony to przede wszystkim:

1. zabezpieczanie cennych, szczególnie chronionych oraz będących w złym stanie zachowania zbiorów Biblioteki Narodowej,
2. udostępnianie kopii egzeplarzy archiwalnych (nieposiadających drugiego egzemplarza użytkowego),
3. realizacja założeń tematycznych poszczególnych kolekcji,
4. udostepnianie materiałów katalogowanych w ramach prowadzonej w BN retrokonwersji katalogów kartkowych,
5. realizacja postulatów użytkowników cBN Polona.



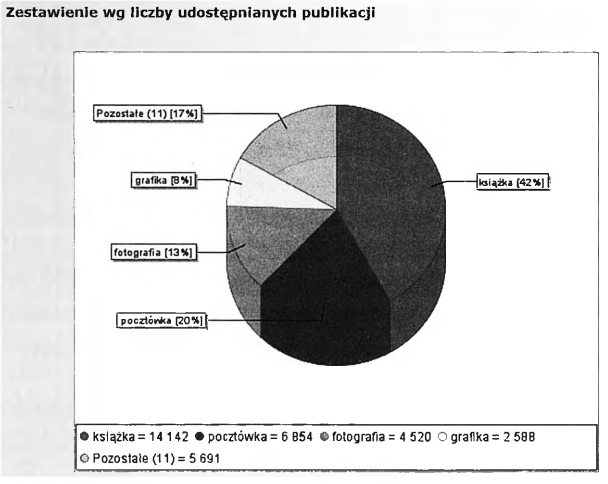
Rys. 2. Prezentacja obiektu w cBN Polona http://polona.pl/dlibra/doccontent2?id=196.

Idea uniwersalności zaważyła również na wyborze narzędzi do prezentacji zbiorów cyfrowych w Internecie. W związku z tym zdecydowano się na połączenie platformy dLibra, opracowanej przez Poznańskie Centrum Superkomputerowo-Sieciowe z działającym w BN Systemem Zbiorów Zdigitalizowanych. Oprogramowanie dLibra wspomaga zarządzanie biblioteką cyfrową oraz umożliwia współpracę z innymi bibliotekami cyfrowymi na płaszczyźnie wzajemnego przeszukiwania zasobów, a także stosowania wspólnych protokołów dostępu do informacji. Z kolei System Zbiorów Zdigitalizowanych BN wykorzystywany jest do prezentacji obiektów cyfrowych w formacie JPEG, czytelnego przez wszystkie systemy komputerowe. SZZ umożliwia tworzenie prostej, intuicyjnej nawigacji po publikacjach, odzwierciedlającej możliwie wiernie strukturę dokumentu oryginalnego (paginację, spisy treści, rozdziały). Użytkownik ma możliwość powiększania obrazu oraz wyświetlenia galerii miniatur wszystkich stron, jak również – podglądu miniatur poszczególnych stron, a także wyboru konkretnego skanu z danego dokumentu (przejścia do strony o wskazanym

numerze). Warto podkreślić, że zastosowany w cBN Polona sposób prezentacji dokumentu pozwala także na nadawanie stronom innej niż standardowa numeracji, wskazywanie cennych opraw, ilustracji czy tablic w tekście. Jest to szczególnie przydatne w przypadku książek zabytkowych.

Aby sprostać oczekiwaniom użytkowników, niezbędne było wprowadzenie innych, dodatkowych formatów prezentacji zasobów cyfrowych. Obecnie sukcesywnie tworzone są alternatywne formaty plików – obok pliku graficznego JPEG, pojawia się plik tekstowy w formacie txt oraz plik PDF z warstwą tekstową. Dzięki temu, każdy może wybrać dogodny sposób oglądania publikacji, a przede wszystkim ma możliwość szybkiego pobrania i zapisania na dysku własnego komputera wybranych dokumentów. Możliwe jest ponadto pełnotekstowe wyszukiwanie w zasobach cBN Polona (dzięki indeksowaniu przez oprogramowanie dLibra plików tekstowych).

Zasoby cBN w listopadzie 2011 r. liczyły 34164 obiektów [stan z dnia 21.11.2011 r.]. Na przestrzeni blisko sześciu lat funkcjonowania Polony roczny wzrost zasobów utrzymywał się na porównywalnym poziomie 6-7 tysięcy obiektów rocznie. Podkreślić należy, że są to w głównej mierze książki oraz materiały ikonograficzne. Struktura zbiorów cBN przedstawia się następująco:



Rys. 3. Obiekty w cBN Polona

1. książka [14142]
2. pocztówka [6854]
3. fotografia [4520]
4. grafika [2588]
5. rysunek [1507]
6. dokument życia społecznego [1321]
7. druk muzyczny [904]
8. stary druk [869]
9. czasopismo [693] 10. rękopis [306]
10. rękopis [306]
11. rękopis muzyczny [54]
12. varia [18]
13. mikrofilm [14]
14. ikonografia [3]
15. dokument elektroniczny [2]

Liczba w nawiasach to łączna liczba udostępnionych w Polonie publikacji zawierających w elemencie DC Typ zasobu podaną wartość atrybutu.

Zbiory cBN Polona porządkowane są według dwóch podstawowych kryteriów: podziału formalnego oraz tematycznego. Przyjęcie takiego rozwiązania podyktowane

było chęcią ułatwienia użytkownikom szybkiej orientacji w zasobach cBN Polona oraz zwrócenia ich uwagi na ciekawe i ważne zasoby Biblioteki Narodowej. Zaznaczyć należy, że każdy obiekt w cBN Polona musi być przyporządkowany do odpowiedniej, jednej kolekcji formalnej. W przypadku kolekcji tematycznych, jeden obiekt może przynależeć do wielu z nich, jak również – nie musi być ujęty w jakiejkolwiek z kolekcji.

Podział formalny odzwierciedla kategoryzację zbiorów powszechnie stosowaną w bibliotekach. Składają się na nią następujące kolekcje:

- Czasopisma
- Grafiki, rysunki, fotografie
- Książki
- Muzykalia
- Rękopisy
- Stare druki
- Ulotki i odezwy

Podział tematyczny z kolei odzwierciedla odmienne podejście do prezentacji zasobów, wynikające z edukacyjno-popularyzatorskiej misji cBN Polona. Podstawowym celem tworzenia kolekcji tematycznych jest przybliżenie użytkownikom historii i rozwoju kultury, literatury polskiej. Dlatego też każdej kolekcji tematycznej towarzyszy tekst wprowadzający, omawiający charakter i tematykę danego zbioru wraz z przedstawieniem ważnych czy ciekawych pozycji.

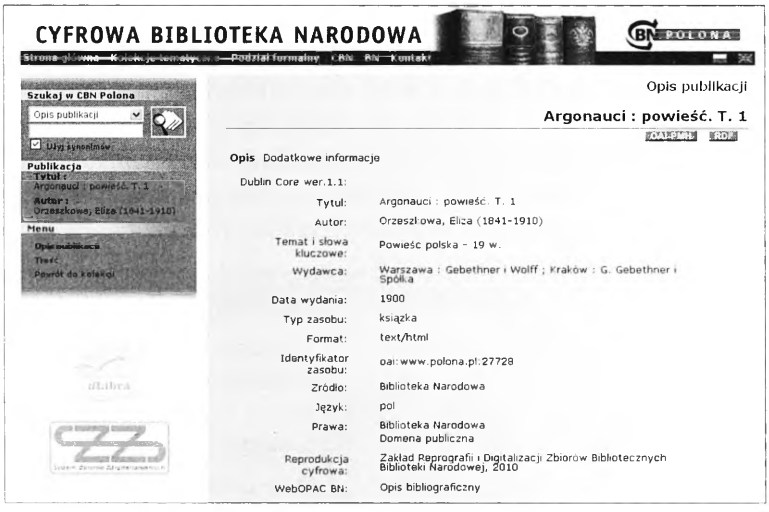
W cBN Polona obecnie (stan na listopad 2011 r.) dostępne są kolekcje tematyczne:

- „Demony" Grzegorza Morycińskiego
- Adam Mickiewicz
- Archiwum Zbigniewa Herberta
- Cyprian Kamil Norwid
- Czasopisma polskie z XIX i I połowy XX w.
- Eliza Orzeszkowa
- Fryderyk Chopin
- Jan Kochanowski
- Juliusz Słowacki
- Karol Lipiński
- Komitet Obrony Robotników
- Kresy Wschodnie
- Kroniki polskie
- Literatura dla dzieci i młodzieży
- Literatura jidysz
- Maria Konopnicka
- Mikołaj Rej
- Powstanie styczniowe
- Skarby Biblioteki Narodowej
- Stanisław Brzozowski
- św. Jadwiga królowa Polski
- Tekla Bądarzewska
- Warszawa w słowie i obrazie
- Wolne lektury
- Wrzesień 1939
- Wydawnictwa konspiracyjne

Istotnym elementem struktury biblioteki cyfrowej jest opis obiektów w niej zamieszczanych – czyli zestaw metadanych opisowych, którymi materiały są opatrzone. W cBN Polona zastosowano standard Dublin Core Metadata Element Set, Ver. 1.1, a przedmiotem opisu jest treść źródła stanowiąca zasób cyfrowy. O wyborze DC zadecydowało przekonanie, iż opis dzieła w cBN służyć ma jedynie celom jego identyfikacji, a także powinien umożliwiać wyszukiwanie obiektu. Opis taki powinien być zwięzły i czytelny. Podstawowym natomiast miejscem, gdzie czytelnik uzyska pełną informację bibliograficzną o źródle zasobu cyfrowego jest katalog komputerowy BN – INNOPAC. Dostęp

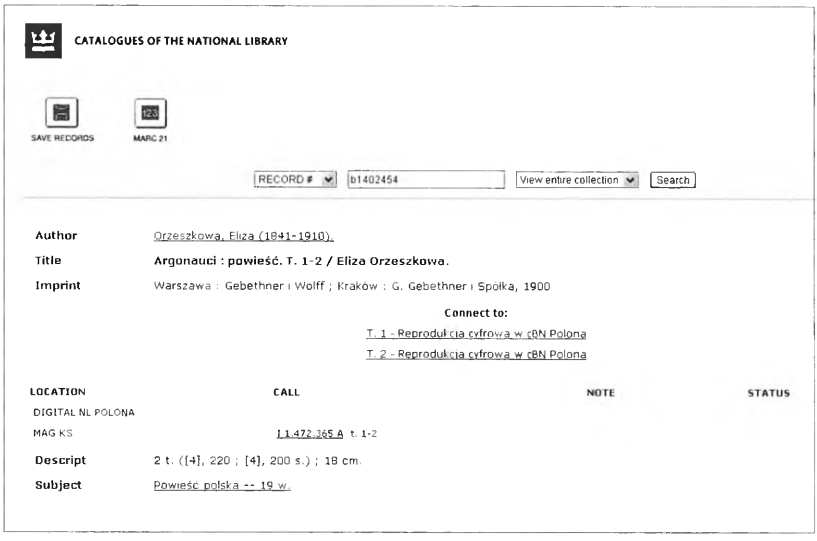
do danych bibliograficznych umożliwiony jest dzięki umieszczeniu powiązań pomiędzy opisem w cBN Polona, a rekordem w katalogowej bazie danych.

Podstawowe 15 elementów (atrybutów) DCMS: *Title*, *Creator*, *Subject*, *Description*, *Publisher*, *Contributor*, *Date*, *Type*, *Format*, *Identifier*, *Source*, *Language*, *Relation*, *Coverage*, *Rights*, zostało w cBN Polona przetłumaczonych następująco: *Tytuł*, *Autor*, *Temat i słowa kluczowe*, *Opis*, *Wydawca*, *Współtwórca*, *Data wydania*, *Typ zasobu*, *Format*, *Identyfikator zasobu*, Źródło, *Język*, *Powiązania*, *Zakres*, *Prawa.* Zestaw ten został poszerzony o dwa elementy: *Reprodukcja cyfrowa,* w którym zamieszczana jest informacja o miejscu i dacie wykonania kopii cyfrowej oraz *WebOPAC BN,* gdzie podawany jest link do opisu bibliograficznego dokumentu źródłowego w katalogu komputerowym BN.



Rys. 4. Opis obiektu cyfrowego w cBN Polona – http://www.polona.pl/dlibra/docmetadata?id=27728&dirids=42.

Ze względu na brak unormowań w zakresie tworzenia metadanych w bibliotekach cyfrowych, jak i zaobserwowane liczne różnice pomiędzy interpretacją schematu Dublin Core w poszczególnych bibliotekach opartych na dLibrze, przygotowano własną implementację DC oraz opracowano zestaw pól formatu MARC21 konwertowanych z INNOPAC-a do opisu obiektu w cBN. Ustalono również, że wszystkie dokumenty źródłowe, których wersja cyfrowa zostanie opublikowana w cBN Polona, będą w katalogu komputerowym INNOPAC BN odpowiednio z nią powiązane poprzez z wykorzystaniem pola 856 formatu MARC21 w rekordzie bibliograficznym (informacja w podpolu *z* treści: „Reprodukcja cyfrowa w cBN Polona"). Dzięki temu użytkownik bazy katalogowej BN ma od razu podaną informację, że dana publikacja jest dostępna online. Takie rozwiązanie umożliwia alternatywny dostęp do zasobów cBN Polona od strony katalogu bibliotecznego.



Rys. 5. Opis bibliograficzny dokumentu źródłowego w katalogu komputerowym wraz z odesłaniem do cBN Polona.

Metadane umożliwiają odnalezienie obiektu w zasobach biblioteki cyfrowej, zarówno poprzez wykorzystanie opcji wyszukiwania prostego (wpisanie słowa do paska wyszukiwarki), jak i wyszukiwania zaawansowanego. Korzystając z zaawansowanego formularza, możliwe jest określenie zakresu wyszukiwania oraz wykorzystanie wyrażeń logicznych (*i, lub*) do łączenia kilku zapytań. Dane umieszczone w elementach *Autor*, *Tytuł* czy *Temat i słowa kluczowe* wykorzystywane do automatycznego generowania indeksów, umożliwiających szybkie przeglądanie zasobów cBN Polona.

Przy ustalaniu formatu opisu obiektów pojawiło się zagadnienie statusu wytwarzanych przez bibliotekę publikacji cyfrowych – czy są one wyłącznie dokumentem prymarnym czy też jest to już nowa jakość – dokument elektroniczny dostępny online? Na potrzeby cBN zadecydowano, że tworzone w postaci cyfrowej i umieszczane w Polonie obiekty nie będą podlegały rejestracji bibliograficznej i bibliotecznej, i tym samym nie będą tworzone dla tych publikacji odrębne opisy bibliograficzne w bazie katalogowej INNOPAC. Wyjątek stanowić będą publikacje wzbogacone o jakąś wartość dodaną, nie będące wierną kopią dokumentów oryginalnych, np. utworzone z wybranych fragmentów dzieła oryginalnego, będące kompilacją kilku dzieł itp. W takich przypadkach, wytworzona publikacja uznana jest za nowy, samoistny dokument elektroniczny i w związku z tym konieczne będzie jego zarejestrowanie i sporządzenie opisu bibliograficznego.

W cBN Polona prezentowane są w zdecydowanej większości dokumenty wolne od zastrzeżeń prawa autorskiego będące w domenie publicznej. Respektując przepisy ustawy o prawie autorskim oraz korzystając z przysługującej bibliotekom licen-

cji ustawowej[1] w cBN udostępnionych jest także wiele współczesnych (chronionych prawem autorskim) i poszukiwanych przez użytkowników publikacji. Dostępne są one wyłącznie w czytelniach Biblioteki Narodowej W przypadku próby skorzystania online z tych utworów, użytkownik cBN Polona otrzymuje informację: „Publikacja chroniona prawem autorskim – reprodukcja cyfrowa dostępna w czytelniach BN". Choć takie rozwiązanie powoduje niedogodność w postaci konieczności osobistego przybycia do siedziby BN, ale jest to jedyna możliwość, by umożliwić czytelnikom dostęp do często jedynych, trudno osiągalnych w tradycyjnej postaci papierowej, egzemplarzy będących w posiadaniu BN.

Należy tutaj podkreślić, że w cBN Polona, dzięki otwartości twórców i ich spadkobierców dostępnych jest online ponad 2000 współczesnych utworów. Zazwyczaj właściciele autorskich praw majątkowych zawierają z Biblioteką Narodową umowy – licencje niewyłączne, tym samym umożliwiając szerokiej publiczności korzystanie z dorobku współczesnych twórców. W przypadku zawarcia takiej umowy, autorskie prawa majątkowe pozostają przy ich właścicielu, natomiast bibliotece udzielana jest zgoda na publikację w cBN objętego taką licencją utworu.

W chwili obecnej użytkownicy cBN Polona mogą zapoznać się m.in. z fotografiami autorstwa Jerzego Benedykta Dorysa[2] i Edwarda Hartwiga[3], dziełami Grzegorza

[1] Art. 28 Ustawy z dnia 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych (Dz. U. z 2006 r. Nr 90, poz. 631 z późn. zm.) zezwala bibliotekom, archiwom i szkołom m.in.: udostępniać zbiory dla celów badawczych lub poznawczych za pośrednictwem końcówek systemu informatycznego (terminali), znajdujących się na terenie jednostek.

[2] **Jerzy Benedykt Dorys** (właściwie Rotenberg) (1901-1990) – wybitny fotografik – portrecista, działacz Polskiego Towarzystwa Fotograficznego przed wojną, po wojnie współzałożyciel współzałożyciel Związku Polskich Artystów Fotografików (1946). Dorys był twórcą unikalnej dokumentacji portretowej wybitnych postaci kultury polskiej. Obok portretów wykonywał zdjęcia architektury oraz pejzaże. Jego przedwojenny cykl „Kazimierz nad Wisłą", który powstawał w latach 1931-1932, uznany został za pierwszy polski reportaż fotograficzny i rezentowany był na wielu światowych wystawach fotograficznych (Nowy Jork 1979, Paryż 1980/81 i 1992). Przekazana Bibliotece Narodowej spuścizna po Benedykcie Jerzym Dorysie (ponad 7000 poz.) jest sukcesywnie digitalizowana i udostępniana w cBN Polona.

[3] **Edward Hartwig** (1909-2003) wybitny fotografik, wywarł duży wpływ na rozwój polskiej fotografii w latach 60. i 70. Był jednym z członków założycieli Związku Polskich Artystów Fotografików, a także członkiem honorowym wielu organizacji fotograficznych na całym świecie. Edward Hartwig przez lata pracował jako fotograf teatralny, a szereg portretów, które wtedy wykonał, zapisało się na trwałe w historii fotografii polskiej. Niezwykła kolekcja Edwarda Hartwiga, w której, obok portretów sławnych polskich aktorów i reżyserów, znajdują się liczne fotografie ze spektaklów teatralnych w Warszawie, Łodzi i Lublinie, dzięki zgodzie spadkobierców, została poddana w 2009 r. digitalizacji i udostępniona w cBN Polona.

Morycińskiego[4], Stanisława Hiszpańskiego[5] oraz utworami znanej autorki literatury dla dzieci – prof. Joanny Papuzińskiej.

Świat technologii, jak również oczekiwania i potrzeby użytkowników Internetu zmieniają się w błyskawicznym tempie. Dostępność treści, już nie tylko w „tradycyjny" komputerowy sposób, ale możliwość sięgania po nie na coraz powszechniejszych urządzeniach mobilnych wyraźnie wskazuje kierunki i tendencje rozwoju bibliotek cyfrowych. Zmiany w sposobach dostępu do informacji naturalnie wpływają na organizowanie treści, zarządzanie nią oraz udostępnianie w bibliotekach cyfrowych. Zagadnienia takie jak: prezentacja treści w wielu formatach, przygotowanie mobilnego, otwartego elektronicznego zasobu oraz organizowanie w przyjazny dla użytkownika sposób dynamicznie rosnących zbiorów cBN to nowe wyzwania dla twórców Polony, ale też i obszary ich intensywnej pracy.

Bibliografia

1. Kolasa, Władysław Marek (2010). *Biblioteki cyfrowe na świecie – powstanie i rozwój* [dok. elektr.] http://eprints.rclis.org/bitstream/10760/16141/1/kolasa_cyrowe.pdf [odczyt: 12.12.2011].
2. Kowalska, Małgorzata (2006). *Digitalizacja zbiorów w bibliotekach polskich – próba oceny doświadczeń krajowych* [dok. elektr.] „EBIB – Elektroniczny Biuletyn Bibliotekarzy" nr 11(81) http://www.ebib.info/2006/81/kowalska.php [odczyt: 12.12.2011].
3. Kowalska, Małgorzata (2007). *Dygitalizacja zbiorów bibliotek polskich*. Warszawa: Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich, 293, [2] s.: il.
4. Stachowska-Musiał, Ewa (2002). *Mikrofilmowanie i digitalizacja zbiorów w Bibliotece Uniwersyteckiej w Warszawie*. „Notes Konserwatorski" 6. s. 182-197.

[4] **Grzegorz Moryciński** (ur. 1936) należy do grona najwybitniejszych twórców współczesnych związanych z nurtem malarstwa metaforycznego. Jego dzieła były wielokrotnie wystawiane w kraju i za granicą, m.in. w Londynie, Rzymie, Brukseli, Melbourne. Zdobył wiele prestiżowych nagród krajowych i międzynarodowych, a jego prace zdobią kolekcje Muzeum Narodowego w Warszawie, Krakowie i Szczecinie, Muzeum Sztuki Współczesnej w Radomiu, Muzeum Archidiecezjalnego w Warszawie, a także galerie narodowe i kolekcje prywatne w Austrii, Niemczech, Francji, Belgii, Holandii, Szwecji, Hiszpanii, USA, Japonii Wenezueli i w Kanadzie.

Cykl „Demony" to pięćdziesiąt prac rysunkowych z lat 2001-2006. Wystawiane w 2006 r. w Bibliotece Narodowej dzieła, dzięki hojności i uprzejmości autora wzbogaciły zbiory Biblioteki, a od 2008 r. w wersji cyfrowej dostępne są dla internatów z całego świata w cBN Polona.

[5] **Stanisław Hiszpański** (1904-1975) – malarz i grafik, od roku 1945 członek rzeczywisty Związku Polskich Artystów Plastyków. Ilustrował wielką literaturę antyczną – Mity Sumeru, Odyseje Homera, Eneidę Wergiliusza, mity słowiańskie, Biblię, fantastykę literacką. Internauci mają okazję zapoznać się w Cyfrowej Bibliotece Narodowej Polona ze zbiorem ilustracji do „Odysei" Homera.

Kujawsko-Pomorska Biblioteka Cyfrowa – studium przypadku

Bożena Bednarek-Michalska

W tekście przedstawiono Kujawsko Pomorską Bibliotekę Cyfrową (KPBC), która była pierwszym w Polsce projektem tak skrupulatnie zaprojektowanym i sfinansowanym z funduszy strukturalnych Unii Europejskiej, z programu: Zintegrowany Program Operacyjny Rozwoju Regionalnego – ZPORR. Przyczyną tego była konieczność związana z poważnym finansowaniem zewnętrznym, a także innowacyjnością zadania, które bez dobrego planu mogłoby ponieść fiasko. Celem projektu było stworzenie biblioteki cyfrowej wspierającej edukację, kulturę i turystykę regionu oraz umożliwienie jak najszybszego dostępu do zasobów wiedzy oraz cennych zabytków kultury piśmienniczej.

Zasób KPBC podzielony został na trzy duże kolekcje zawierające:

- regionalia – cyfrowe wersje książek i zbiorów: ikonograficznych, kartograficznych, muzycznych oraz dokumentów życia społecznego dotyczących Kujaw, Pomorza i Ziemi Dobrzyńskiej, miast regionu (Bydgoszcz, Grudziądz, Włocławek, Toruń, Rypin, Inowrocław i innych), Baltica, Vilniana, Copernicana;
- materiały dydaktyczne – cyfrowe kopie wybranych podręczników akademickich, monografii i artykułów naukowych tworzonych w regionie przez wszystkie uczelnie wyższe;
- dziedzictwo kulturowe – kopie wybranych najcenniejszych i najczęściej wykorzystywanych pozycji: inkunabułów, starodruków, rękopisów, zbiorów ikonograficznych, kartograficznych oraz emigracyjnych pochodzących ze zbiorów bibliotek naukowych i publicznych regionu.

Kujawsko-Pomorska Biblioteka Cyfrowa służy naukowcom, studentom, uczniom i wszystkim mieszkańcom regionu kujawsko-pomorskiego, którzy potrzebują dostępu do wiedzy i informacji w pracy zawodowej czy życiu prywatnym. Ze względu na to, że uczelniom regionu zależy na jak najszybszym udostępnieniu druków własnych

w związku z tym wprowadzono procedury, które umożliwiają sprawne skanowania współczesnych utworów wydanych przez oficyny uczelniane lub samych autorów. Publikacje umieszczone w KPBC mają uregulowany status prawny. KPBC powstaje w oparciu o oprogramowanie *dLibra*, które jest rozwijane od lat przez ośrodek PAN – Poznańskie Centrum Superkomputerowo-Sieciowe.

Projekt i partnerzy

Kujawsko-Pomorska Biblioteka Cyfrowa KPBC jest projektem realizowanym przez Bibliotekę Uniwersytecką w Toruniu (koordynatora) oraz 10 partnerów regionalnych:

- Biblioteka Uniwersytetu Kazimierza Wielkiego w Bydgoszczy
- Biblioteka Collegium Medicum im. Ludwika Rydygiera UMK w Bydgoszczy
- Wojewódzka Biblioteka Publiczna – Książnica Kopernikańska w Toruniu
- Biblioteka Miejska im. Wiktora Kulerskiego w Grudziądzu
- Centrum Kultury Dwór Artusa w Toruniu
- Centrum Sztuki Współczesnej „Znaki Czasu" w Toruniu
- Teatr Baj Pomorski w Toruniu
- Miejsko-Powiatowa Biblioteka Publiczna w Rypinie
- Biblioteka Główna Uniwersytetu Technologiczno-Przyrodniczego w Bydgoszczy
- Biblioteka Państwowej Wyższej Szkoły Zawodowej we Włocławku

Zakłada także współpracę z innymi bibliotekami działającymi w ramach Konsorcjum Bibliotek Naukowych Regionu Kujawsko-Pomorskiego.

Zasady gromadzenia i ochrony

KPBC ma w tej chwili w swoim zasobie ponad 40 000 obiektów cyfrowych różnego typu, od dzieł naukowych po popularnonaukowe czy literackie, od grafik po pocztówki i mapy, rękopisy, stare druki, filmy czy czasopisma. W zasadzie wiele rzeczy daje się ucyfrawiać a system przechowa każdy format. Polityka archiwizacji KPBC opiera się o kilka ważnych założeń. Jednym z najważniejszych jest udostępnienie materiałów, które mogą wspomóc naukę i dydaktykę i ich dobór opiera się o dezyderaty czytelnicze. Jeśli jest prośba studenta czy profesora, zapotrzebowanie na materiał, to się go skanuje i upublicznia, pod warunkiem, że prawo na to zezwala.

Od paru lat BG UMK systematyczne pozyskuje prawa autorskie do dzieł toruńskich profesorów i buduje spersonalizowane kolekcje. Wynikło to z zapotrzebowania naszego środowiska. Kolekcje takie posiadają już prof. Aleksander Jabłoński http://kpbc.umk.pl/dlibra/publication?id=29737, Karol Górski, Krzysztof Mikulski, Andrzej Majdowski, Włodzisław Duch, Adam Jakubowski, w przygotowaniu są kolekcje prof. Wilhelminy Iwanowskiej, Romana Ingardena, Władysława Dziewulskiego, Jerzego Dygdały, Janusza Małłka i innych.

Innym ważnym powodem, dla którego bibliotekarze skanują obiekty jest polityka zabezpieczenia cennych druków, książek czy czasopism. Udostępnianie obiektów

fizycznych powoduje ich degradację, by temu zapobiec przygotowuje się ich postać cyfrową – surogat, który doskonale nadaje się do wielokrotnego wykorzystania nie powodując uszczerbku dla oryginału.

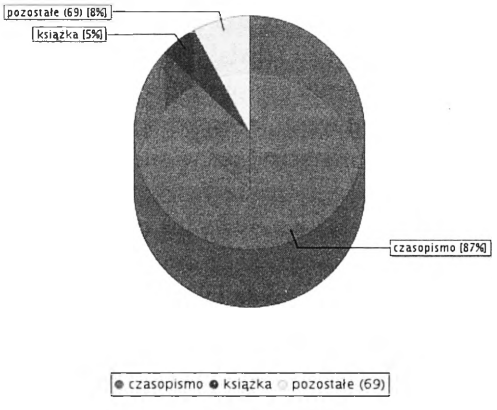
Kolejnym warunkiem szybkiego skanowania jest zniszczenie i zaczytanie materiałów bibliotecznych uniemożliwiające wykorzystanie ich w normalnych warunkach. *Ustawa o prawie autorskim* zezwala takie obiekty skanować, dla uzupełnienia zasobu. Następnym ważnym argumentem za upublicznianiem dzieł jest promocja badań i nauki regionu. Im więcej prac naukowych naszych pracowników znajdzie się w sieci, tym więcej będą mieli czytelników i tym samym cytowań. Praktyki pokazują, że użytkownicy Internetu wchodzą do czasopism i książek wydanych nie tylko na bieżąco, ale i lata temu – dzięki takiej operacji wiele dzieł uzyskuje „drugie życie" i ponownie wchodzi do obiegu czytelniczego, co jest niezwykle ważne dla autora.

Statystyka KPBC

Statystyki w KPBC prowadzone są na bieżąco i dostępne są pod adresem: http://kpbc.umk.pl/dlibra/pubstats. Pozwalają one dokonywać analiz rocznych i wprowadzać ewentualne zmiany. Również na bieżąco utrzymywany jest kontakt w czytelnikami i użytkownikami KPBC zarówno przez e-mail, jak i inne formy komunikacji bezpośredniej czy pośredniej. Dezyderaty czytelników i ich uwagi traktowane są bardzo poważnie.



Rys. 1. Przyrost publikacji w KPBC od 2005 r.



Rys. 2. Typy publikacji w KPBC

Procesy pracy opracowane dla KPBC w Bibliotece Uniwersyteckiej w Toruniu

Procesy pracy związane z upublicznianiem obiektów cyfrowych w KPBC są odmienne dla każdej z instytucji uczestniczącej w projekcie. Zależy to od liczby planowanych do udostępnienia pozycji, wielkości pracowni digitalizacji, liczby pracowników i sprzętu. W niektórych instytucjach przy tej pracy zatrudnia się jedną osobę w innych 16, dlatego nie ma jednolitego modelu organizacyjnego.

Jeśli chodzi o Bibliotekę Uniwersytecką w Toruniu, to zespół pracujący dla KPBC liczy 16 osób i działa wewnątrz BG UMK w strukturze rozproszonej koordynowanej przez szefa projektu. Na zespół składają się pracownicy innych działów, którzy wykonują pracę na rzecz biblioteki cyfrowej w ramach swoich normalnych obowiązków. Planowania szczegółowe procesów pracy, szczególnie wówczas, kiedy mamy do czynienia z dużym zespołem jest trudne. Im bardziej kompetentny zespół, tym mniej szczegółów się opisuje, ponieważ ludzie wiedzą, co i jak należy zrobić. Niemniej jednak problemy występują stale a do dyskusji o organizacji pracy trzeba ciągle powracać, ponieważ wszystko ulega zmianom zarówno koordynacja, jak i skład zespołu.

Proces pracy w KPBC można opisać następująco:

- Selekcja i typowanie dokumentów do cyfryzacji (selekcjonerzy) zgodnie z przyjętymi założeniami;
- Tworzenie listy obiektów przygotowanych do dygitalizacji w postaci tabel (selekcjonerzy);
- Kolejkowanie dokumentów do procesu dygitalizacji i kontrola kolejki – (redaktor);
- Przygotowanie techniczne dokumentu i dostarczenie do pracowni (selekcjonerzy);
- Skanowanie i archiwizacja w formacie TIFF – (technicy);
- Przetwarzanie plików powstałych w procesie skanowania OCR i DjVu (informatycy, technicy);
- Tworzenie opisu bibliograficznego do lokalnego katalogu Horizon w formacie MARC21, konwersja do dLibry w DublinCore (katalogerzy);
- Sprawdzenie statusu prawnego dokumentu (redaktor);
- Publikacja na platformie cyfrowej (redaktor);
- Kontrola metadanych w dLibrze i Horizonie oraz melioracja zasobu i poprawki (główny kataloger);
- Koordynacja wszystkich prac (koordynator).

Etapy digitalizacji dokumentów

1. **Procedury zalecane do stosowanie w Bibliotece Uniwersyteckiej w Toruniu**
   a. Typowanie dokumentów;
   b. Określenie co, jak i kiedy powinno być zdigitalizowane;
   c. Stworzenie tabel zawierająca do 15 pozycji z dokładnymi danymi do obróbki (jakość skanowania, kolor, inne);
   d. Przesyłanie tabel na wspólny dysk dostępny w wew. sieci (umowny format nazwy pliku zawierającego tabele: *grafika_003_doc*).

Wykonanie: bibliotekarze odpowiedzialni za poszczególne kolekcje.

2. **Kolejkowanie dokumentów (tabelek)**
   a. kolejkowanie tabel na dysku X, nadawanie priorytetów, co idzie w jakiej kolejności do skanowania;
   b. informowanie pracowni o kolejce i możliwości realizacji zadania;
   c. stała kontrola procesu digitalizacji i przestrzeganie procedur;
   d. zamawianie dokumentów analogowych do skanowania i dostarczanie do pracowni digitalizacji.

Wykonanie: redaktor KPBC lub selekcjoner.

3. **Skanowanie i archiwizacja**
   a. Skanowanie obiektów wg tabel z kolejki;
   b. przetwarzanie wyniku skanowania programami dołączonymi do skanerów (korekta obrazu, kadrowanie, itp.);
   c. uzupełnianie obiektu o metadane techniczne; format plików, typ nośnika i data nagrania, rozdzielczość, kolor, itp.;
   d. zapisanie pliku pod standardowa nazwą; nadawanie sygnatur nośnikom archiwalnej wersji cyfrowej dokumentu; przesyłanie plików macierzystych na serwer – magazynowanie trwałe plików;
   e. przesłanie drugiej kopii gotowych plików tiff do dalszego przetwarzania w celach udostępnienia na platformie *dlibra*;
   f. oddanie materiałów analogowych do działu opracowania formalnego.

Wykonanie: pracownicy Pracowni Digitalizacji.

4. **Przetwarzanie plików powstałych w procesie skanowania**
   a. Przetwarzanie plików tiff do formatów DjVu, html, PDF, jpg i innych zaplanowanych w procesie typowania i przygotowania;
   b. OCR dla obiektów, które mogą być objęte tym procesem i dadzą dodatni efekt;
   c. Wykonanie miniatury strony tytułowej;
   d. wystawienie plików dla redaktora do upublicznienia – na dysk Y.

Wykonanie: pracownicy Pracowni Digitalizacji, informatycy.

5. **Opracowanie opisu bibliograficznego**
   a. skatalogowanie obiektu w Horizonie w formacie Marc21;
   b. export opisów z formatu Marc21 do DC do dLibry;

c. stała melioracja i sprawdzanie jakości opisów w dLibrze;
d. wprowadzanie korekt do dLibry, przeglądanie indeksów.

Wykonanie: kataloger, bibliotekarze odpowiedzialni za poszczególne kolekcje specjalne.

6. Upublicznienie obiektu na platformie cyfrowej

a. sprawdzenie poprawności wykonania plików oraz opisów;
b. łączenie obiektu oraz miniatury z opisem bibliograficznym;
c. oznaczenie kolekcji;
d. uzupełnianie opisu o dane prawne;
e. upublicznienie obiektu.

Wykonanie: redaktor KPBC.

7. Kontrola metadanych, linkowanie i przesyłanie opisu do NUKAT

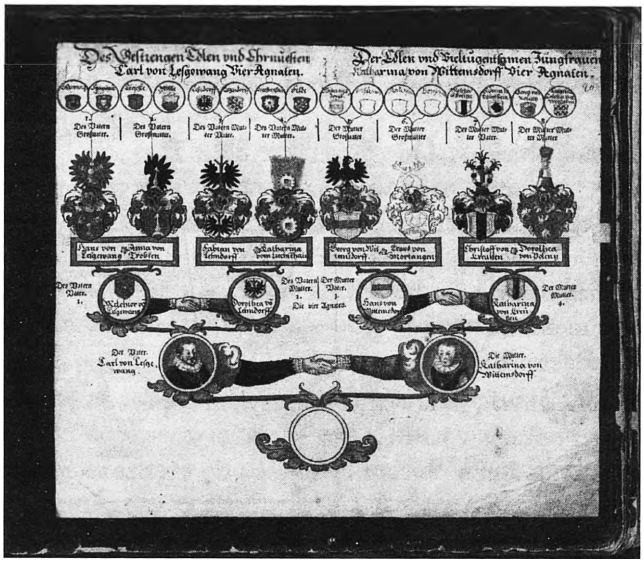
a. melioracja opisów w KPBC – raz w roku;
b. wprowadzanie linku obiektu cyfrowego do Horizona i istniejących rekordów w bazie Nukat;
c. koordynacja poprawek.

Wykonanie: główny kataloger.

Trzeba wiedzieć, że niektóre procesy pracy przebiegają równolegle, inne w następstwie czasu. Pojawiają się tzw. „wąskie gardła" – procesy, które zatrzymują prace i bez dodatkowych nakładów sił czy finansowanych nie da się ich usprawnić. To też należy przewidzieć lub elastycznie na nie zareagować, jeśli się pojawią. Takim „wąskim gardłem" może być przetwarzanie obrazów i OCR, jak w przypadku KPBC. Przy zastopowaniu prac mamy do wyboru w zależności od przyczyny:

- Zmodyfikować zadanie;
- Usprawnić procesy pracy;
- Zlikwidować zadanie;
- Zatrudnić nowego pracownika lub zmienić skład zespołu;
- Zdobyć dodatkowe środki.

Wszystkie problemy związane z zarządzaniem procesami pracy i KPBC wymagają kompromisów i rozmów wewnątrz instytucji. Dzięki nim mogą następować pozytywne zmiany i rozwój. Natomiast są także problemy, które wychodzą poza instytucję rodzimą, do nich należą wszelkie relacje związanie z pozyskiwaniem utworów do KPBC, przepisy prawne, nieujednolicone formy zapisu danych i inne. Te trzeba uzgadniać na poziomie krajowym a czasem, jak w przypadku określania statusu prawnego dzieła.



Rys. 3. Przykładowy obiekt z KPBC: Hennenberger Johann: *Stemmata genealogica praecipuarum in Prussia Familiarum Nobilium*; rękopis.

Zarządzanie prawami autorskimi w KPBC

Jest to jedno z najtrudniejszych zadań bibliotekarzy cyfrowych przy tworzeniu kolekcji. Zmusza do studiowania przepisów prawa i uczestniczenia w debatach na ten temat organizowanych przez prawników. Samo przeczytanie przepisów niczego nie daje, bibliotekarze nie znają języka prawa i nie potrafią go interpretować. Dlatego przy takich projektach najlepiej jest zatrudnić prawnika, lub wykształcić pracownika biblioteki w tym kierunku. By dobrze wywiązać się z tego trudnego zadania w roku 2004 zamówiono dla KPBC ekspertyzy i opinie prawne w Instytucie Prawa Własności Intelektualnej UJ wskazując jednocześnie na standardy obowiązujące w Europie i USA, by prawnicy mogli się do nich odnieść. Poradzono się także prawników uczelnianych, dzięki którym powstały jasno określone zasady pozyskiwania i udostępniania dzieł w KPBC.

W chwili obecnej mamy taki stan, że opracowaliśmy nie tylko politykę ogólną udostępniania KPBC http://kpbc.umk.pl/dlibra/text?id=polityka, ale nadaliśmy status prawny każdemu obiektowi. Mamy bibliotekarzy-prawników, którzy dyskutują nad różnymi rozwiązaniami i jednego bibliotekarza wykształconego kierunkowo w tym zakresie. W planie są regularne szkolenia z tego zakresu dla większej grupy pracowników. Prace nad tymi rozwiązaniami trwały w Toruniu 5 lat, wymagały wielu konsultacji także zagranicznych.

Zgodnie z prawem autorskim w KPBC starsze obiekty idą do Internetu bez ograniczeń (po 70 latach od śmierci autora, czasopisma po 70 latach od daty druku) i maja wówczas status *domena publiczna (ponad 35 tys.)*, o młodsze za każdym razem pytamy oraz podpisujemy licencje z wydawcą książki czy czasopisma albo z autorem,

czy spadkobiercą, a czasem z oboma na raz. Każdy obiekt to indywidualna historia prawna. Coraz częściej uzyskujemy zgodę właścicieli praw majątkowych na zastosowanie nowoczesnych wolnych licencji Creative Commons. W KPBC jest już ich ponad 240. W obu przypadkach status dzieła określa konkretna licencja, przy której dokładnie podajemy zasady korzystania z danego materiału.

Zarządzanie prawami jest trudne, żmudne, kosztowne i czasem nieefektywne. Wszystkie licencje trzeba przechowywać w postaci papierowej, prowadzić szeroką korespondencję z autorami i wydawcami, dzwonić, przekonywać, wyjaśniać i negocjować. Jak w wydawnictwie. Kiedyś bibliotekarze w ogóle nie zajmowali się takimi sprawami, bo niczego samodzielnie nie skanowali i nie upubliczniali na nowym polu eksploatacji. Nowa dLibra ma moduł samoarchiwizacji, dzięki któremu każdy autor może sam zdeponować utwór w bibliotece cyfrowej. Przy zastosowaniu tego rozwiązania czas na rozmowy z wersja autorami skraca się do minimum.

Dla ułatwienia pracy warto przygotować sobie serię standardowych wzorów pism, które potem modyfikując rozsyła się do właścicieli praw majątkowych. Warto prowadzić bazę danych ewidencjonującą historię negocjacji, ponieważ wskazują one na „dokonywanie szczególnej staranności" w wyjaśnianiu prawnej sytuacji materiałów przeznaczonych do skanowania. Jest to niezbędne w trakcie podejmowania decyzji np. przy upublicznianiu dzieła osieroconego. Warto poprosić o pomoc w pozyskiwaniu praw autorskich instytucje współpracujące czy inne biblioteki z regionu.

Na uczelni dobrze jest przyjąć ogólną politykę publikowania elektronicznego czy archiwizowania, zgodną z zaleceniami Ruchu Open Access, która może przyspieszyć pozyskiwanie praw. Pomocne w tym zakresie mogą być przepisy prawne związane z utworem powstającym w stosunku pracy. Modele takich polityk są opracowane przez Ruch Open Access. Trudnym może okazać się przekonanie rektorów, że taka polityka jest niezbędna i z tym trzeba się liczyć.

Najnowsze sygnały idące z MNiSW wskazują, że polityka państwa może zmierzać w tym kierunku, co dałoby bibliotekarzom potężne narzędzie przyspieszające udostępnianie współczesnych zasobów nauki w bibliotekach cyfrowych. Hasło otwarta nauka staje się modne i ważne.

Na zakończenie

Twórcy KPBC stale poszukują nowych rozwiązań, które wzbogacają nie tylko te bibliotekę cyfrową, ale i inne w Polsce. Poszukiwania innowacji prowadzą ich za granicę, gdzie rozwija się wiele inicjatyw technologicznych czy prawnych, które można adoptować do warunków krajowych. Wydaje się, że te kontakty dają obopólną korzyść zarówno nam, jak i naszym partnerom. Najważniejszym jest, żeby nie spocząć na laurach i zawsze poszukiwać nowych wyzwań, poprawiać jakość usług w zależności od potrzeb użytkowników i zadań instytucji, które powinna realizować. Najpoważniejszym wyzwaniem bibliotek w Polsce jest opracowanie długoterminowej procedury konserwacji i ochrony zasobu cyfrowego – do tego trzeba się teraz przymierzyć.

Małopolska Biblioteka Cyfrowa – regionalne repozytorium wiedzy

Wojciech Kowalewski, Władysław Marek Kolasa

W tekście przedstawiono pokrótce historię oraz zasoby Małopolskiej Biblioteki Cyfrowej. Dodatkowo zaprezentowano także jej funkcję jako regionalnego repozytorium wiedzy.

Wstęp

Dynamiczny rozwój technologii informatycznych stwarzający warunki do masowego przenoszenia treści drukowanych na papierze na nośnik elektroniczny, jak również wykorzystanie szerokopasmowego dostępu do sieci globalnej umożliwiające wręcz nieograniczone udostępnianie plików w formie zapisu rastrowego, zapoczątkował w XXI w. ruch na rzecz budowania rozproszonych sieci bibliotek cyfrowych w Polsce. Zadania tego podjęły się głównie naukowe, a z czasem również publiczne książnice, które (wypełniając misję mediacyjną oraz) sprawując informacyjną i edukacyjną funkcję, mogły stworzyć odpowiednie warunki do zaistnienia digitalizacyjnych laboratoriów i pracowni (przeszkolenie personelu, zakup oprogramowania zarządzającego obiektami cyfrowymi, dostęp do zbiorów bibliotecznych, itd.). Obecność bibliotek cyfrowych kształtuje więc postać globalnych zasobów elektronicznych w XXI w., stara się również porządkować wirtualny chaos informacyjny, zmierzając ku budowaniu semantycznego ładu danych cyfrowych.

Digitalizacja dokumentów pozwoliła na łatwiejsze przygotowywanie pełnotekstowych baz danych, które z kolei umożliwiły szybszy dostęp do zasobów nauki i kultury. W latach 2005-2010 rozpoczęto formowanie bibliotek cyfrowych dla konkretnych profili tematycznych. Kierowano się przy tym możliwościami dostępu do określo-

nych zbiorów bibliotecznych, stanem ich zachowania, już zbudowanymi kolekcjami cyfrowymi, a także sugestiami środowiskowymi (zapytania i prośby czytelników; ustalenia instytucji współpracujących z bibliotekami cyfrowymi itp.). Spowodowało to, iż spośród kilkudziesięciu instytucji zaangażowanych w zarządzanie elektroniczną treścią, część aktywnie działających bibliotek, skupiła się głównie na digitalizacji prasy lokalnej i regionalnej, zagrożonej głównie ze względu na złą jakość papieru, na którym była ona drukowana (tzw. kwaśny papier). Jedną z tych bibliotek jest Małopolska Biblioteka Cyfrowa, która realizując intencjonalne dążenia portalu stanowiącego *silva rerum* małopolskiego e-regionalizmu, dąży do upowszechnienia dziedzictwa kulturowego regionu w formie cyfrowej [Kowalewski 2007; Kowalewski; Kolasa 2008].

Wzrastająca powszechność dostępu do technologii cyfrowych w bibliotekach, niesie ze sobą nie tylko wzrost potencjału intelektualnego społeczeństwa, umacnia również nowy paradygmat światopoglądowy polskiego bibliotekarstwa w XXI w., otwartego na zmiany społeczno-kulturowe i nieustanny postęp technologiczny. Aby więc móc kształtować rozwój biblioteki cyfrowej należy rozpoznać czytelników korzystających z jej zasobów, ich sposób komunikacji, konstruowania zapytań, a nawet procesów rozumowania. W tym celu organizatorzy Małopolskiej Biblioteki Cyfrowej monitorują zapytania wyszukiwawcze czytelników agregowane przez narzędzia wyszukiwawcze systemu dLibra [Kowalewski 2010].

Analiza tych zapytań wskazuje, iż spora część użytkowników cyfrowej książnicy stosuje zapytania dotyczące historii regionu. Poszukujący informacji kulturalnej, regionalnej lub naukowej, to najczęściej naukowcy, nauczyciele i dziennikarze, ale też historycy, genealodzy oraz miłośnicy małych ojczyzn. W zdecydowanej większości są to osoby związane z Małopolską. Występuje również grupa czytelników spoza regionu. Są to internauci skupieni wokół zbiorów dotyczących tematyki literackiej, literacko-naukowej, kulturalnej, społecznej, politycznej, filozoficznej. Reprezentują ją czytelnicy poszukujący tytułów prasowych tj.: *Kurier Literacko-Naukowy, Czas, Życie Literackie, Ilustrowany Kurier Codzienny, Język Polski, Rocznik Krakowski, Tygodnik Powszechny, Życie, Pamiętnik Literacki, Rocznik Podhalański, Wierchy, Józefa Czecha Kalendarz Krakowski, Dziennik Polski* itd.

Małopolską Bibliotekę Cyfrową od grudnia 2006 do stycznia 2011 r. odwiedziło ponad 5 200 000 internautów. Dzienne odwiedziny kształtują się na poziomie od 1900 do 2700 użytkowników pobierających nawet 5 GB danych. W tym kontekście ważnym krokiem ku budowaniu cyfrowych kolekcji jest kontynuowanie digitalizacji ważnych czasopism polskich z XIX i XX w. Prócz już udostępnionych wysokonakładowych gazet: *Czas* (1848-1939), *Dziennik Polski* (1945-), *Tygodnik Powszechny* (1945-), udostępniono również mniejsze wydawnictwa, tj.: *Życie Literackie* (1951-1991), *Życie* (1897-1900), *Kurier Literacko-Naukowy* (1924-1939), *Niwa* (1872-1905), *Świat* (1888-1895), *Kraj* (1882-1909), *Kalendarz Krakowski Józefa Czecha* (1832-1917), *Pamiętnik Literacki, Język Polski, Djabeł, Wierchy, Rocznik Krakowski, Przekrój* (1945-2000), *Wiadomości Literackie* (1924-1939) oraz 'gadzinowy' *Goniec Krakowski*.

Niewykluczone, że plany digitalizacyjne obejmą również tytuły, które na stałe wpisały się do dziejów polskiego dziennikarstwa, m.in.: *Gazeta Krakowska* (1794-1848) Jana. A. Maja. Do spraw decyzyjnych należeć będzie natomiast kwestia skanowania pism współczesnych (po 1945 roku): *Echo Krakowa* (1946-1997), *Gazeta Krakowska* (1949-1990), *Tempo* (1948-1999), *Czas Krakowski* (1990-1997).

Przeszłość

Wojewódzka Biblioteka Publiczna w Krakowie w ramach Programu Ministra Kultury i Dziedzictwa Narodowego (Dziedzictwo kulturowe Priorytet 4 – Tworzenie zasobów cyfrowych dziedzictwa kulturowego), rozpoczęła w roku 2009 realizację projektu *Utworzenie cyfrowego zasobu dziennika Czas* [Kowalewski; Kolasa 2010]. Zadanie to objęło digitalizację (dziennika) *Czas* wydawanego w latach 1848-1939 oraz udostępnienie jego elektronicznej wersji w Małopolskiej Bibliotece Cyfrowej. W procesie skanowania wykorzystane zostały roczniki (Czasu) znajdujące się w posiadaniu Biblioteki Książąt Czartoryskich, należącej do grona bibliotek naukowych, których zbiory tworzą Narodowy Zasób Biblioteczny. Było to możliwe dzięki zawartej umowie o współpracy w ramach Małopolskiej Biblioteki Cyfrowej pomiędzy Muzeum Narodowym w Krakowie a Wojewódzką Biblioteką Publiczną w Krakowie oraz pisemnie wyrażonej zgody przez Zarząd Fundacji XX Czartoryskich na skanowanie dziennika *Czas*.

Celem nadrzędnym projektu była ochrona i zachowanie materialnego dziedzictwa kulturowego poprzez utrwalanie zasobów w formie cyfrowej, a także poszerzenie oferty biblioteki o usługi dostosowywane do aktualnych potrzeb użytkowników i wymogów stawianych przez szybko zmieniające się standardy zdobywania wiedzy. Zwiększenie ogólnodostępnych zasobów Małopolskiej Biblioteki Cyfrowej stało się w tym kontekście równie ważnym zadaniem wpływającym na wzrost wartości informacyjnej ogólnoświatowych zasobów WWW. Dostęp do zdigitalizowanego *Czasu* miał na celu przyciągnięcie nowych czytelników do biblioteki cyfrowej; był to dziennik zróżnicowany tematycznie, dotykający tematyki politycznej, społecznej, historycznej i literackiej.

Realizacja projektu wymogła podjęcia działań w następujących zakresach:

Digitalizacja Czasu za lata 1848-1939

Skanowanie zbiorów odbyło się w pracowni digitalizacyjnej Małopolskiej Biblioteki Cyfrowej, realizacji zadania podjęli się natomiast pracownicy Wojewódzkiej Biblioteki Publicznej w Krakowie z Zespołu MBC. Do zdigitalizowania wyznaczono 224 oprawne woluminy *Czasu*. W celu zagwarantowania jak najlepszej jakości prezentowanych dokumentów oraz zapewnienia odpowiednich warunków podczas skanowania cennych oryginałów konieczne było wyposażenie laboratorium w niezbędne

urządzenia skanujące – skaner wielkoformatowy, a także sprzęt komputerowy wraz z oprogramowaniem do edycji graficznej. Digitalizacja pisma odbyła się zatem przy zastosowaniu najnowocześniejszych technologii, wykorzystywanych do skanowania zbiorów bibliotecznych. W tym celu zakupiono specjalistyczny skaner dziełowy, umożliwiający skanowanie kolorowe w formacie A2 z rozdzielczością 600 dpi.

Wysoka jakość skanowanego obrazu, cechująca się dużą rozpiętością tonalną oraz dużą ostrością była podstawowym założeniem digitalizacji dziennika *Czas*; miało to odzwierciedlenie w postaci zwiększania pojemności magazynowania plików danych. Dlatego prócz nabycia urządzeń skanujących do pracowni digitalizacyjnej, konieczne było zakupienie urządzeń do masowego gromadzenia zbiorów cyfrowych (macierze dyskowe).

Opracowanie metadanych dokumentów i udostępnienie materiału online

Po procesie skanowania i przetwarzania obiektów w laboratorium, pliki cyfrowe trafiały do redaktora zajmującego się przygotowaniem opisów bibliograficznych. Następnie udostępniano zbiory w Internecie. Proces opracowania opisów do *Czasu* był realizowany przy wykorzystaniu standardu metadanych zdefiniowanych do opisu zbiorów internetowych opartych na zestawie (15) elementów zgodnych z *Dublin Core Metadata Element Set.* Opracowane w ten sposób metadane wydawnictw zostały udostępnione w rozproszonej sieci polskich bibliotek cyfrowych PIONIER, a następnie w bibliotece cyfrowej Europeanie, która będąc elementem europejskiej infrastruktury naukowej oferuje dostęp do zdigitalizowanych dokumentów dziedzictwa kulturowego znajdujących się w różnych instytucjach kultury w Europie.

Internetowe udostępnianie zeskanowanych numerów *Czasu* było realizowane na bieżąco, w miarę postępu prac nad przetwarzaniem i elektronicznym opracowaniem kolejnych tomów dziennika. Ze względu na czasochłonność zadania, planowane działania nie zostały zakończone w 2009 r. i trwały nieprzerwanie do połowy 2010 r.

Przyszłość

Kolejnym dużym projektem realizowanym w ramach Wieloletniego Programu Rządowego Kultura+ Priorytet Digitalizacja, stało się zdigitalizowanie *Ilustrowanego Kuryera Codziennego (IKC).* O wytypowaniu *IKC* do skanowania zdecydowała przede wszystkim jego pozycja wśród XX-wiecznej prasy polskiej. Był to jeden z ważniejszych w Polsce dzienników, wydawany w latach 1910-1939. Według zgodnej opinii historyków i badaczy kultury *IKC* był pismem prezentującym wysoki poziom merytoryczny i dziennikarski oraz uchodził za jedną ważniejszych (prócz *Czasu*) polskich gazet znanych i komentowanych w stolicach europejskich. Podobnie jak większość polskich gazet I połowy XX wieku *IKC* jest również narażony na chemiczną degradację papieru (kwasowa hydroliza celulozy), która nieuchronnie postępuje – digitalizacja pis-

ma stała się jedynym sposobem zabezpieczenia go dla potomnych. Z uwagi na dużą poczytność pisma, nie zachowały się do dziś jego kompletne archiwa, jednak spora ich część znajduje się w archiwach i bibliotekach, z którymi współpracuje Małopolska Biblioteka Cyfrowa. Przeprowadzony wywiad wykazał, iż w zbiorach Archiwum Państwowego w Krakowie znajduje się archiwum *IKC* za lata 1925-1938. Zdigitalizowanie i udostępnienie tego zasobu w MBC będzie w przyszłości stanowiło podstawę do ubiegania się o wypożyczenie pozostałych woluminów z lat 1910-1924 oraz 1939 z innych bibliotek w kraju.

Tworzenie elektronicznych publikacji na podstawie skanowanego *IKC* będzie odbywać się na bieżąco w miarę postępu prac nad przetwarzaniem i opracowaniem zasobu. Zamieszczenie wszystkich obiektów potrwa do połowy 2012 r. i będzie realizowane środkami i siłami WBP. Pozyskany w wyniku skanowania *IKC* materiał posłuży również do zaprezentowania wystawy wirtualnej, której pomysłodawcą i realizatorem będzie członek zespołu MBC, specjalista w zakresie historii czasopiśmiennictwa w Polsce. W trakcie digitalizowania *IKC* uwzględnia się konieczność współpracy na poziomie krajowym z instytucjami i organizacjami zajmującymi się zachowaniem i ochroną dziedzictwa kulturowego i naukowego. Będzie to służyć koordynacji działań w dziedzinie digitalizacji, dążeniu do zachowania wspólnych standardów oraz wymianie doświadczeń przy rozwiązywaniu niektórych problemów.

Kierunki rozwoju

Do głównych zasobów dziedzictwa kulturowego Małopolski, które z obiektywnych przyczyn wymagają zabezpieczenia należy prasa. Wynika to z trzech przyczyn. Po pierwsze zbiory te z uwagi na zły stan zachowania i postępującą degradację papieru należy zabezpieczyć dla potomnych. Po wtóre – zasoby biblioteczne prasy małopolskiej (w tym ogólnopolskiej) są niemal wyłącznie dostępne na terenie Małopolski, głownie w Krakowie. Po trzecie – w odróżnieniu od książek uregulowania prawne w odniesieniu do czasopism są precyzyjne, co pozwala zaplanować opracować i zaplanować ten proces od strony prawnej. Na terenie Małopolski istnieją dwa ośrodki, które są w stanie zrealizować to przedsięwzięcie: z jednej strony Wojewódzka Biblioteka Publiczna w Krakowie wraz z partnerami operująca Małopolską Biblioteką Cyfrową, która specjalizuje się w zbiorach XX-wiecznych i współczesnych; z drugiej Biblioteka Jagiellońska, która zajmuje się zbiorami do 1918 r. W wyniku uzgodnień opracowano wstępny podział prac dotyczący obiektów najważniejszych i najczęściej wykorzystywanych. Biblioteka Jagiellońska – zobowiązała się zdigitalizować i udostępnić 58 tytułów, głównie lwowskich i XIX-wiecznych (w tym m.in. *Gazeta Lwowska, Nowa Reforma, Piast*)[1] a Wojewódzka Biblioteka Publiczna – 67 tytułów, głównie krakowskich i XX-wiecznych (w tym m.in. *Czas, Ilustrowany Kurier Codzienny, Goniec Krakowski, Dziennik Polski, Tygodnik Powszechny, Przekrój* i in.) oraz ok. 120 tytułów sublokalnych (prasy powiatowej, gminnej i dzielnicowej). Dotychczas z wykazu przynależne-

[1] Wykaz: http://jbc.bj.uj.edu.pl/dlibra/dlibra/text?id=info-projekt-mrpo [odczyt: 15.04.2011].

go do WBP zdigitalizowano *Czas* i tytuły wydawane po 1945; do realizacji pozostał jedynie *Ilustrowany Kuryer Codzienny* oraz *Goniec Krakowski*. W dalszej perspektywie należy starać się objąć akcją wszystkie wydawane w Małopolsce tytuły, a pełna liczba obiektów, która winna być objęta tym procesem jest znaczna (ok. 5000 tytułów).

Aktualnie w najlepszej sytuacji jest stan zabezpieczenia małopolskich gazet ulotnych i seryjnych. Według przeprowadzonych w Krakowie badań – wydrukowano 194 takie gazety – czyli ok. 30% wszystkich pism wydawanych w Polsce (w tym: 24 tytuły w XVI w., 94 w XVII oraz 60 w XVIII i 16 o nieustalonej dacie. Zasób ten jest już w pełni zdigitalizowany i udostępniony w portalu Cyfrowej Biblioteki Druków Ulotnych Polskich i Polski Dotyczących z XVI, XVII i XVIII wieku [Zawadzki 1977-1990]. Nieco gorzej przedstawia się stan zabezpieczenia gazet wydawanych w schyłkowych latach XVIII wieku (razem w Krakowie i Lwowie wydawano ok. 15 tytułów), lecz z wyjątkiem *Gazety Krakowskiej* 1796-1846 i lwowskiego *Dziennika Patriotycznych Polityków* 1792-1798 – nie ma wśród nich pozycji szczególnie ważnych dla kultury narodowej [Hombek 2011].

W najgorszym stanie zachowania i udostępniania są pozycje XIX-wieczne i prasa okresu Polski Odrodzonej 1918-1939. W okresie tym ukazywało się ok. 4600 tytułów; w zbliżonych proporcjach pisma wydawane w Krakowie i we Lwowie (po ok. 2300 tytułów). Z czego w Krakowie do 1864 r. – 76 tytułów [Jarowiecki 2003], a w latach autonomii – 1091 tytułów [Jakubek 2004] oraz 1124 w okresie II RP. We zaś Lwowie: ok. 96 do 1884 r. [Jarowiecki 2008]; 1119 w latach autonomii [Jarowiecki 2002] oraz 1173 w dwudziestoleciu międzywojennym [Jarowiecki; Góra 1994].

Spośród najstarszych gazet krakowskich (do 1864 r.) prawidłowo zabezpieczony i zdigitalizowany jest wymieniony już *Czas* (1848-1939); wśród najważniejszych dezyderatów dotychczas nieobjętych akcją scalania i digitalizacji należy: przede wszystkim *Gazeta Krakowska* Jana Maja oraz gazety Konstantego Majeranowskiego: *Pszczółka Krakowska* (1819-1822), *Krakus* (1822), *Muza Nadwiślańska* (1823), *Pielgrzym z Tęczyna* (1823), *Pszczółka Polskiej* (1823), *Rozrywki Przyjemne i Pożyteczne* (1826-1827), *Kurier Krakowski Płci Pięknej i Literaturze Poświęcony* (1827), *Goniec Krakowski* (1828-1830) i *Wanda* (1829-1830).

Z ponad tysiąca pism wydawanych w Krakowie w okresie autonomii, zdigitalizowano do tej pory kilka, w tym najważniejsze pismo tego okresu – dziennik *Czas* oraz organ modernistów *Życie* (1897-1900); częściowo socjalistyczny *Naprzód* (1894-1939), *Przegląd Polski* (1866-1914), *Diabeł* (1869-1922) i *Krytyka* (1899-1914) oraz objęte planem BJ: *Nowa Reforma* i *Piast*. Wciąż pozostało co najmniej kilkadziesiąt pism o dużym znaczeniu dla kultury narodowej, w tak ważne gazety i czasopisma jak: *Głos Narodu, Gazety Powszechna, IKC* (1910-1939); *Nowiny dla Wszystkich/Nowiny* (1903-1914); *Przegląd Powszechny* (1884-1939); *Prawda* (1895-1918); *Przegląd Wszechpolski* (1895-1905); *Ojczyzna* (1903-1929), *Polak* (1896-1909) oraz liczne pisma literackie i artystyczne (Świat, *Museion*) i naukowe (w tym wydawnictwa seryjne Akademii Umiejętności) [szerzej: Jarowiecki; Kolasa 2006, s. 25-53].

W najgorszej sytuacji są gazety i czasopisma okresu II RP. Spośród 1124 tytułów jedynie kilka zostało objętych procesem digitalizacji (*Kurier Literacko-Naukowy, Czas*)

lub jest w trakcie skanowania, m.in. *Nowa Reforma, Piast*. Zdecydowana większość, w tym pisma najważniejsze wciąż czekają na realizację. Do najpilniejszych dezyderatów należą przede wszystkim dzienniki: *Ilustrowany Kurier Codzienny, Głos Narodu, Naprzód* (po 1918); popularne tygodniki wydawane przez koncern *IKC*: *Tempo Dnia* (1933-1939); Światowid (1924-1939); *Na Szerokim Świecie* (1928-1939); *Wróble na Dachu* (1930-1939); *Tajny Detektyw* (1931-1934); *Raz, Dwa, Trzy...* (1931-1939) i *Magazyn As* (1935-1939); wiele czasopism literackich, jak: *Zwrotnica , Linia, Maski, Helion; Gazeta Literacka* (1926-1927); *Gazeta Literacka* (1931-1934); *Nasz Wyraz, Biuletyn Literacki, Krak* (1930); *Sztuki Piękne, Wieś* (1934); szeroka gama wydawnictw religijnych (razem 52 tytuły, w tym *Przegląd Powszechny*); naukowych i fachowych (330 tytułów) [szerzej: Kolasa 1990] oraz ok. 170 pism żydowskich [Brzoza 1993, s. 71-110], w tym najważniejsza polskojęzyczna gazeta syjonistów polskich *Nowy Dziennik* (1918-1939).

W okresie okupacji wydawano w Krakowie 158 pism konspiracyjnych [Jarowiecki 1978] szczegółowo opisanych i przeanalizowanych przez Jerzego Jarowieckiego w monografii *Konspiracyjna prasa w Krakowie w latach okupacji hitlerowskiej 1939-1945* (Kraków 1980). Dotychczas żadne pismo z tego okresu nie zostało poddane digitalizacji. Wydaje się jednak, że z uwagi na charakter tych wydawnictw (zwykle powielane efemerydy) digitalizacja prasy konspiracyjnej nie wydaje się ani pilna, ani szczególnie celowa. Można jedynie rozważyć digitalizację długowiecznego *Dziennika Polskiego* (1940-1945) – konspiracyjnego organu Stronnictwa Demokratycznego. Nadto warte uwagi są dwa tytuły gadzinowe: *Goniec Krakowski* (1939-1945) i *Gazeta Żydowska* (1940-1942) oraz niemieckojęzyczny organ Dystryktu Krakowskiego *Krakauer Zeitung* (1939-1945).

W okresie 1945-1989 ukazywało się w Krakowie ok. 400 tytułów. Stan zabezpieczenia tej grupy jest stosunkowo dobry. Dotychczas zdigitalizowano prawie cały zasób głównej małopolskiej gazety *Dziennik Polski* (od 1945), cały *Tygodnik Powszechny, Życie Literackie* (1951-1989), *Dziennik Literacki* (1947-1950), *Znak* (od 1946); trwają prace nad skanowaniem *Przekroju* (od 1945). Warto dodać, że gazety wydawane w tym okresie wciąż pozostają w zakresie ochronny prawnej – stąd do ich zdigitalizowania i udostępnienia niezbędne były odpowiednie umowy, które wynegocjowano dla MBC. Wśród kilku dezyderatów dla tego okresu trzeba wskazać przede wszystkim na wydawane w Krakowie u progu PRL: *Odrodzenie i Twórczość*, partyjny dziennik wojewódzki *Gazeta Krakowska* (od 1949); sportowe *Tempo* (1959-2008), i *Echo Krakowa* (1946-1996). Realizacja ostatnich postulatów jest jednakże uzależniona od pozyskania stosownych umów licencyjnych od wydawców lub właścicieli praw majątkowych (głównie Polskapresse).

Współczesna prasa lokalna i regionalna

Ostatnia grupą wydawnictw objętych procesem zabezpieczenia jest małopolska prasa lokalna. W latach 1990-2011 ukazywało się w Małopolsce 498 takich tytułów, z czego 185 to gazety informacyjne i informacyjno-publicystyczne, inne grupy są mniej liczne [Kolasa; Michalska 2007, s. 219-254]. Aktualnie stan zabezpieczenia tej grupy jest stosunkowo dobry. W latach 2006-2010 Małopolska Biblioteka Cyfrowa podpisała z wydawcami szereg umów i deklaracji, co umożliwiło bieżącą współpracę z najważniejszymi redakcjami z terenu Małopolski. W wyniku tej współpracy MBC pozyskuje i udostępnia ok. 120 tytułów treści ogólnej (czyli 64% ich ogólnej liczby). Wśród planów z tego zakresu warto odnotować pisma mniejszości narodowych, w szczególności: słowackojęzyczny *Zivot* (od 1958*), Biuletyn Ormiańskiego Towarzystwa Kulturalnego* (od 1993), romski *Dialog-Pheniben* oraz łemkowską *Watrę* (od 1992). Realizacja tego postulatu, podobnie jak dla prasy okresy PRL, uzależniona od pozyskania stosownych umów licencyjnych od wydawców.

Udostępnienie lokalnych i regionalnych zbiorów w bibliotece cyfrowej nie jest przedsięwzięciem okazjonalnym, bowiem celem Małopolskiej Biblioteki Cyfrowej jest także upowszechnianie współczesnej prasy lokalnej i regionalnej Małopolski, poprzez gromadzenie druku cyfrowego [Kowalewski 2008]. Do obowiązków Wojewódzkiej Biblioteki Publicznej w Krakowie, będącej organizatorem MBC należy już nie tylko gromadzenie lokalnych wydawnictw oraz dokumentów życia społecznego z regionu w formie tradycyjnej, ale też, co podyktowane jest rozwojem technologicznym i wpływem nowoczesnych środków masowego przekazu, konsolidacja tych wydawnictw w formie elektronicznej. Dzięki temu działaniu wydawnictwa regionalne poddawane są archiwizacji i remediacji, a więc ponownemu włączeniu do obiegu informacji. Istotną funkcję pełni przy tym próba semantycznego uporządkowania rozproszonych w sieci e-wydawnictw, która realizuje podstawowe założenia użytkowników Web 2.0.

Wydawcy prasy lokalnej i regionalnej nie od dziś zmagają się z wieloma problemami, głównie finansowymi, które bezpośrednio przekładają się na niski nakład wydawniczy oraz ograniczony kolportaż prasy. W rezultacie, pisma te docierają do stosunkowo małego audytorium czytelniczego. Wobec powyższych trudności, digitalizacja i cyfrowa publikacja regionalnych i lokalnych wydawnictw wydaje się być oczywista i konieczna. Udostępnienie wydawnictw w globalnej sieci WWW zachęcić ma również regionalnych pasjonatów do tworzenia własnych zbiorów cyfrowych obiektów rzadkich i zagrożonych degradacją (archiwalia, rękopisy, stare druki), poprzez ich fotografowanie i skanowanie. Celom powyższym służy Małopolska Biblioteka Cyfrowa, której zadaniem jest również zintegrowanie współczesnej prasy regionalnej i lokalnej. Przedsięwzięcie to jest konieczne, bowiem wydawnictwa lokalne i regionalne nie są systematycznie gromadzone ani przez Biblioteką Narodową ani przez inne biblioteki naukowe w Polsce. W efekcie rejestracji bibliograficznej wymyka się spora część publikacji, pozostała natomiast niknie w lawinie informacji.

Wybrane procedury w ramach realizowanych projektów Małopolskiej Biblioteki Cyfrowej

Zakres wykonywanych czynności:

1. skanowanie obiektów,
2. zarządzania plikami w przestrzeni sieciowej (przetwarzanie i edycja obrazów; walidacja, inwentaryzacja oraz archiwizacja danych),
3. konwersja obiektów cyfrowych do formatu prezentacyjnego,
4. edycja plików w zakresie wymaganych procedur MBC,
5. kontrola, inwentaryzacja i archiwizacja danych.

Wytyczne techniczne (ogólne):

1. Dokumenty są digitalizowane według ustalonych procedur skanowania (skanowanie stron pojedynczo przy częściowym zaciemnieniu laboratorium, strony pojedyncze przygotowywane są w jednym pliku).
2. Nazwy plików zapisywane są w następującej strukturze katalogów: rok/miesiąc/nr kolejny pliku (np. 1925/08/0001.tif).
3. Obiekty przed dokonaniem czynności edytorskich są poddane kontroli kompletności i jakości skanowanych stron. Zajmujący się tym redaktor przeprowadza walidację zdigitalizowanego materiału na zasadzie porównania materiału z oryginałem.
4. Stosowana jest korekta graficzna (kadrowanie, edycja tonalna itd.) a następnie archiwizacja na macierzy dyskowej w formie plików TIFF 6.0 (bez kompresji), o rozdzielczości 300 dpi i głębi kolorów 24-bit RGB.
5. Pliki zawierają metadane techniczne.
6. Kopie plików są przekazywane instytucji udostępniającej zbiory na pamięciach przenośnych.
7. Obiekty poddawane są procedurom nałożenia znaku wodnego oraz kompresji do formatu prezentacyjnego (DjVu). Stosowana jest procedura rozpoznawania druku (OCR).
8. Przygotowany dokument DjVu zapisywany jest jako plik rozdzielony. Nazwa tego pliku zawiera odpowiednią notację, np. ab_00001 (nie stosuje się spacji, wielkich/małych liter, ani znaków diakrytycznych).
9. Dla przygotowanych publikacji sporządzane są opisy katalogowe (metadane opisowe) w schemacie DublinCore.
10. Po przydzieleniu opisów bibliograficznych do odpowiedniej kolekcji w systemie dLibra, następuje połączenie opisów z treścią oraz publikacja w bibliotece cyfrowej.

Zakończenie

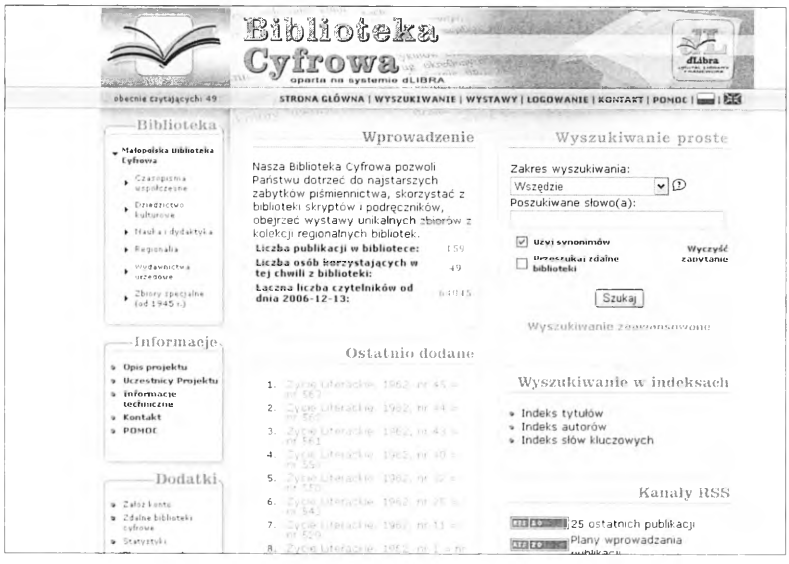
Skanowanie cennych dla polskiej kultury zbiorów bibliotecznych i muzealnych wytyczyło nowe ścieżki i dążenia polskich bibliotek w XXI wieku. Dzięki skanowaniu i udostępnianiu elektronicznej treści w Internecie zaczęły kształcić się pokolenia Polaków, którzy korzystając z wydobytej i przetworzonej do formy bitowej wiedzy kreują demokratyczne idee wolności i godności ludzkiej społeczeństwa obywatelskiego.

Ruch na rzecz budowania sieci rozproszonej bibliotek cyfrowych spowodował, że bibliotekarze cyfrowi zaczęli wzajemnie organizować się i, bez rządowego wsparcia oraz centralnej koordynacji zadań, tworzyć lokalne *entourage* bibliotek cyfrowych, nadając im z czasem określony rys historyczny i profilując przyszłe kierunki rozwoju.

Procesy digitalizacji zbiorów bibliotecznych rozpoczynano jednak na niskobudżetowych skanerach zlokalizowanych w bibliotecznych magazynach przystosowanych pod przyszłe pracownie. W momencie, gdy pojawiły się możliwości wykorzystania środków unijnych, masowa digitalizacja papieru nabrała wyraźnego tempa. Czas ten został dobrze wykorzystany przez bibliotekarzy, wśród których wykształcili się zawodowcy zdolni przenieść w postać bitową niemal milion publikacji.

Biblioteki cyfrowe stały się popularne, a coraz powszechniejszy dostęp do urządzeń mobilnych (e-czytnik, tablet) sprawił, iż przed elektronicznym drukiem pojawiły się kolejne wyzwania. Stoją one *nolens volens* przed twórcami Małopolskiej Biblioteki Cyfrowej, którzy są świadomi rosnących oczekiwań użytkowników elektronicznych książnic. Dostęp do cyfrowej treści musi stać się bardziej powszechny w czasach „mobilnej rewolucji". Tak, aby czytelnik miał możliwość jeszcze szybszego odnalezienia go i bezproblemowego (od)czytania.

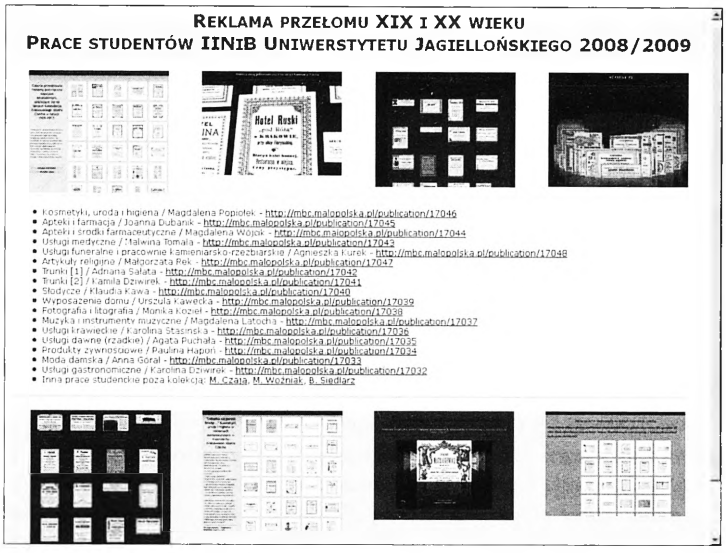
Aneks



Rys. 1. Wygląd strony internetowej Małopolskiej Biblioteki Cyfrowej (02.2007).



Rys. 2. Wygląd strony internetowej Małopolskiej Biblioteki Cyfrowej (09.2011).



Rys. 3. Prezentacja obiektów w formie kolekcji wystawy eksponowanej na stronie głównej MBC.

Bibliografia

1. Brzoza, Czesław (1990). *Polityczna prasa Krakowa 1918-1939.* Kraków: Uniwersytet Jagielloński, 199 s.
2. Brzoza, Czesław (1993). *Materials toward the bibliography of the jewish press in Cracow (1918-1939).* [W:] *Bibliographies of Polish Judaica.* Cracow: Jagiellonian University. Research Center of Jewish History and Culture in Poland, s. 71-110.
3. Hombek, Danuta (2001). *Prasa i czasopisma polskie XVIII wieku w perspektywie bibliologicznej.* Kraków: TAiWPN „Universitas", 383 s.
4. Jakubek, Mariusz (2004). *Prasa krakowska 1795-1918 : bibliografia.* Kraków: Wydaw. Naukowe DWN: Archiwum Uniwersytetu Jagiellońskiego, 387 s.
5. Jarowiecki, Jerzy (1978). *Katalog krakowskiej prasy konspiracyjnej 1939-1945.* Kraków: Wydawnictwo Naukowe WSP, 183 s.
6. Jarowiecki, Jerzy (2002). *Prasa lwowska w latach 1864-1918: bibliografia.* Kraków: Wydaw. Naukowe Akademii Pedagogicznej, 553 s.
7. Jarowiecki, Jerzy (2003). *Prasa w Krakowie : tradycja i współczesność (część pierwsza: do roku 1918).* „Rocznik Historii Prasy Polskiej", z. 1, s. 11-52.
8. Jarowiecki, Jerzy (2008). *Dzieje prasy polskiej we Lwowie do 1945 roku.* Kraków: Księgarnia Akademicka, 502 s.
9. Jarowiecki, Jerzy; Góra, Barbara (1994). *Prasa lwowska w dwudziestoleciu międzywojennym : próba bibliografii.* Kraków: Wydaw. Naukowe WSP, 422 s.
10. Jarowiecki, Jerzy; Kolasa, Władysław M. (2006). *Najważniejsze polskie gazety do wybuchu pierwszej wojny światowej w kontekście zabezpieczenia ich dla potomnych : charakterystyka, badania, zasoby.* [W:] Jazdon, Artur; Chachlikowska Aldona red. *Gazety:*

zasoby, opracowanie, ochrona, digitalizacja, promocja/informacja. Poznań: Biblioteka Uniwersytecka, s. 25-53.

11. Kolasa, Władysław M. (1990). *Prasa krakowska na warsztatach historyków.* [W:] *350 lat prasy polskiej* [w druku].
12. Kolasa, Władysław M.; Michalska, Mariola (2005). *Katalog małopolskich mediów lokalnych i regionalnych 2005.* Kraków: Wojewódzka Biblioteka Publiczna, 213 s.
13. Kolasa, Władysław M.; Paszko, Artur (2007). *Małopolskie media lokalne i regionalne w świetle analiz statystycznych.* [W:] *Polskie media lokalne na przełomie XX i XXI wieku.* Kraków: Wojewódzka Biblioteka Publiczna, s. 219-254.
14. Kowalewski, Wojciech (2007). *U progu Małopolskiej Biblioteki Cyfrowej.* „Małopolska" t. 9, s. 249-254.
15. Kowalewski, Wojciech (2008). *E-regionalizm a Małopolska Biblioteka Cyfrowa.* „Małopolska" t. 10, s. 301-309.
16. Kowalewski, Wojciech (2010). *Użytkownicy biblioteki cyfrowej w świetle analizy statystyk wyszukiwawczych MBC.* [W:] Mazurek, Cezary; Stroiński, Maciej; Węglarz, Jan red. *Polskie Biblioteki Cyfrowe 2009: materiały z konferencji zorganizowanej w dniach 8-10 grudnia 2009 r. przez Bibliotekę Kórnicką PAN, Poznańską Fundację Bibliotek Naukowych, Poznańskie Centrum Superkomputerowo-Sieciowe.* Poznań: Ośrodek Wydawnictw Naukowych, s. 31-34.
17. Kowalewski, Wojciech; Kolasa, Władysław M. (2010). *Dziennik Czas w Małopolskiej Bibliotece Cyfrowej.* „Małopolska" t. 12, s. 345-350.
18. Kowalewski, Wojciech; Kolasa, Władysław M. (2008). *Małopolska Biblioteka Cyfrowa – regionalne repozytorium informacji o kulturze.* „Notes Wydawniczy" nr 1, s. 32-36.
19. Zawadzki, Konrad (1977-1990). *Gazety ulotne polskie i Polski dotyczące XVI-XVIII wieku.* T. 1-3. Wrocław: Zakład Narodowy im. Ossolińskich, 228, 290, 97 s. Wersja elektroniczna na platformie Cyfrowej Biblioteki Druków Ulotnych Polskich i Polski Dotyczących z XVI, XVII i XVIII wieku [UW]: http://cbdu.id.uw.edu.pl/ [07.01.2011 – 2088 poz.].

Śląska Biblioteka Cyfrowa – analiza przypadku

Remigiusz Lis

W tekście przedstawiono Śląską Bibliotekę Cyfrową, regionalny zasób cyfrowy, który stanowi międzyinstytucjonalną organizację, tworzoną na zasadzie współdzielenia zasobów.

1. Wstęp

W dniu 19 sierpnia 2011 r. minęło pięć lat od opublikowania na platformie Śląskiej Biblioteki Cyfrowej pierwszej publikacji, co rozpoczęło proces ciągłego powiększania jej zasobów. Towarzyszyła temu sukcesywna rozbudowa jej infrastruktury, otwarcie na współpracę z instytucjami z różnych obszarów kultury i nauki oraz intensywna promocja przedsięwzięcia w zarówno w środowisku bibliotekarskim, jak i poza nim.

Z pewnej perspektywy można powiedzieć, że piśmiennictwo przechowywane w śląskich instytucjach kultury, a także poza nimi, jest wyjątkowo dobrym materiałem dla szerokiej, cyfrowej prezentacji w globalnej sieci. Śląsk bowiem to region o bogatej historii, obszar wielu historycznych państwowości, narodowości i wyznań, miejsce współistnienia różnorakich ruchów politycznych i społecznych. Ich obywatele, wyznawcy i uczestnicy w celu utrwalenia biegu życia codziennego i niekiedy trudnych dziejów dawali im świadectwo w bogatym piśmiennictwie, które, mimo strat i wojennych zawieruch, przetrwało w zasobach regionalnych instytucji, urzędów i kolekcjach osób prywatnych. Cyfrowy zasób tworzony na postawie tych dokumentów i udostępniony w Internecie cieszy się zainteresowaniem i jest użyteczny dla publiczności z całego świata, legitymując tym samym dalsze starania o rozwój regionalnej biblioteki cyfrowej.

Istotnymi uwarunkowaniami formowania regionalnej biblioteki cyfrowej były: niedostatek środków, wspomniane rozproszenie piśmiennictwa w wielu miejscach oraz kwestia ewentualnej współpracy wielu różnych instytucji, które często wcześniej ze sobą nie współpracowały.

Pierwszy problem, powszechnie dotykający twórców cyfrowych zasobów do dziś, mógł zostać chociaż częściowo zniwelowany przez przyjęcie zasady współdzielenia zasobów. Kwestia druga znalazła rozwiązanie w formule organizacyjnej oraz dojrzałych narzędziach zdalnej współpracy (oprogramowanie biblioteki cyfrowej, narzędzia komunikacji). Kwestia kooperacji instytucji stanowiła zarówno wyzwanie, jak i szansę. Z jednej strony inicjowana współpraca wymagała częstokroć wychodzenia poza instytucjonalne egoizmy, z drugiej – mogła się oprzeć o posiadane przez uczestników zasoby: doświadczenia współpracy na innych polach (współkatalogowanie, tworzenie sieciowych baz danych), umiejętności w zakresie realizacji projektów, informatyzacji systemów, czy wreszcie pomyślnych działań w zakresie tworzenia własnych zasobów cyfrowych.

Niniejsza analiza przypadku Śląskiej Biblioteki Cyfrowej zostanie dokonana z podkreśleniem wcześniejszych doświadczeń w zakresie digitalizacji zbiorów bibliotecznych jej organizacyjnego koordynatora – Biblioteki Śląskiej, nie umniejszając wartości digitalizacyjnych doświadczeń i osiągnięć innych instytucji obecnie współtworzących ŚBC. Zamiarem tej analizy jest bowiem ilustracja ciągłości procesu rozwojowego regionalnej biblioteki cyfrowej (obejmującej obecnie kilkudziesięciu uczestników), który został przygotowany na podstawie wcześniejszych prób i doświadczeń instytucji inicjującej.

Dla ścisłości należy dodać, że pilotażową inicjatywą o podobnym (regionalnym) charakterze na obecnym obszarze funkcjonowania ŚBC, która prezentowała kilkadziesiąt cyfrowych dokumentów w dostępnej w Internecie bazie danych, była *Cieszyńska Biblioteka Wirtualna* [dostępna do dziś pod adresem: http://archiwum.kc-cieszyn.pl/biblioteka/biblioteka.html], utworzona przez Książnicę Cieszyńską. Funkcjonowała ona w latach 2003-2007 i prezentowała regionalne źródła dla użytkowników zainteresowanych historią ziemi cieszyńskiej. Projekt jednak, wykorzystujący autorskie oprogramowanie, był trudny w rozwijaniu przez instytucję, dla której digitalizacja nie była kluczową kompetencją i został zakończony. Jej cyfrowe treści po przystąpieniu Książnicy do ŚBC zostały w niej opublikowane i poszerzone o dalsze publikacje.

2. Digitalizacja zbiorów i projektowe koncepcje digitalizacyjne w latach poprzedzających powołanie ŚBC (1998-2006)

Śląska Biblioteka Cyfrowa nie była pierwszą próbą ustanowienia serwisu prezentującego w Internecie zasób piśmienniczy. Z chwilą zaistnienia ŚBC, Biblioteka Śląska posiadała już tradycje w zakresie digitalizacji zbiorów bibliotecznych, a w latach poprzedzających powstanie biblioteki cyfrowej projektowano kompleksowe systemy

digitalizacji i udostępniania zbiorów cyfrowych. Pomimo, że próby te nie zakończyły się utworzeniem realnie działającego serwisu, pozwoliły one na zapoznanie się personelu Biblioteki Śląskiej z tworzeniem cyfrowych zasobów, wymogami technologicznymi procesu digitalizacji, wskazały także na szereg jego specyficznych uwarunkowań ekonomicznych.

Bazą techniczną przedsięwzięć digitalizacyjnych Biblioteki Śląskiej była oddana do użytku w nowym gmachu w 1998 r. pracownia reprograficzna Działu Ochrony, Konserwacji i Archiwizacji, wyposażona w ówcześnie nowoczesny sprzęt do skanowania zbiorów: 2 skanery planetarne oraz cyfrowy aparat. Głównym sposobem wykorzystania sprzętu było świadczenie usług reprograficznych dla czytelników.

Digitalizacja wykonywana w ramach pracowni nie opierała się na żadnym programie cyfryzacji dokumentów do przełomu lat 2003/2004, kiedy postanowiono rozpocząć skanowanie pewnych zasobów Działu Zbiorów Specjalnych. Cyfrowe kopie co cenniejszych i częściej wykorzystywanych przez czytelników zabytkowych dokumentów, nagrywano na płyty CD i udostępniano w miejsce oryginałów w czytelni Działu Zbiorów Specjalnych, jednak bez publicznego udostępniania treści w sieci i możliwości skopiowania na nośniki zewnętrzne.

Pierwszym systemowym podejściem do kwestii digitalizacji była opracowana w 1999 r. przez ówczesny Dział Komputeryzacji Procesów Bibliotecznych koncepcja o roboczych nazwach *Wirtualna Biblioteka Śląska* i *Śląska Biblioteka Wirtualna*, która dzięki współpracy z jednym z polskich integratorów systemów informatycznych uzyskała postać funkcjonalnej dokumentacji systemu. W ramach projektu przewidywano zeskanowanie wybranego zabytkowego piśmiennictwa oraz dzieł współczesnych ze zbiorów Biblioteki Śląskiej. Dodatkowym celem było umożliwienie dostępu do tekstów dla osób niewidomych poprzez interfejs dźwiękowy oraz wydruki alfabetem Braile'a, w oparciu o uzyskany i odpowiednio przetworzony cyfrowy zasób.

Projekt *Śląska Biblioteka Wirtualna*, pomimo udanego początku prac, nie doczekał się realizacji. Po oszacowaniu kosztów projektu okazało się bowiem, że jego wartość (pamięci masowe, koszty licencji złożonego oprogramowania, specjalistycznych urządzeń) przekracza zdolności finansowe Biblioteki Śląskiej. Bardzo cennym zwyczajem okazało się dokładne dokumentowanie projektowych założeń, szacunków, typów kategorii zbiorów oraz nawet najdziwniejszych pomysłów, które następnie mogły być wykorzystane przy następnych projektach.

Na przełomie lat 2004/05, w związku z uruchomieniem środków funduszu Europejskiego Obszaru Gospodarczego, Biblioteka Śląska opracowała projekt pod nazwą *Śląska Internetowa Biblioteka Zbiorów Zabytkowych* (ŚIBZZ). Stanowił on całościowe podejście do kwestii digitalizacji zabytkowych zabiorów bibliotecznych. Projekt przewidywał utworzenie biblioteki cyfrowej udostępniającej zasób szerokiej publiczności oraz zdolnej do długotrwałego przechowywania zasobu cyfrowego w oparciu o zasoby Biblioteki Śląskiej.

Zgodnie z wnioskiem aplikacyjnym projekt miał przynieść rozwiązania w głównych obszarach problemowych: potencjalnej degradacji piśmiennictwa zabytkowego oraz ograniczonego dostępu do tych zbiorów dla naukowców i czytelników. Celem

technicznym projektu było utworzenie kompleksowego rozwiązania służącego do pozyskiwania, opracowania i prezentacji cyfrowej postaci zbiorów zabytkowych. Rozwiązanie miało być realizowane przez system złożony z czterech wielkoformatowych skanerów dziełowych i jednego stanowiska fotografii cyfrowej, infrastruktury sprzętowej zapewniającej przechowywanie i archiwizację zbiorów cyfrowych oraz oprogramowania specjalnego do opracowania bibliograficznego zbiorów, jak i do ich prezentacji w dedykowanym serwisie WWW. Przepływ skanów od operatorów skanerów do masowych pamięci długoterminowego składowania regulowało oprogramowanie typu *workflow,* łączące wyodrębnione stanowiska operacyjne systemu. Założono także możliwość rejestrowania i włączania do systemu zdalnych końcówek skanujących, na wypadek przyszłej rozbudowy systemu i współpracy z innymi instytucjami. Projekt wymagał dostosowania pomieszczeń: założono powstanie rozbudowanej pracowni digitalizacji, zapasową serwerownię z biblioteką taśmową w innym budynku Biblioteki do długoterminowego przechowywania zasobów, modernizację serwerowni głównej. Planowano zatrudnić jedenaście osób: czterech skanerzystów i fotografika, trzech korektorów-grafików, dwóch bibliotekarzy oraz informatyka-administratora systemu. Żmudnym zadaniem okazało się typowanie zbiorów do digitalizacji, której dokonali bibliotekarze Zbiorów Śląskich i Zbiorów Specjalnych, w kategoriach zbiorów wybranych do digitalizacji w ramach próby projektowej z 1999 r.

Projekt Śląskiej Internetowej Biblioteki Zbiorów Zabytkowych, o szacowanej wartości ponad 12 mln zł, złożono w listopadzie 2005 r. pod ocenę instytucji sterującej – Ministerstwa Kultury i Dziedzictwa Narodowego. Niestety, nie został on rekomendowany do dofinansowania, pomimo że uzyskały je projekty z niższą ilością punktów. Niemniej stanowił on gotową do wykorzystania koncepcję, która po niewielkiej aktualizacji mogła być wykorzystana w sytuacji dostępu do większych środków finansowych.

3. Ustanowienie i rozwój Śląskiej Biblioteki Cyfrowej (2006-2011)

Inicjatywa powołania Śląskiej Biblioteki Cyfrowej przypadła w 2006 r., w okresie intensywnego rozwoju polskiego systemu bibliotek cyfrowych. Funkcjonowało już szereg cyfrowych regionalnych jak Wielkopolska Biblioteka Cyfrowa, czy Kujawsko-Pomorska Biblioteka Cyfrowa oraz biblioteki instytucjonalne: Politechniki Wrocławskiej, Uniwersytetu Wrocławskiego, Politechniki Łódzkiej, Politechniki Warszawskiej. Ustanawiano je w różny sposób – przy tworzeniu bibliotek regionalnych wykorzystywano istniejące tradycje współpracy oraz bogate zasoby wiodącej instytucji. Cyfrowe biblioteki instytucjonalne tworzone były na bazie własnych zasobów instytucji z znaczącym uwzględnieniem treści współczesnych (prace naukowe, skrypty). Digitalizacja i budowa bibliotek cyfrowych była podejmowana przede wszystkim akademickie środowisko bibliotekarskie. Wspólnym mianownikiem tych inicjatyw było wykorzystywanie oprogramowania dLibra, rozwijanego przez Poznańskie Cen-

trum Superkomputerowo-Sieciowe. Niski koszt licencji i przynajmniej początkowo niewysokie wymagania sprzętowe dawały szansę na rozpoczęcie tworzenia biblioteki cyfrowej z wykorzystaniem bieżących środków instytucji, bez konieczności pozyskiwania znacznych środków zewnętrznych.

Wstępny pomysł utworzenia ŚBC na wstępnym etapie nie rozstrzygał o międzyinstytucjonalnym kształcie biblioteki cyfrowej, choć sygnalizował taką możliwość. Zasób o charakterze regionalnym w podstawowej wersji można było utworzyć w oparciu o zbiory Biblioteki Śląskiej, jednakże wówczas byłby on ograniczony do piśmiennictwa starszego, nieobjętego rygorami prawa autorskiego. W tym przypadku poza zasobem znalazłyby się nieposiadane przez Bibliotekę Śląską zbiory innych bibliotek regionu, wydawnictwa współczesne, w tym także treści wytwarzane przez ośrodki naukowe oraz szereg instytucji, dla których gromadzenie i udostępnianie piśmiennictwa nie było celem głównym. Istotne stało się więc opracowanie sposobu pozyskiwania treści zewnętrznych.

Najprostszą koncepcyjnie postacią takiego mechanizmu byłoby skanowanie zbiorów innych instytucji siłami i środkami Biblioteki Śląskiej oraz publikacja tych zbiorów w serwisie biblioteki cyfrowej. Rozwiązanie to jednak powodowałoby stosunkowo wysokie koszty pozyskiwania i tworzenia zasobu (organizacja, pracownia mobilna lub transport dokumentów, negocjowanie odrębnych porozumień z właścicielami zasobów). Powyższe uwarunkowania spowodowały, że ten sposób realizacji biblioteki cyfrowej uznano za trudny do praktycznego rozwinięcia na obecnym etapie przedsięwzięcia.

16 maja 2006 r. powstał w Bibliotece Śląskiej zespół do spraw Śląskiej Biblioteki Cyfrowej. W skład 8-osobowego zespołu weszli bibliotekarze działu Zbiorów Specjalnych, Zbiorów Śląskich, Działu Gromadzenia i Opracowania oraz informatycy Działu Informatycznego Biblioteki.

Istotnym problemem planowanej ŚBC był dostęp do treści współczesnych, szczególnie tych wydawanych i posiadanych przez ośrodki naukowe. Biblioteka Śląska, pełniąc funkcję biblioteki wojewódzkiej, była naturalnym partnerem dla bibliotek publicznych województwa śląskiego. Jednakże pole ściślejszej współpracy z bibliotekami akademickimi było mocno ograniczone, nie miało szerszych podstaw i nie było poparte znajomością specyfiki oczekiwań i możliwości ośrodków akademickich względem digitalizacji i sieciowej publikacji piśmiennictwa.

Z tego względu przedstawiono propozycję współpracy nad utworzeniem ŚBC Bibliotece Uniwersytetu Śląskiego, największej uczelni państwowej Górnego Śląska. Biblioteka UŚ, prócz zaawansowania w procesie komputeryzacji procesów bibliotecznych, posiadania doświadczeń we współpracy nad rozległymi systemami katalogowymi (NUKAT) i bazami danych (baza naukowych i fachowych polskich czasopism elektronicznych ARIANTA), podejmowała także prace z zakresu digitalizacji katalogów kartkowych i dokumentów, udostępniając je na stronie WWW.

Na roboczym spotkaniu przedstawicieli obu instytucji 6 czerwca 2006 r. ustalono, że nad projektem będzie współpracował międzyinstytucjonalny zespół pracowników Biblioteki Śląskiej oraz Biblioteki Uniwersytetu Śląskiego. Formalną podstawą współ-

pracy stało się *Porozumienie o współpracy przy utworzeniu Śląskiej Biblioteki Cyfrowej*, podpisane przez obie instytucje w dniu 20 lipca 2006, zawierające także deklarację wypracowania otwartej na inne instytucje formuły biblioteki cyfrowej.

W ramach wstępnych przygotowań rozważono ponownie kwestię oprogramowania biblioteki. Obok niskich kosztów licencji, rozwinięte funkcjonalności oraz sprawne wsparcie świadczone przez zespół programistów PCSS, były istotnymi argumentami za wykorzystaniem dLibry. Na krótko przed rozpoczęciem negocjacji zakupu dLibry zainstalowano i przetestowano system otwartego kodu Greenstone, jednakże stopień jego rozwoju oraz stosunkowo słabe wsparcie, przy braku zasobów kadrowych i finansowych na samodzielne dostosowanie i rozwój oprogramowania, przesądziły o zaprzestaniu dalszych testów tego systemu i zakupie przez Bibliotekę Śląską oprogramowania autorstwa PCSS.

W trakcie kilkumiesięcznej współpracy powstały dwie koncepcje funkcjonowania ŚBC. Koncepcja BUŚ proponowała utworzenie miał być realizowany w sposób scentralizowany. Nad doborem treści do digitalizacji i publikacji w bibliotece cyfrowej miała czuwać Rada Programowa, która decydowałaby o merytorycznej stronie zasobu, doborze materiału, jej rozbudowie, zakresie i zasięgu. Zakładano także umowny podział obszarów digitalizacji pomiędzy instytucje. Włączanie do ŚBC zbiorów innych bibliotek miało być poprzedzone ustaleniem profilu akceptowalnych treści.

Ostatecznie w toku dyskusji przyjęto bardziej liberalny model funkcjonowania ŚBC, zaproponowany przez Bibliotekę Śląską. Ze względu na terytorialne rozproszenie treści uznano, że najlepszym rozwiązaniem będzie przyznanie potencjalnym uczestnikom całkowitej autonomii w zakresie wyboru dzieł do digitalizacji oraz udostępnienie im możliwości samodzielnego, zdalnego publikowania treści w bibliotece. Podobnie za zbędną uznano instancję włączania do kooperacji nowych instytucji-uczestników za zgodą centralnego ciała, przyjmując, że wystarczającym probierzem intencji nowych członków ŚBC będzie podpisanie przez nich przygotowanego *Porozumienia o współtworzeniu ŚBC*. Dokument, ten o charakterze wielostronnej konwencji, wymagany ze względów formalno-prawnych (współpraca instytucji), uregulował takie obszary działania ŚBC jak misja, cele, obszar działania, strukturę organizacyjną, podział zadań, domeny internetowe (sbc.org.pl oraz digitalsilesia.eu) , utrzymanie infrastruktury i komunikacji między Uczestnikami.

Po analizie kosztów utrzymywania instalacji, przedstawiono także Uczestnikom następujące zasady współpracy:

I. Dla Uczestnika – ŚBC jest nieodpłatną usługą biblioteki cyfrowej.
II. Wszyscy publikują na tych samych zasadach.
III. Wspólny zasób tworzy wiele rozproszonych ośrodków.
IV. Nie ma centralnej kontroli i selekcji publikacji.
V. Łatwo zostać Uczestnikiem ŚBC, łatwo można zrezygnować.
VI. Cyfrowy zasób jest identyfikowany jako przynależny do Uczestnika.
VII. Sposób pozyskiwania, przygotowanie zasobu (przy zachowaniu standardów) i tempo jego publikowania jest autonomiczną decyzją każdego Uczestnika.

VIII. Każdy Uczestnik rozwiązuje kwestię praw autorskich (stwierdzenie wygaśnięcia lub umowa z autorem, wydawcą).

IX. Wszyscy zachowują prawa do swojego zasobu.

Wszystkie ustalenia zostały opisane w *Porozumieniu* [pełny tekst *Porozumienia* dostępny pod adresem: http://www.sbc.org.pl/dlibra/text?id=porozumienie] zaprezentowanym w dniu 24 października 2006 r. na zwołanym spotkaniu 19 instytucji. Po ustaleniu drobnych poprawek do tekstu Porozumienia w ciągu dwóch tygodni rozpoczął się proces przyjmowania nowych Uczestników ŚBC.

Zgodnie z innymi zapisami *Porozumienia* centralną platformę sprzętowo-programową utrzymuje koordynator organizacyjny ŚBC – Biblioteka Śląska. Początkowo był to sprzęt o niewielkiej wydajności, stopniowo rozbudowywany, w dalszych zaś latach wymieniony na serwerowy system wysokiej dostępności. Biblioteka Śląska obsługuje także formularz kontaktowy ŚBC, którego wiadomości dotyczące innych Uczestników lub całej biblioteki cyfrowej są przesyłane do właściwego adresata lub publikowane na forum ŚBC. Bieżącymi sprawami ŚBC zajmuje się wyłoniony przez Zebranie Uczestników 7-osobowy Zespół Koordynacyjny. W ramach ŚBC funkcjonuje również Kolegium Naukowe, formowane z przedstawicieli instytucji naukowych, tworzących bibliotekę cyfrową, którego zadaniem jest doradztwo i opiniowanie w zakresie publikacji treści naukowych, edukacyjnych i dydaktycznych. Zarówno Zespół Koordynacyjny jak i Kolegium Naukowe pracują w 3-letnim okresie kadencyjnym. Najwyższą 'władzą' ŚBC jest odbywane raz do roku Zebranie Uczestników, które większością głosów przedstawicieli każdej z instytucji uczestniczących rozstrzyga propozycje przedkładane przez Uczestników lub Zespół Koordynacyjny.

Ważną kwestią dla ustanowienia i rozwoju biblioteki cyfrowej stało się dostarczenie uczestnikom nowej wiedzy i umiejętności z zakresu tworzenia cyfrowych zasobów. Dokonało się to w drodze organizowanych szkoleń dla redaktorów i skanerzystów, prowadzonych przez zespół instytucji założycielskich. W pierwszym półroczu istnienia ŚBC odbyły się trzy takie szkolenia, obejmujące 47 osób z różnych instytucji. W późniejszym okresie szkolenia były i są kontynuowane w sposób ciągły, w ramach utworzonej Społecznej Pracowni Digitalizacji.

W celu wsparcia Uczestników utworzono także forum internetowe ŚBC, którego najistotniejszą jego częścią jest podforum *Informacje o ŚBC*. Są tam zawarte informacje dotyczące: instalacji oprogramowania aplikacji redaktora, ustanowienia kont redaktorskich, szkoleń, zasad korzystania ze Społecznej Pracowni Digitalizacji, czy uzyskania wsparcia. Zamieszczone są też tam m. in. wzory umów licencyjnych z autorami, wydawcami i właścicielami dzieł, a także słownik typów dokumentów oraz interpretacja formatu Dublin Core.

Jedną z pierwszych inicjatyw nowo powstałej ŚBC było przygotowanie nowych wersji językowych interfejsu biblioteki cyfrowej. Ponieważ spodziewano się przystępowania do ŚBC instytucji z Czech (co ostatecznie nie nastąpiło), jako pierwszą opracowano wersję czeską, a następnie niemiecką oraz francuską. Tłumaczenia wykonane przez pracowników Biblioteki Śląskiej, realizowane były w serwisie PCSS, a następnie

nieodpłatnie udostępnione innym bibliotekom cyfrowym w kraju, pod warunkiem wskazania autorów tłumaczeń.

Ważną aktywnością osób zaangażowanych w tworzenie ŚBC jest stała promocja przedsięwzięcia. W latach 2006-2011 członkowie Zespołu Koordynacyjnego przedstawili w ramach konferencji, spotkań i warsztatów ok. 50 prezentacji ŚBC i innych jej projektów w 26 miastach w Polsce i za granicą [prezentacje wygłoszone przez koordynatorów ŚBC od 2006 r. są dostępne pod adresem: http://www.slideshare.net/SlaskaBC/presentations]. Informacje o bibliotece cyfrowej zamieszczono w ogólnopolskich publikatorach oraz portalach internetowych. Tematyka digitalizacji oraz organizacji i funkcjonowania ŚBC została wprowadzona do programu nauczania Instytutu Bibliotekoznawstwa i Informacji Naukowej UŚ.

W roku 2007, po ustanowieniu Federacji Bibliotek Cyfrowych ŚBC została przyłączona do krajowego systemu bibliotek cyfrowych. Bibliotekarze cyfrowi z ŚBC wnieśli znaczny wkład w rozwój tego systemu, w postaci wspomnianych wersji językowych interfejsu, pomysłów mechanizmu koordynacji digitalizacji w kraju, standardów opracowania cyfrowych zasobów, usprawnień interfejsu użytkownika oraz promocji digitalizacji i form organizacyjnych bibliotek cyfrowych. Cztery instytucje współtworzące ŚBC są reprezentowane w Konsorcjum Polskie Biblioteki Cyfrowe, a 2 osoby brały udział w jego zespołach roboczych ds. metadanych oraz synonimów.

Od 2010 r. zasoby ŚBC są widoczne w Europeanie oraz w serwisach DART-Europe i ViFaOst (Wirtualna Biblioteka Europy Wschodniej).

4. Projekty Uczestników ŚBC

Forum i blog Biblioteka 2.0

W lutym i marcu 2007 r. staraniem pracowników Biblioteki Śląskiej tworzących ŚBC uruchomiono dwa ogólnodostępne serwisy internetowe: forum Biblioteka 2.0 [http://forum.biblioteka20.pl] oraz blog o takiej samej nazwie [http://blog.biblioteka20.pl]. Oba serwisy zostały poświęcone stosowaniu nowych technologii w bibliotekach, przy czym na forum rozwinięto przede wszystkim problematykę tworzenia bibliotek cyfrowych. Wśród tematów dyskusji pojawiły się takie kwestie jak sprzęt i oprogramowanie stosowane w procesach digitalizacji, analiza metadanych formatu Dublin Core, wątki o dobrych i złych praktykach w zakresie digitalizacji, a także informacje o ciekawych inicjatywach digitalizacyjnych, wartych uwagi serwisach internetowych (forach, blogach, portalach) oraz publikacjach dotyczących stosowania elektronicznych mediów. Na forum tym także dyskutowane zasady krajowej koordynacji tworzenia cyfrowych zasobów oraz podjęto kwestie prawno-autorskie. Jedno z podforów dedykowano komunikacji z użytkownikami, zachęcając ich do zgłaszania sugestii co do kształtu polskiego systemu bibliotek cyfrowych i propozycji digitalizacji konkretnych dzieł. Forum, dzięki otwartej formule umożliwiło komunikację

i wymianę doświadczeń aktywniejszym środowiskom tworzącym krajowy system bibliotek cyfrowych.

Społeczna Pracowania Digitalizacji ŚBC

W 2007 r. Biblioteka Śląska przystąpiła do prowadzonego przez Ministerstwo Kultury i Dziedzictwa Narodowego programu *Mecenat 2007* z wnioskiem projektowym pt. *Społeczna Pracownia Digitalizacji Śląskiej Biblioteki Cyfrowej przy Bibliotece Śląskiej – organizacja i wyposażenie*. Celem projektu było uruchomienie oraz wyposażenie pracowni digitalizacji, której działalność wspierałaby rozwój ŚBC, skupiającej wówczas 20 instytucji uczestniczących [o SPD więcej w: http://www.slideshare.net/SlaskaBC/spoeczna-pracownia-digitalizacji-bc]. W aplikacji projektowej podkreślono, że prócz celu materialnego, jakim jest cyfrowy zasób, ważnym aspektem projektu jest stymulowanie dóbr społecznych: edukacji medialnej i branżowej, integracji społecznej, podnoszenie kompetencji pracowników instytucji kultury i nauki oraz wspieranie rozwoju społecznego zaufania.

W ramach projektu SPD, pod koniec 2007 r., zakupiono 5 zestawów komputerowych ze skanerami i programami do obróbki dokumentów graficznych i tekstowych oraz wyposażenie meblowe. Opracowano regulamin korzystania z Pracowni, a do opieki nad nią skierowano dwóch etatowych pracowników Biblioteki. Jednocześnie podjęto kroki zmierzające do ustanowienia zespołu skanerzystów-wolontariuszy. Pierwsze próby nie były zachęcające, bowiem zgłaszający się zainteresowani, po przyuczeniu do tworzenia zasobów cyfrowych, rezygnowali ze współpracy po zeskanowaniu i opublikowaniu dokumentów interesujących ich osobiście. W celu powołania zespołu bardziej stałego, zaproponowano współpracę przy digitalizacji grupie seniorów – studentów Uniwersytetu Trzeciego Wieku. Pilotażowa grupa 5 osób, po przejściu szkolenia z podstaw obsługi komputera, usług internetowych oraz obsługi skanerów, pomyślnie rozpoczęła, pod okiem opiekunów, stałą współpracę z zespołem bibliotekarzy cyfrowych Biblioteki Śląskiej. Po prezentacji SPD i ŚBC przed ogólnym spotkaniu słuchaczy UTW na Uniwersytecie Śląskim, grupa powiększyła się do 14 osób, które przy nikłej rotacji i wysokim stopniu osobistego zaangażowania, są stałym zespołem tworzącym cyfrowe zasoby ŚBC. Dla seniorów SPD jest także miejscem spotkań, rozwijania własnych zainteresowań z wykorzystaniem stale dostępnego Internetu oraz międzypokoleniowej integracji.

Druga grupą wspierającą tworzenie zasobów ŚBC w ramach SPD są studenci bibliotekoznawstwa Uniwersytetu Śląskiego, których ścieżka edukacyjna obejmuje zagadnienia ochrony zbiorów. W ramach programowych staży mogą praktycznie zapoznać się z procesem digitalizacji i publikacji w bibliotece cyfrowej przy skanowaniu i obróbce zbiorów Biblioteki Śląskiej.

SPD wspiera też Uczestników ŚBC, a także, w miarę możliwości czasowych, inne instytucje pragnące rozpocząć programy digitalizacji. Pracownicy instytucji Uczestników korzystają z kompleksowych szkoleń przysposabiających do digitalizacji i kooperacji w ramach biblioteki cyfrowej. Uczestnicy także mogą skorzystać na miejscu

ze sprzętu i oprogramowania w celu skanowania własnych zbiorów lub przesłać przez sieć do Pracowni skany do obróbki przez opiekunów. Wszystkie usługi SPD są nieodpłatne.

W ciągu 3,5 roku funkcjonowania w ramach SPD wytworzono siłami wolontariuszy blisko 23% publikacji ŚBC (tj. ok. 300 tys. skanów), przeszkolono 40 osób z 18 instytucji, zapewniono staże i praktyki 34 uczniom i studentom, przetworzono blisko 100 tys. skanów innych uczestników ŚBC oraz udostępniono im sprzęt i oprogramowanie w wymiarze ponad 200 godzin.

Śląska Internetowa Biblioteka Zbiorów Zabytkowych (2008)

We wrześniu 2008 r. Biblioteka Śląska złożyła do Ministerska Kultury i Dziedzictwa Narodowego wniosek aplikacyjny dla projektu repozytorium zbiorów zabytkowych, który stanowił wznowienie projektu o takiej samej nazwie z roku 2005. Aplikację poprzedziła aktualizacja dokumentacji technicznej oraz czynności wyboru zasobu. Na nowo także określono cele projektu, uzgadniając je z wymogami Programu Operacyjnego Infrastruktura i Środowisko, działanie 11.1: *Ochrona dziedzictwa kulturowego o znaczeniu ponadregionalnym*.

Projekt, choć w głównych założeniach był identyczny z projektem niezrealizowanym w 2005 r. (pracownia digitalizacji, rozbudowa infrastruktury, oprogramowanie *workflow* na otwartej licencji, zatrudnienie zespołu), musiał brać pod uwagę funkcjonowanie Śląskiej Biblioteki Cyfrowej, która w swoim zasobie prezentowała już obiekty zabytkowe. Analiza funkcjonalności obu zamierzeń wskazała, że dzięki drobnym modyfikacjom pierwotnych założeń uda się, dzięki realizacji ŚIBZZ, utworzyć kompleksową bibliotekę cyfrową, realizującą funkcje długoterminowego zachowania cyfrowych zasobów, z wykorzystaniem już istniejącego, sprawnego mechanizm publikowania i udostępniania za pomocą aplikacji oprogramowania ŚBC – systemu dLibra. Dlatego postanowiono dokonać integracji, w drodze rezygnacji w projekcie ŚIBZZ z realizacji interfejsu dostępowego oraz utworzenia modułu integracyjnego pomiędzy dLibrą a oprogramowaniem *workflow*. Ponieważ *workflow* stanowi krytyczną, przechowującą zasób długoterminowy aplikację biblioteki cyfrowej, Biblioteka kontraktując oprogramowanie, zagwarantowała sobie licencyjną możliwość wglądu w kod źródłowy oraz prawo do jego modyfikacji w drodze prac własnych lub zleconych. Zapewnia to pełną kontrolę instytucji nad przechowywanymi danymi oraz uniezależnia ją od firm zewnętrznych – właścicieli zamkniętego kodu.

Ze środków projektu w postępowaniu konkursowym zatrudniono także 11 osób do obsługi poszczególnych etapów tworzenia zasobów i administracji systemem, których doświadczenia przy digitalizacji zróżnicowanych typów zbiorów pozwoli na zebranie znacznego kapitału kompetencji. Przeprowadzono także prace adaptacyjne pomieszczeń w celu utworzenia rozbudowanej pracowni digitalizacji, modernizacji serwerowni głównej oraz utworzenia zapasowej serwerowni z urządzeniami archiwizacyjnymi, zlokalizowanej w innym budynku. Materialnie w ramach projektu powstanie ok. 27 tysięcy cyfrowych publikacji, prezentujących inkunabuły, starodruki,

rękopisy, grafikę, kartografię, pocztówki, druki ulotne oraz czasopisma, co pozwoli na poszerzenie zasobu ŚBC do ok. 60 tys. publikacji (od stanu obecnego). Projekt ma być realizowany do końca roku 2012 [więcej o projekcie: http://sibzz.bs.katowice.pl].

Ochrona i konserwacja cieszyńskiego dziedzictwa piśmienniczego

W styczniu 2009 r. do ŚBC przystąpiła Książnica Cieszyńska, realizująca rozpoczęty niezależnie projekt *Ochrona i konserwacja cieszyńskiego dziedzictwa piśmienniczego*. Realizację rozpoczęto w 2007 r. siłami konsorcjum złożonego z Książnicy Cieszyńskiej (lidera) oraz 5 innych cieszyńskich instytucji [szczegółowy opis projektu: http://www.eog.kc-cieszyn.pl/index.php].

Projekt finansowany z funduszy Mechanizmu Finansowego Europejskiego Obszaru Gospodarczego oraz z dotacji pochodzącej z programu Promesa Ministra Kultury i Dziedzictwa Narodowego obejmował szereg działań kompleksowego zabezpieczenia oraz wprowadzenie do obiegu naukowego i kulturalnego zbiorów instytucji cieszyńskich. W ramach projektu wykonano także digitalizację ok. 570 000 stron dawnej prasy, rękopisów oraz starych druków, w tym nieco ponad 400 000 ze zbiorów samej Książnicy. Po przystąpieniu Książnicy do ŚBC prezentacyjne zasoby wykonane w ramach projektu zostały włączone do jej zasobu. Prócz zbiorów pozyskanych w ramach projektu, zasób ŚBC został także powiększony o treści ze wspomnianego projektu Książnicy Cieszyńskiej *Cieszyńska Biblioteka Wirtualna* (98 dokumentów, ok. 13 tys. skanów), realizowanego samodzielnie w latach 2003-2007.

RID (Rozszerzony Internetowy Dostęp)

Projekt Biblioteki Uniwersytetu Śląskiego, finansowany w ramach Regionalnego Programu Operacyjnego, Priorytet II *Społeczeństwo informacyjne*, Działanie 2.2. *Rozwój elektronicznych usług publicznych* został opracowany i rozpoczęty w 2008 r.

Jego celem było rozszerzenie dostępu oraz powiększenie poziomu niezawodności dostępu do Śląskiej Biblioteki Cyfrowej, zwłaszcza grupie osób niedowidzących i niewidzących. Celami szczegółowymi była poprawa funkcjonalności ŚBC poprzez rozszerzony mirroring zasobu ŚBC, dostosowanie zasobu do potrzeb osób niepełnosprawnych oraz digitalizacja zbiorów.

Zakresie zwiększenia funkcjonalności BBC projekt przewidywał utworzenie systemu mirorra ŚBC – synchronizowanego cyklicznie z zasobami publikowanymi w systemie podstawowym utrzymywanym przez Bibliotekę Śląską, utworzenie opartej na odrębnym systemie z oprogramowaniem dLibra Biblioteki Cyfrowej UŚ publikującej lokalnie cyfrowe zasoby chronione prawem autorskim. W ramach projektu zbudowano zaplecze serwerowe oparte na maszynach o nadmiarowej konfiguracji. W celu udostępnienia zbiorów cyfrowych osobom niewidzącym, przewidziano strumieniowy przekaz dźwięku na urządzenia użytkowników.

Projekt został zakończony, a w jego efekcie zasoby Śląskiej Biblioteki Cyfrowej stały się dostępne z dwóch systemów, a w razie problemów technicznych możliwe jest

udostępnianie zasobów ŚBC za pośrednictwem mirrora [więcej o projekcie: http://rid.bg.us.edu.pl].

Cyfrowy Beskidzki Świat

Po przeprowadzeniu aplikacji w marcu 2009 r., Książnica Beskidzka zrealizowała ze środków Ministra Kultury i Dziedzictwa Narodowego projekt *Cyfrowy Beskidzki Świat*, którego celem było powiększenie kolekcji cyfrowych. Projekt obejmował digitalizację dokumentów życia społecznego Podbeskidzia – regionaliów, zabezpieczenie zbiorów poprzez digitalizację i tworzenie cyfrowych kopii zabytków kultury piśmienniczej, popularyzowanie oraz zwiększenie dostępności zbiorów Książnicy Beskidzkiej (digitalizacja starodruków). W trakcie trwania projekt został rozszerzony o digitalizację regionalnego tygodnika *Kronika Beskidzka* za lata 1956-1991, poprzedzone uzyskaniem licencji od wydawcy. Większość wytworzonych cyfrowych publikacji zostało opublikowanych w ŚBC.

Digitalizacja zasobów dziedzictwa kulturowego – kronik szkolnych

We wrześniu 2008 r. jedna z instytucji uczestniczących ŚBC – Związek Górnośląski, jako partner, wraz z liderem Urzędem Miasta Katowice, złożył aplikację do Regionalnego Programu Operacyjnego o dofinansowanie w ramach działania 4.1. *Infrastruktura kultury* projektu dotyczącego zachowania w formie cyfrowej kronik szkolnych przechowywanych w katowickich szkołach. Projekt obejmował digitalizację dokumentów, obróbkę graficzną, tłumaczenie tekstów z języka niemieckiego, przygotowanie prezentacyjnych wersji publikacji cyfrowej (DjVu) oraz umieszczenie ich w ŚBC. Kroniki przechowywane były w szkołach i gimnazjach oraz w Muzeum Miasta Katowice – w 16 lokalizacjach. Zasób piśmienniczy stanowiło 29 ksiąg o łącznej objętości 6060 stron, z czego 1556 w języku niemieckim. Najstarsze z kronik sięgały roku 1810 [więcej o projekcie: http://www.slideshare.net/SlaskaBC/digitalizacja-zasobow-dziedzictwa-kulturowego-kronik-szkolnych].

Po zakończeniu projektu jego techniczna baza pozostała w wyposażeniu Związku Górnośląskiego w celu kontynuacji prac nad kronikami nieobjętymi projektem oraz digitalizacji innych zasobów piśmiennictwa.

Digitalizacja kronik cechowych Muzeum Miejskiego w Żywcu

W okresie od czerwca 2009 do czerwca 2010 r. Biblioteka Śląska zrealizowała pilotażowy projekt digitalizacji 'mobilnej', wykonywanej poza pracowniami BŚ. We współpracy z Miejskim w Żywcu dokonano digitalizacji przechowywanych przez muzeum zabytkowych ksiąg cechowych, rękopisów dokumentujących powstanie, organizację, skład osobowy i dzieje cechów rzemieślniczych, działających na terenie Żywca. Ze względu na brak celowych funduszy, w realizacji wykorzystano sprzęt kompu-

terowy z SPD, a do przeprowadzenia digitalizacji oddelegowano jednego z opiekunów SPD oraz skanerzystę z Działu Ochrony, Konserwacji i Archiwizacji Biblioteki Śląskiej. W celu obniżenia wzajemnych kosztów przedsięwzięcia Muzeum zapewniło ekipie skanującej warunki pobytu, natomiast Biblioteka Śląska – przekazała Muzeum archiwalne skany rękopisów w formacie TIFF do dowolnego wykorzystania. W trakcie trzech sesji wyjazdowych zespół mobilny wykonał ponad 7000 skanów 51 egzemplarzy rękopisów, które po przetworzeniu trafiły do zasobu ŚBC [więcej o projekcie: http://www.slideshare.net/SlaskaBC/mobilna-pracownia-digitalizacji].

Współpraca z Fundacją Nowoczesna Polska na rzecz projektu „Wolne Lektury"

We wrześniu 2010 r. Biblioteka Śląska podpisała porozumienie z FNP o współpracy w zakresie przygotowania skanów dla portalu Wolne Lektury. W ramach współpracy seniorzy z SPD zeskanowali kilkadziesiąt istotnych dla polskiej i światowej kultury utworów piśmienniczych ze zbiorów Biblioteki Śląskiej, które po dokładnym rozpoznaniu tekstu przekazano do publikacji w serwisie Wolnych Lektur.

5. Podsumowanie

Wynikiem 5 lat pracy nad tworzeniem regionalnego zasobu cyfrowego jest uformowanie i funkcjonowanie biblioteki cyfrowej, stanowiącej międzyinstytucjonalną, elastyczną organizację o wysokim stopniu wirtualizacji, w oparciu o zasadę współdzielenia zasobów. Zapewnia ona stosunkowo swobodny przepływ wiedzy, limitowany jedynie zdolnością uczestników do jej absorpcji. Stawia to niekiedy przed jej uczestnikami pionierskie wyzwania w zakresie komunikacji, koordynacji działań, a także mechanizmów międzyinstytucjonalnego transferu kompetencji, wsparcia i dobrych praktyk. ŚBC tworzą ludzie, którzy najczęściej znają z korespondencji mailowej lub wypowiedzi na forum, a ich działania często mają dorywczy, 'pulsujący' charakter. Trudno także oszacować sprzętowy lub kadrowy potencjał ŚBC, bowiem zmienia się on z dnia na dzień – w zależności od działań podejmowanych przez uczestników.

Bibliotekę cyfrową regionu śląskiego charakteryzuje klastrowa struktura [bliższa charakterystyka klastra ŚBC: http://www.slideshare.net/SlaskaBC/tworzenie-biblioteki-cyfrowej-w-modelu-klastrowym-dowiadczenia-bc], dzięki której wiele instytucji, pomimo istnienia pól konkurencji, może realizować wspólne przedsięwzięcia, dysponując przy tym pełną autonomią w zakresie stopnia zaangażowania swoich zasobów. Autonomia działań sprzyja także elastyczności działania i wzbogacaniu wspólnego przedsięwzięcia o nowe treści, bowiem każdy z uczestników może niezależnie od innych wypracowywać własne formy docierania do dokumentów (współpraca z kolejnymi instytucjami, osobami prywatnymi) oraz realizować własne projekty. Przynosi to także specyficzną, 'fraktalną' dynamiką działania i rozwoju przedsięwzięcia, które tym samym ma wysokie zdolności adaptacyjne: wykorzystuje okazje dotarcia do

nowych, niszowych dokumentów, dobrowolnie angażuje kolejne instytucje i osoby prywatne, jednocześnie promując je oraz ich wkład, umożliwia niezależne lub kooperacyjne pozyskiwanie środków – przekraczając granice instytucjonalne i resortowe.

Uzyskana dzięki tej organizacyjnej formie synergia działań, poparta intensywną promocją, ułatwia 'budowanie marki' biblioteki cyfrowej, co z kolei stanowi znaczną pomoc przy ubieganiu się o budżetowe lub projektowe zasoby na jej dalszy rozwój.

Należy zaznaczyć, że tego rodzaju organizacja stawia specyficzne – bo prócz branżowych, także psychospołeczne – wymagania aktywnie rozwijającym ją ludziom, które są stosunkowo łatwo sprawdzalne w codziennej współpracy. Te najbardziej pożądane to: przedsiębiorczość, techniczna i społeczna innowacyjność, ponadinstytucjonalne i bezosobowe myślenie zadaniowe, nakierowanie na działania projektowe oraz wysoka zdolność do negocjacyjnej kooperacji i wielokierunkowej komunikacji. Dynamika środowiska nowych mediów, przy ciągle pojawiających się okazjach rozwoju (ciekawy, wart wykonania pomysł, realizacja trudnych projektów, czy zwykłe, wymagające szybkiej interwencji problemy sprzętowo-programowe) wymagają także zdolności do działania w warunkach 'wersji beta': stałej zmienności, poprawek i ulepszeń. Często kłoci się to z mentalnym skłonnościami przeciętnego pracownika tradycyjnej, statycznej instytucji.

Podstawą trwałości i bieżącego funkcjonalnego utrzymania ŚBC jest wkład koordynatora organizacyjnego – Biblioteki Śląskiej, udostępniającego platformę sprzętową i programową oraz ośrodki dystrybucji kompetencji (forum ŚBC, Społeczna Pracownia Digitalizacji), co umożliwia stopniowe, lecz stałe podnoszenie umiejętności uczestników w zakresie digitalizacji i, w efekcie, przyrost różnorodnego zasobu. Nieodpłatne udostępnienie tych zasobów oraz brak wstępnych wymogów włączenia do przedsięwzięcia, obniżyło próg wejścia do ŚBC – obecnie wspólny zasób tworzy 50 instytucji województwa śląskiego i opolskiego, w tym biblioteki publiczne (od wojewódzkiej po miejsko-gminną), biblioteki akademickie szkół państwowych i prywatnych, niebiblioteczne instytucje kultury, stowarzyszenie, wydawnictwa oraz związek wyznaniowy [pełna lista uczestników dostępna pod adresem: http://www.sbc.org.pl/dlibra/text?id=participants]. W zasobie ŚBC można także znaleźć dokumenty z kolekcji prywatnych, pozyskane staraniem Uczestników.

Rozwój ŚBC jako przedsięwzięcia niskobudżetowego i zaprojektowanego na obniżanie kosztów współpracy jest możliwy dzięki realizacji przez jej uczestników innowacyjnych projektów, które pozwalają na jakościowe skoki w zakresie poszerzenia treści, poprawy infrastruktury lub zatrudnienia specjalistów. Uczestnicy w ciągu 5 lat istnienia ŚBC zrealizowali lub realizują projekty na łączna kwotę ponad 10 milionów złotych.

Chociaż bieżące funkcjonowanie ŚBC, dzięki mocnemu instytucjonalnemu ufundowaniu jest zapewnione, należy wspomnieć także o słabszych stronach takiego rozwiązania, wynikających z jego istoty (luźny charakter kooperacji) lub specyfiki środowiska bibliotekarskiego i ogólniej, instytucji kultury. Oparcie inicjatywy na samoorganizacji autonomicznych uczestników skutkuje zróżnicowaniem zaan-

gażowania we współpracę, znacznymi (jak na możliwości) dysproporcjami w liczbie tworzonych zasobów, różnicami w wyposażeniu, jakości opracowania zasobów oraz często długotrwałym uruchamianiem procesu digitalizacji, szczególnie tam, gdzie digitalizacja jest aktywnością marginalną. Niedostatek środków bieżących na działalność instytucji kultury oraz brak możliwości rozwoju dla zdolnych, innowacyjnych pracowników powoduje, że traktują oni te miejsca pracy jako krótkotrwały wstęp do 'prawdziwej' kariery zawodowej. Skutkuje to znaczną rotacją kadr i utratę kapitału kompetencji, która także dotyczy osób pracujących w instytucjach uczestników na rzecz ŚBC. Konieczne zatem jest stałe odnawianie umiejętności, bez gwarancji, że istotnie zmieni to stopień zaangażowania uczestnika we współpracę.

Niemniej Śląska Biblioteka Cyfrowa, utworzona w warunkach ograniczonych środków, jest pomyślnie rozwijającą się inicjatywą. Przy oczekiwanej ogólnej poprawie cywilizacyjnego poziomu kraju, szczególnie w sferze finansowania kultury, w dalszym ciągu będzie mogła czerpać z wiedzy, doświadczeń i umiejętności osób i instytucji zaangażowanych w jej tworzenie, wnosząc jednocześnie istotne dla nowoczesnych społeczeństw dobra wolnego dostępu do kulturowych treści oraz stymulacji społecznego kapitału.

Jagiellońska Biblioteka Cyfrowa – opis przypadku

Łukasz Mesek

W tekście przedstawiono genezę, prace przygotowawcze i wdrażające oraz codzienną praktykę w Jagiellońskiej Bibliotece Cyfrowej. Zaprezentowane procedury czy listy zakupów pozwalają na ocenę sytuacji i przygotowanie się do przedsięwzięcia okiem pragmatyka, a miesięczne zestawienia statystyczne pomogą na szacunkowe określenie ilości dokumentów, zabezpieczonych i przetworzonych terabajtów danych oraz udostępnionych w wolnym Internecie cyfrowych reedycji dokumentów w większości wchodzących w skład domeny publicznej.

Geneza projektu

Powstanie Jagiellońskiej Biblioteki Cyfrowej (JBC) jest powiązane z wynikami przeprowadzonych w latach 2002-2003 w Bibliotece Jagiellońskiej badań oceniających stan zachowania całego księgozbioru. Około 13% badanych książek i czasopism zostało uznanych za obiekty w bardzo złym stanie, co uzasadniało konieczność wyłączenia ich z udostępniania. Rozwiązaniem tego problemu była digitalizacja najbardziej zniszczonych zasobów i ich publiczne udostępnienie w postaci kopii cyfrowych. Zadanie to wpisywało się w ogłoszony w 2007 r. *Program Rozwoju Uniwersytetu Jagiellońskiego*, który do najważniejszych przedsięwzięć rozwojowych zaliczył konieczność szybkiego wprowadzenia nowych rozwiązań z zakresu technik informatycznych, pozwalających na łączenie tradycji uczelni z nowoczesnością na najwyższym poziomie (priorytet 5 Programu *Informatyzacja uczelni* powiązanie z zadaniem 5.03: *Rozwijanie systemów dostępu do Internetu i baz danych przez pracowników, doktorantów i studentów*) przewidziany do realizacji w ciągu 5 następnych lat. Kolejnym impulsem było powstanie Federacji Bibliotek Cyfrowych (FBC) w dniu 22 czerwca 2007 roku. FBC

zrzesza biblioteki cyfrowe w Polsce, udostępniające swoje publikacje na stworzonej w Poznańskim Centrum Superkomputerowo-Sieciowym (PCSS) oprogramowaniu dLibra, które w ośrodkach akademickich zaczęły powstawać w latach 2002-2007, a pierwszą z nich jest powstała w 2002 r. Wielkopolska Biblioteka Cyfrowa. Bliższe zapoznanie się z projektem FBC i porównanie systemu organizacji repozytoriów cyfrowych na świecie, było czynnikiem przekonującym do włączenia się do współpracy w ramach jednej ogólnonarodowej platformy, skupiającej z każdym rokiem coraz większą liczbę bibliotek. Szczególnie idea agregatora danych, oferującego wspólne narzędzie wyszukiwawcze, pozwalające na koordynację prac na poziomie ogólnopolskim wpisywało się w program rozwoju Biblioteki Jagiellońskiej czynnie uczestniczącej w tworzeniu wspólnej bazy opisów bibliograficznych w ramach Narodowego Uniwersalnego Katalogu Centralnego (NUKAT).

1. Przygotowania projektowe

W ramach kompleksowego zabezpieczenia zbiorów w 2008 r. powstał plan prac nad stworzeniem JBC obejmujący:

1. wyszukanie odpowiedniego źródła finansowania dla projektu na dużą skalę, w którym znalazłyby się pieniądze na doposażenie pracowni Foto w Oddziale Zabezpieczenia Zbiorów Biblioteki Jagiellońskiej, jak również na wyposażenie Oddziału Komputeryzacji i Oddziału Dokumentów Audiowizualnych w odpowiedni sprzęt komputerowy, niezbędny do utrzymania i rozwijania JBC,
2. opracowanie koncepcji działania JBC tak, aby zabezpieczane i publikowane dokumenty dostosowane były do standardów na najwyższym światowym poziomie,
3. zapoznanie się z pracą innych bibliotek cyfrowych w Polsce od strony organizacji pracy, używanego sprzętu i oprogramowania oraz zaplecza kadrowego.

Zapotrzebowanie na dobrej jakości sprzęt do digitalizacji jak i komputery do obróbki kopii cyfrowych i ich publikacji online, a także zabezpieczenie kopii cyfrowych w postaci obrazów pozyskiwanych bezpośrednio ze skanerów, czyli plików master w formacie TIFF (Tagged Image File Format), wymagało znacznych inwestycji. Takie możliwości finansowe zapewniał tylko program współfinansowany ze środków Unii Europejskiej. Jedynym poważnie rozważanym i potencjalnie pewnym miejscem do aplikowania o dofinansowanie było któreś z ministerstw, dysponujących funduszami europejskimi. Taką możliwość stwarzał ogłoszony przez Ministerstwo Kultury i Dziedzictwa Narodowego nabór projektów w ramach Programu Operacyjnego Infrastruktura i Środowisko na lata 2007-2013, Priorytet XI. *Kultura i Dziedzictwo Kulturowe* Działanie 11.1. *Ochrona i zachowanie dziedzictwa kulturowego o znaczeniu ponadregionalnym* ogłoszony w 2008 r.

W przygotowaniu do tego projektu została opracowana koncepcja działania JBC, która obejmowała następujące zadania:

1. Typowanie, czyli wyznaczanie publikacji do skanowania.
2. Katalogowanie formalne i rzeczowe skanowanych publikacji w Komputerowym Katalogu Zbiorów Bibliotek Uniwersytetu Jagiellońskiego (KKZBUJ)

w formacie MARC21 i umieszczanie opisów bibliograficznych w ogólnopolskiej bazie NUKAT.

3. Przygotowanie do skanowania i mikrofilmowania, polegające na wykonaniu napraw i niezbędnych zabiegów konserwatorskich oraz zabezpieczających w pracowni konserwacji i introligatorni Oddziału Zabezpieczenia Zbiorów (OZZ).
4. Skanowanie wytypowanych publikacji.
5. Archiwizacja kopii cyfrowych, czyli zabezpieczenie kopii archiwalnych plików TIFF na taśmach magnetycznych przechowywanych w bibliotece taśmowej.
6. Archiwizacja w postaci kopii mikrofilmowych zeskanowanych dokumentów.
7. Konwersja obrazów cyfrowych w postaci plików TIFF do formatu plików publikacyjnych (DjVu).
8. Katalogowanie w programie dLibra polegające na konwersji opisów bibliograficznych wprowadzonych do KKZBUJ w formacie MARC21 do formatu Dublin Core, obsługiwanego przez oprogramowanie dLibra.
9. Publikacja polegająca na dodawaniu przekonwertowanych do formatu DjVu kopii cyfrowych do bazy utrzymywanej w oprogramowaniu dLibra.
10. Zabezpieczenie oryginalnych dokumentów (w przypadku czasopism, ułożenie ich w pudłach z tektury bezkwasowej w pozycji poziomej).
11. Zabezpieczenie mikrofilmów w magazynie podręcznym Oddziału Dokumentów Audiowizualnych (ODA), nadanie im sygnatur i kodów kreskowych oraz stworzenie w KKZBUJ rekordów zasobu i rekordów egzemplarzy podpiętych do rekordu bibliograficznego.

W celu realizacji tych zadań w Oddziale Komputeryzacji Biblioteki Jagiellońskiej powstał System Obsługi Jagiellońskiej Biblioteki Cyfrowej (SOJBC). W systemie tym nanoszone są informacje o każdym zakończonym etapie zabezpieczenia dokumentu, co pozwala na kontrolę i komunikację pomiędzy różnymi oddziałami zaangażowanymi do tak dużego przedsięwzięcia.

Uczestnictwo w organizowanych przez Poznańskie Centrum Superkomputerowo-Sieciowe warsztatach dotyczących wykorzystania oprogramowania dLibra do tworzenia bibliotek cyfrowych, a także uruchomienie Federacji Bibliotek Cyfrowych, skupiających głównie biblioteki cyfrowe działające na bazie tego oprogramowania skierowało działania organizatorów JBC w kierunku implementacji stosowanych i rozwijanych na gruncie polskim rozwiązań. Należy tu nadmienić, że pod względem unifikacji oprogramowania i systemów tworzonych do udostępniania i przeszukiwania zasobów bibliotek cyfrowych rozwiązania polskie są chlubnym ewenementem w skali światowej. Nigdzie na świecie (oprócz projektu Europeana) nie powstało jeszcze zintegrowane narzędzie do przeszukiwania zasobów cyfrowych na skalę kraju.

Wizyta w Bibliotece Uniwersytetu Adama Mickiewicza w Poznaniu oraz w Bibliotece Politechniki Poznańskiej, które to biblioteki tworzą główny trzon największej w Polsce Wielkopolskiej Biblioteki Cyfrowej przekonały o słuszności podążania przetartym szlakiem i uczestniczenia w ogólnopolskiej inicjatywie. Dodatkowe doświadczenia zostały wyniesione po zapoznaniu się z działaniem Narodowej Biblioteki

Cyfrowej w Pradze, a także przy przeglądaniu zasobów najbardziej reprezentatywnych bibliotek cyfrowych na świecie w tym Biblioteki Kongresu, Biblioteki Narodowej w Londynie i w Paryżu. Dlatego zdecydowano się na zakup i implementację oprogramowania dLibra w wersji 4.0 już w 2008 r., kiedy prace nad aplikowaniem o środki z MKiDN dopiero się rozpoczęły. Pozwoliło to na zapoznanie się z możliwościami oprogramowania i prace przygotowujące do publikacji dokumentów dotychczasowo zabezpieczonych w Bibliotece Jagiellońskiej przez OZZ.

2. Dobór dokumentów

Do digitalizacji w projekcie zostały wytypowane jedynie czasopisma. Podyktowane to było katastrofalnym stanem zachowania niektórych tytułów, co wynika ze słabej kondycji papieru i ich ogromnej poczytności. Udostępnianie czytelnikom oryginałów aż do końca XX wieku, poczyniło nieodwracalne spustoszenie, do tego stopnia, że niektóre woluminy nie są już możliwe do skanowania. Ilość tytułów czasopism przechowywanych w Bibliotece Jagiellońskiej jest ogromna, a więc potrzebne było dodatkowe kryterium doboru materiału. Wyselekcjonowano tytuły będące unikatami na skalę kraju, a ponadto zostało dobrane 10 czasopism z regionu Galicji cieszących się największą poczytnością. Wyróżnia się wśród nich Gazeta Lwowska, dziennik ukazujący się nieprzerwanie przez 128 lat, którego niemal pełnym zasobem może poszczycić się Biblioteka Jagiellońska. W projekcie znalazły się ponadto wszystkie zdigitalizowane do sierpnia 2008 r. dokumenty, przechowywane na dyskach optycznych. Podyktowane to zostało troską o zabezpieczenie zbiorów, których przyszłość w tej postaci jest niepewna. Według specjalistów trwałość informacji zachowanych na nośnikach optycznych, szczególnie, jeśli są one wypalane a nie tłoczone, wynosi 5 lat. Po tym czasie informacja musi zostać przegrana na nowy nośnik. Dla porównania trwałość informacji przechowywanych na dyskach twardych to ok. 20 lat, na taśmach magnetycznych w bibliotekach taśmowych 30-40 lat, a na mikrofilmie 125 lat [*Digital imaging guidelines* http://www.lva.virginia.gov/agencies/records/electronic/digital_imaging.pdf]. Ten ostatni wynik zadecydował o systemie kompleksowego zachowania digitalizowanego materiału w postaci mikrofilmów. Tym sposobem projektowi JBC wyznaczono najwyższy z możliwych statusów zabezpieczenia zbiorów i obejmuje on następujące ich postaci:

a) konserwacja i zabezpieczenie oryginału w pudłach z tektury bezkwasowej,
b) archiwizowanie kopii cyfrowej w postaci pliku TIFF na taśmie magnetycznej,
c) konwersja do publikowalnego pliku DjVu, przechowywanego w programie dLibra,
d) zabezpieczenie w postaci kopii negatywowej na mikrofilmie.

Ogólna liczba dokumentów w projekcie obejmuje:

- 239 tytułów czasopism,
- 221 tytułów książek,
- 165 jednostek zbiorów graficznych,

- 33 mapy,
- 292 jednostki zbiorów muzycznych,
- 217 jednostek dokumentów życia społecznego
- 695 rękopisów,
- 545 starych druków.

W ramach projektu *Bezpieczne i kompleksowe udostępnienie zasobów cyfrowych Uniwersytetu Jagiellońskiego w sieci Internet. Etap I*, który powstał niejako przy okazji tworzenia projektu Jagiellońska Biblioteka Cyfrowa, zostało przeznaczonych do zabezpieczenia i powszechnego udostępnienia w JBC kolejne 100 tytułów czasopism, z czego 42 tytuły, należą do jedynych ocalałych w Polsce.

3. Dobór sprzętu

Aby spełnić wszystkie postawione w rozdziale 2 warunki należało zapewnić odpowiedni sprzęt dla pracowni Foto Biblioteki Jagiellońskiej. Ogromne znaczenie miał dobór sprzętu o jak najlepszych parametrach. Dlatego w planach zakupów znalazły się skanery z najwyższej półki. Rozważane były produkty różnych firm z czego postawione kryteria spełniały urządzenia firm Zeutschel oraz Canon. Aby zapewnić na przyszłość sprzęt o parametrach umożliwiających digitalizację większości zbiorów Biblioteki Jagiellońskiej do zakupu zostały wytypowane jednostki skanujące dokumenty w różnych formatach. Dodatkowo zaplanowany do zakupu został konwerter do tworzenia mikrofilmów, pozwalający na dodatkowe zabezpieczenie materiału w postaci negatywowych kopii mikrofilmowych.

Kolejnym aspektem była kwestia przechowywania wytworzonych kopii cyfrowych zbiorów charakteryzujących się ogromnymi rozmiarami. Dla przykładu jedna strona zeskanowanej w formacie TIFF Gazety Lwowskiej (od 1850 ukazującej się w formacie V wysokość grzbietu 40 cm, a od roku 1870 wysokość grzbietu 47 cm) to średnio 50 MB. Waga jednego woluminu zdigitalizowanej gazety wynosi ok. 50 GB, co przy 300 woluminach daje 15 TB danych. W projekcie do zabezpieczenia zostało przewidzianych 1220 woluminów gazet, co szacunkowo daje wielkość 61 TB. Przechowywanie danych w takiej ilości na macierzy dyskowej, z ekonomicznego punktu widzenia, jest zupełnie nie opłacalne, szczególnie, że przyrost danych w postaci zabezpieczonych cyfrowo obiektów zwiększa się i będzie się zwiększał. Dlatego zaprojektowano rozwiązanie, w którym zabezpieczone kopie cyfrowe są przechowywane w bibliotece taśmowej. Pozwala to na stały dostęp do danych, a jednocześnie, przy utworzeniu dodatkowej kopii cyfrowej, przechowywanie jej ze względów bezpieczeństwa poza Biblioteką Jagiellońską. Prowadzi to do zaoszczędzenia środków finansowych przy zakupie macierzy dyskowej o mniejszej pojemności i skierowania ich na zakup taśm, na których kopie cyfrowe są przechowywane. Macierz robocza o pojemności 18 TB pozwala na bezawaryjną pracę bieżącą, polegającą na obróbce graficznej obrazów cyfrowych, ich konwersji do formatu, publikowalnego (DjVu) oraz przechowywanie tymczasowych plików przeznaczonych do stworzenia zabezpieczającej kopii mikro-

filmowej. Jest to rozwiązanie optymalne, praktykowane w bibliotekach cyfrowych zarówno w Polsce jak i na świecie.

Do pracy przy obróbce graficznej uzyskanych kopii cyfrowych wymagany był sprzęt pozwalający na wydajną i optymalną pracę. W ramach projektu zakupionych zostało 5 stacji roboczych HP Z400 o następujących parametrach:

- procesor Intel Xeon /Quad-Core/ W3550 3,06Ghz,
- pamięć RAM 3x2GB DDR3 ECC 1333Mhz,
- dysk twardy 1 TB.

W ramach przygotowania JBC do sprawnego działania pod planowaną do zakupu wersją 5 oprogramowania dLibra zakupione zostały trzy serwery Serwer stelażowy HP ProLiant DL380 G7 Special, z których każdy ma przydzieloną odpowiednią funkcję: indeksowania, trybu dostępu i przechowywania danych.

Jeszcze w fazie projektowej, w 2008 r. zostało zakupione oprogramowanie Document Express Enterprise pozwalające na wsadowe konwertowanie dużej ilości plików w formacie TIFF do niewielkich wagowo plików DjVu. Wybór tego oprogramowania i zastosowanie formatu DjVu podyktowany był również doświadczeniami zdobytymi w innych bibliotekach cyfrowych, ale głównie możliwościami, jakie format ten oferuje. Główną zaletą tego formatu jest ponad stukrotnie mniejsza przestrzeń (waga pliku) w porównaniu z plikiem otrzymywanym bezpośrednio ze skanera, bez utraty jakości przedstawianego obrazu. Dodatkowo oprogramowanie pozwala na jednoczesne z konwersją wykonanie zadania OCR (Optical Text Recognition), które w skrócie sprowadza się do rozpoznania tekstu przedstawionego na obrazie, możliwości jego przeszukiwania i edytowania. Dotyczy to oczywiście tylko tekstu drukowanego, ale co ważne obejmującego większość europejskich języków, włącznie z rozpoznawaniem cyrylicy. Silnik OCR ReadIRIS w tym oprogramowaniu nie jest najlepszym dostępnym na świecie silnikiem, co odzwierciedla się w ilości błędów, czyli nierozpoznanych liter, ale zaimplementowanie silnika AABBYY oferowanego w oprogramowaniu Fine Reader przedstawiało zbyt wiele trudności przy konwersji do formatu DjVu. Przeprowadzone testy[1] pokazały, że stosunek poprawnie rozpoznanych wyrazów w zaznaczonym tekście przez silnik ReadIRIS i silnik AABBYY kształtuje się odpowiednio na poziomie 82% i 89%, co zadecydowało o nie inwestowaniu środków i czasu w uzyskanie jak najlepszego wyniku OCR. Wskazując producentom niszę, jaką są polskie biblioteki cyfrowe używające głównie formatu DjVu do publikacji kopii cyfrowych swoich zasobów, przygotowali oni wersję 11 programu Fine Reader pozwalającą na zastosowanie silnika AABBYY w formacie DjVu. Silnik ten zacznie działać na indywidualnych stanowiskach bibliotekarzy cyfrowych na początku 2012 r.

[1] Testy przeprowadzone przez pracowników Oddziału Komputeryzacji i Oddziału Dokumentów Audiowizualnych Biblioteki Jagiellońskiej na próbce z dwóch czasopism: *Goniec Krakowski* i *Podgórzanin*

4. Opracowanie procesu technologicznego

Proces technologiczny został opracowany na podstawie zadań jakie zostały wyznaczone w koncepcji działania JBC i przedstawia się następująco:

1. Typowanie
   Wyznaczanie publikacji do skanowania.
   Ludzie: pracownicy odpowiednich oddziałów.
   Sprzęt: PC-ty pracowników, serwer SOJBC.
   Wyjście: lista publikacji do skanowania.
2. Katalogowanie w formacie MARC 21
   Katalogowanie skanowanych publikacji w VIRTUI.
   Ludzie: osoby katalogujące dla JBC.
   Sprzęt: PC-ty pracowników, serwer SOJBC, serwer VIRTUI.
   Wejście: lista publikacji do skanowania.
   Wyjście: lista publikacja skatalogowanych w VIRTUI.
3. Przygotowanie oryginałów do skanowania
   Wykonanie napraw i zabiegów konserwatorskich oraz zabezpieczających.
   Ludzie: pracownicy konserwacji i introligatorni Oddziału Zabezpieczenia Zbiorów
   Sprzęt: PC-ty pracowników, serwer SOJBC.
   Wejście: lista publikacji do skanowania.
   Wyjście: lista publikacji po zabiegach konserwatorskich na serwerze SOJBC.
4. Skanowanie
   Skanowanie wytypowanych publikacji.
   Ludzie: osoby skanujące dla JBC.
   Sprzęt: skanery, PC-ty skanujących, serwer SOJBC.
   Wejście: lista publikacji do skanowania.
   Wyjście: lista publikacji zeskanowanych; pliki TIFF na serwerze SOJBC.
5. Archiwizacja kopii cyfrowych
   Tworzenie kopii archiwalnych plików TIFF.
   Ludzie: administratorzy serwera SOJBC.
   Sprzęt: serwer SODJBC, streamer lub macierz dyskowa.
   Wejście: lista publikacji zeskanowanych; pliki TIFF na serwerze SOJBC.
   Wyjście: baza informująca, gdzie można znaleźć pliki TIFF każdej pozycji z JBC; pliki TIFF na taśmach lub na macierzy.
6. Archiwizacja kopii mikrofilmowych
   Tworzenie kopii mikrofilmowych ze skanów konwertowanych do formatu JPG
   Ludzie: osoba wykonująca mikrofilmy.
   Sprzęt: serwer SODJBC, konwerter, urządzenie wywołujące.
   Wejście: lista publikacji zeskanowanych; pliki JPG na serwerze SOJBC.
   Wyjście: mikrofilmy w opakowaniu z tektury bezkwasowej z nadaną sygnaturą i kodem kreskowym.

7. Konwersja
   Tworzenie plików publikacyjnych (DjVu).
   Ludzie: bibliotekarze cyfrowi.
   Sprzęt: serwer SOJBC, PC-ty bibliotekarzy cyfrowych, komputer z oprogramowaniem Document Express Enterprise.
   Wejście: lista publikacji zeskanowanych; pliki TIFF na serwerze SOJBC.
   Wyjście: lista publikacji skonwertowanych, pliki DjVu na serwerze SOJBC.
8. Katalogowanie w formacie Dublin Core
   Katalogowanie w programie dLibra polegające na konwersji opisów bibliograficznych wprowadzonych do KKZBUJ w formacie MARC21 do formatu Dublin Core
   Ludzie: bibliotekarze cyfrowi.
   Sprzęt: serwer SOJBC, PC-ty bibliotekarzy cyfrowych, serwer VIRTUI, serwer dLibry.
   Wejście: lista publikacji zeskanowanych na serwerze SOJBC ; KKZBUJ.
   Wyjście: lista opisów bibliograficznych w formacie Dublin Core w programie dLibra.
9. Publikacja
   Dodawanie publikacji do oprogramowania dLibra.
   Ludzie: bibliotekarze cyfrowi.
   Sprzęt: serwer SOJBC, serwer VIRTUI, serwer dLibry, PC-ty bibliotekarzy cyfrowych.
   Wejście: lista publikacja skatalogowanych w VIRTUI, lista publikacji skonwertowanych, pliki DjVu na serwerze SODJBC.
   Wyjście: lista publikacji w dLibrze, pliki DjVu w dLibrze.
10. Zabezpieczenie oryginałów
    Zabezpieczenie oryginalnych dokumentów (w przypadku czasopism, ułożenie ich w pudłach z tektury bezkwasowej w pozycji poziomej).
    Ludzie: konserwatorzy, magazynierzy.
    Sprzęt: PC-ty pracowników, serwer SOJBC, tektura bezkwasowa.
    Wejście: lista publikacji zeskanowanych.
    Wyjście: opakowania z tektury bezkwasowej z nadaną sygnaturą i kodem kreskowym.
11. Zabezpieczenie mikrofilmów
    Wprowadzenie rekordów zasobu i rekordów egzemplarza dla mikrofilmów do KKZBUJ i zabezpieczenie ich w magazynie podręcznym Oddziału Dokumentów Audiowizualnych.
    Ludzie: bibliotekarze Oddziału Dokumentów Audiowizualnych.
    Sprzęt: serwer SOJBC, serwer VIRTUI, PC-ty bibliotekarzy.
    Wejście: lista publikacja skatalogowanych w KKZBUJ.
    Wyjście: lista publikacji w KKZBUJ uzupełniona o rekordy zasobu i rekordy egzemplarzy.

5. Opis formalny publikacji w JBC

Bardzo ważnym aspektem przy tworzeniu biblioteki cyfrowej jest stworzenie poprawnego opisu zgromadzonych i publikowanych dokumentów. Założeniem przy tworzeniu JBC jest, że warunkiem uzyskania poprawnego opisu bibliograficznego w formacie Dublin Core, jest stworzenie opisu w formacie MARC 21 i wprowadzenie do bazy NUKAT. Działanie to ma dwie główne zalety, nie wspominając o oczywistej, jaką jest napełnianie bazy danych opisami dokumentów przechowywanych w bibliotece:

- stworzenie opisu bibliograficznego pozwala na podpięcie rekordu zasobu dla czasopism i rekordu egzemplarza dla pozostałych jednostek bibliotecznych, co ułatwia użytkownikowi dotarcie do informacji o publikacji, do których często droga, prowadząca przez zeskanowany katalog podstawowy, lub kartkowy, jest skomplikowana,
- opis w formacie MARC 21 stworzony przez specjalistę, konwertowany automatycznie do formatu Dublin Core, pozwala na uniknięcie błędów, możliwych do popełnienia przez redaktora integrującego obraz cyfrowy, z metadanymi w dLibrze.

Drugi z tych aspektów czyni twórcę rekordu bibliograficznego w formacie MARC 21 w pełni odpowiedzialnym za jakość opisu, który dołączony do kopii cyfrowej dokumentu zostanie publicznie pokazany światu. Dlatego tak istotne jest, aby wszystkie dokumenty, które w przyszłości będą digitalizowane i publikowane w JBC miały utworzone rekordy bibliograficzne w KKZBUJ.

5a. Przygotowanie modelu metadanych

Model metadanych w formacie Dublin Core dla JBC został oparty na wytycznych Dublin Core Metadata Element Set, Version 1.1. Analizując przypadki zastosowania tego formatu do opisu obiektów w innych bibliotekach cyfrowych nasunęła się potrzeba ujednolicenia opisu atrybutów i wyznaczenia pewnego standardu dla wszystkich publikacji w JBC. Kolejnym problemem był brak niektórych elementów w tej wersji formatu, co doprowadziło do utworzenia podatrybutów dookreślających publikację.

Przy założeniu, że w pierwszej kolejności powstaje opis bibliograficzny dzieła w formacie MARC 21, zaistniała potrzeba takiego modelu Dublin Core, który zawierałby wszystkie pożądane przez bibliotekarzy i użytkowników elementy przejmowane z opisu bibliograficznego w KKZBUJ.

Dla lepszej orientacji przedstawiono kilka przykładów opisów w formacie Dublin Core odbiegających od przyjętego standardu.

- Tytuł : Bayki i przypowieści Ignacego Krasickiego
  - Wariant tytułu : Bajki i przypowieści Ignacego Krasickiego
- Autor : Krasicki, Ignacy (1735-1801)
- Wydawca : Drukarnia A. Marcinowskiego ; Towarzystwo Typograficzne

  - Miejsce wydania/powstania : Wilno
- Data : 1820
- Język : pol
- Temat i słowa kluczowe : literatura 18 w.
- Typ zasobu : książka
- Format : image/vnd.djvu
- Źródło : Biblioteka Jagiellońska, 101 I
  - Identyfikator dokumentu cyfrowego : DIGDRUK000029
- Katalog Biblioteki Jagiellońskiej : kliknij tutaj, żeby przejść
- Lokalizacja oryginału : Biblioteka Jagiellońska
- Status prawny : Domena publiczna (public domain)
- Identyfikator zasobu : oai:jbc.bj.uj.edu.pl:3081

W przedstawionym przykładzie został dodane następujące podatrybuty:

- wariant tytułu – odpowiednik pola 246 w formacie MARC 21,
- miejsce wydania – odpowiednik podpola \a z pola 246 w formacie MARC 21,
- identyfikator kopii cyfrowej dokumentu (odpowiednik sygnatury) pod jaką figuruje dokument w systemie SOJBC.

- Tytuł : Mazurek cis-moll. Op. 50 nr 3 na fort. dedykowany L. Szmitkowskiemu
- Autor : Chopin, Fryderyk (1810-1849)
- Opis : Autograf roboczy
- Data : 1841-1842
- Zakres : 19 w.
- Temat i słowa kluczowe : autografy muzyczne 19 w. ; rękopisy 19 w.
- Typ zasobu : rękopis
- Format : image/vnd.djvu
- Źródło : Biblioteka Jagiellońska, Muz. Rkp. 2202
  - Identyfikator dokumentu cyfrowego : DIGMUZ000114
  - Wersja dźwiękowa utworu : kliknij tutaj, żeby przejść
- Lokalizacja oryginału : Biblioteka Jagiellońska
- Status prawny : Domena publiczna (public domain)
- Identyfikator zasobu : oai:jbc.bj.uj.edu.pl:937

W ramach współpracy z Instytutem Fryderyka Chopina do każdego z rękopisów tego kompozytora opublikowanego w JBC dołączony został odnośnik do wersji dźwiękowej utworów udostępnianego na stronie Instytutu. W atrybucie 'źródło' utworzono podatrybut 'wersja dźwiękowa utworu', obsługiwany przez odnośnik przekierowujący użytkownika do zapisu dźwiękowego. W taki sposób w JBC planowane jest obsługiwanie wszelkich wersji i adaptacji dokumentów posadowionych na innych serwerach.

- Tytuł : Człowiek – media – środowisko
- Autor : Biernacki, Wojciech

- Wydawca : Instytut Geografii i Gospodarki Przestrzennej Uniwersytetu Jagiellońskiego
  - Miejsce wydania/powstania : Kraków
- Data : 2010
- Język : pol
- Zakres : 21 w.
- Temat i słowa kluczowe : mass media ; środowisko
  - Klasyfikacja systematyczna : Geografia
- Typ zasobu : e-książka
- Format : application/pdf
- Źródło : Uniwersytet Jagielloński. Instytut Geografii i Gospodarki Przestrzennej
  - Identyfikator dokumentu cyfrowego : DIGONLINE000010
- Katalog Biblioteki Jagiellońskiej : kliknij tutaj, żeby przejść
- Status prawny : Copyright
  - Właściciel praw : Instytut Geografii i Gospodarki Przestrzennej UJ
- Identyfikator zasobu : oai:jbc.bj.uj.edu.pl:22849

Przytoczony opis dotyczy dokumentu, którego wersja cyfrowa posadowiona jest na innym serwerze, a JBC pełni tylko rolę pośrednika, w dotarciu do pełnotekstowego dzieła. W opisie tego dokumentu pojawiają się dwa, (oprócz wcześniej wspomnianych) wymagające objaśnienia podatrybuty. Pierwszy z nich 'klasyfikacja systematyczna' został wprowadzony na potrzeby utworzenia dodatkowej klasyfikacji tematycznej dokumentów publikowanych w JBC. Opracowanie klasyfikacji systematycznej do trzeciego poziomu w głąb zostało przygotowane przez Oddział Opracowania Przedmiotowego (OOP) Biblioteki Jagiellońskiej, na podstawie katalogu systematycznego BJ.

Jak widać z przytoczonych przykładów dostosowanie schematu Dublin Core do potrzeb JBC uzależnione jest w głównej mierze od rodzaju poszczególnych dokumentów. Zastosowanie odpowiednich zasad opisu w formacie MARC 21 w zależności od rodzaju dokumentu implikuje postać opisu kopii cyfrowej tegoż dokumentu w formacie Dublin Core. Różnice te nie są znaczne z uwagi na porównywalnie niewielką ilość danych przejmowanych z opisu w MARC 21 do Dublin Core, co nie zmienia faktu, iż inaczej wygląda opis dla starego druku czy rękopisu, a inaczej dla mapy lub czasopisma.

Zasady wypełniania poszczególnych pól schematu powinny ulec dalszemu doprecyzowaniu dzięki opublikowanemu w 2011 r. schematowi metadanych PLMET.

5b. Proces tworzenia opisu

Tworzenie opisu w formacie Dublin Core odbywa się na podstawie konwersji danych zawartych w powstałym wcześniej opisie w formacie MARC 21. Procedura ta sprawdza się nie tylko w samym tworzeniu opisu, ale także na poziomie identyfikacji dokumentu, co w przypadku dokumentów ciągłych ma ogromną wagę. Należy tu pod uwagę wziąć choćby częste zmiany tytułów czy redaktorów, zmiany miejsca

publikacji, czy wreszcie istnienie pokaźnej liczby dodatków samoistnych i niesamoistnych. Ustalenie tych szczegółów przy opracowaniu czasopisma z autopsji pozwala na odwzorowanie postaci wydawniczej czasopisma 'ponownie publikowanego' w bibliotece cyfrowej.

Konwersja z formatu MARC 21 następuje na podstawie tabeli przekładu, gdzie informacje z poszczególnych pól i podpól są przejmowane do odpowiednich atrybutów w formacie Dublin Core. Nie wszystkie dane z formatu MARC 21 są przejmowane, gdyż format Dublin Core jest opisem bardziej ogólnym mającym do spełnienia odmienne zadanie. Służy on do identyfikacji kopii cyfrowej obiektu w sieci Internet, zarówno na poziomie Federacji Bibliotek Cyfrowych, Europeany, czy wyszukiwarek internetowych. Na najbardziej ogólnym poziomie, dla przykładu najpopularniejszej wśród wyszukiwarek: Google – roboty, czyli urządzenia przeszukujące sieć, co jakiś czas wchodzą na stronę biblioteki cyfrowej indeksując pojawiające się na niej informacje. Tymi informacjami są dane zawarte w opisach bibliograficznych. Stworzenie opisu w formacie Dublin Core jest o tyle istotne, iż format ten w odróżnieniu od formatu MARC 21 jest w pełni przeszukiwalny i możliwy do indeksowania przez wyszukiwarki internetowe. Jeśli spojrzymy na przeciętnego użytkownika bibliotek cyfrowych i odtworzymy drogę jego poszukiwań, to okaże się, że w 80% przypadków dotarł on do materiałów zgromadzonych w bibliotece cyfrowej poprzez stworzenie zapytania w którejś z wyszukiwarek internetowych [zob. forum: http://forum.biblioteka20.pl].

Mając na uwadze zachowania wyszukiwawcze użytkowników, ilość, a szczególnie jakość metadanych powinna być dobierana tak, aby trafność zapytania była jak największa. Okazuje się, że ten komunał nie zawsze sprawdza się w praktyce i dokonanie korekty opisu po imporcie metadanych z katalogu bibliotecznego ma na celu jego potwierdzenie. W Bibliotece Jagiellońskiej pracują nad tym redaktorzy, odpowiedzialni za dokument, który przygotowują od momentu korekty graficznej, aż do jego publikacji w JBC. Taki system pracy pozwala bibliotekarzom na dokładny wgląd do dokumentu, zapoznanie się z jego treścią, ale co najważniejsze w przypadku czasopism, jego strukturą.

Przykłady na uzupełnienie metadanych informacjami o obiekcie cyfrowym można znaleźć w kilku atrybutach. Jednym z nich jest atrybut 'data' uzupełniany o datę dzienną dla każdego indywidualnego numeru czasopisma. Umożliwia to stworzenie zapytania wyszukiwawczego, którego wynikiem są wszystkie tytuły czasopism, które ukazały się konkretnego dnia w danym roku. Zaleta takiego zapytania jest oczywista dla historyka, który poszukuje informacji dotyczącej zdarzenia mającego swoją określoną datę oraz opisu tego zdarzenia w prasie lokalnej, regionalnej, czy ogólnonarodowej, zróżnicowanego pod względem np. orientacji politycznej.

Kolejnym jest atrybut 'źródło' uzupełniany o informację miejsca przechowywania dokumentu, wraz z jego sygnaturą i zawarty w nim podatrybut 'identyfikator dokumentu cyfrowego'. Ten ostatni pozwala na sprawne zarządzanie kopiami cyfrowymi dokumentów w Bibliotece Jagiellońskiej i łatwe dotarcie do plików master w przy-

padku zamówienia tych plików przez czytelnika, np. w celu wykorzystania ich do publikacji.

Przygotowanie plików z metadanymi do masowego ładowania publikacji wykorzystywane jest głównie w przypadku czasopism. Polega to na odpowiednim nazwaniu plików, odzwierciedlającym strukturę czasopisma. Dla przykładu struktura taka wygląda następująco:

czasopisma – CZAS
DIGCZAS000001_7545_IV_czas – identyfikator dokumentu cyfrowego i sygnatura oryginału
DIGCZAS000001_1939 – rok w obrębie tytułu
DIGCZAS000001_1939_001 – numer w obrębie roku
DIGCZAS000001_1939_001_0001.tiff – pliki w numerze
.
.
.
DIGCZAS000001_1939_001_0006.tiff – pliki w numerze
DIGCZAS000001_1939_002 – numer w obrębie roku
DIGCZAS000001_1940 – rok w obrębie tytułu
DIGCZAS000252_434_V_czas – identyfikator dokumentu cyfrowego i sygnatura oryginału

Zachowanie takiej struktury ma na celu odwzorowanie struktury czasopisma, charakteryzującego się np. dodatkami samoistnymi (wydzielanymi i publikowanymi osobno) i niesamoistnymi (publikowanymi razem z numerem) oraz ułatwienie organizacji procesu masowego ładowania publikacji.

6. Opis rzeczowy publikacji

Powtarzając wyżej przytoczony komunał, iż należy tak skonstruować opis dokumentu, aby był on dla czytelnika łatwy do znalezienia, i porównamy go ze sposobem konstruowania zapytań przez użytkowników, obejmującym z reguły pojedyncze słowa, frazy, a czasami nawet zdania w języku naturalnym, to aby sprostać ich potrzebom wyszukiwawczym najlepszym sposobem jest konstruowanie opisu rzeczowego dokumentu według metod wyszukiwania przez nich przyjętych. Z góry odrzucone zostaje stworzenie opisu przedmiotowego w postaci zdań języka naturalnego, natomiast zastosowanie języka swobodnych słów kluczowych z możliwością użycia fraz wyszukiwawczych dostosowuje się do tych potrzeb wyszukiwawczych. W wielu bibliotekach cyfrowych zauważyć można, że zastosowanie języka swobodnych słów kluczowych pociąga za sobą znaczną niekonsekwencję w ich użyciu, aby nie powiedzieć dowolność. Brak kontrolowanego słownictwa powoduje wystąpienie słów kluczowych w liczbie pojedynczej i w liczbie mnogiej, stosowanie synonimów bez żadnej wersji ujednoliconej, czy stosowanie różnych przypadków we frazach wyszukiwawczych. Prowadzi to do chaosu informacyjnego, co jest do opanowania przy niewielkiej ilości dokumen-

tów opracowywanych przez jednego bibliotekarza, ale w repozytoriach osiągających 20 tys. publikacji, przy których tworzeniu zatrudnionych jest kilka osób, przestaje być możliwe do kontrolowania.

Z tego względu był to ważny aspekt, jakim na poziomie projektowania należało się zająć i co w początkowej wersji doprowadziło do powstania listy haseł przedmiotowych, która zaimplementowana do dLibry miała spełniać rolę katalogu przedmiotowego. Lista tych haseł została przedstawiona w aneksie do tego rozdziału. Kolejne prace objęły próbę konwersji języka haseł przedmiotowych KABA do języka swobodnych słów kluczowych.

6a. Możliwości zastosowania języka KABA w Jagiellońskiej Bibliotece Cyfrowej

Analizując proces opracowania przedmiotowego składający się z następujących czynności:

- analiza dokumentu,
- selekcja informacji,
- charakterystyka słowna (słowa kluczowe),
- przekład na jiw (jhp KABA),
- charakterystyka wyszukiwawcza dokumentu w jhp KABA (opis przedmiotowy dokumentu),

nasuwa się oczywisty wniosek, że słowa kluczowe opisujące treść dokumentu powstają przed translacją do języka informacyjno-wyszukiwawczego. Fakt ten jest częściowo wykorzystywany i w trakcie tworzenia opisu przedmiotowego w rekordzie bibliograficznym słowa kluczowe są w nim umieszczane w polu lokalnym na poziomie bazy KKZBUJ. W procesie konwersji z formatu MARC 21 do formatu Dublin Core pole to jest uwzględnione w tablicy translacji i zastosowane słowa kluczowe trafiają do atrybutu 'temat i słowa kluczowe'. Dodatkowo zastosowany został podatrybut 'klasyfikacja systematyczna' gdzie również z pola lokalnego w rekordzie w formacie MARC 21 trafia hasło przedmiotowe z listy umieszczonej w aneksie.

6b. Reguły przekładu języka KABA na język swobodnych słów kluczowych

W ramach prac nad translacją języka haseł przedmiotowych KABA do języka swobodnych słów kluczowych powstało opracowanie mające na celu ujednolicenie zasad tworzenia słów kluczowych wykorzystywanych w formacie Dublin Core w Jagiellońskiej Bibliotece Cyfrowej w oparciu o język haseł przedmiotowych KABA.

1. Wszystkie dokumenty wprowadzone do JBC posiadają słowa kluczowe, częściowo oparte na jhp KABA oraz klasyfikację systematyczną zawartą w polu 595 rekordu bibliograficznego w Komputerowym Katalogu Zbiorów Bibliotek UJ (KKZBUJ).

2. Klasyfikacja systematyczna z pól 595 rekordu bibliograficznego w KKZBUJ zostaje rozbudowana i wprowadzona do osobnego indeksu rzeczowego.
3. W instrukcji użytkownika JBC znajdzie się informacja, iż opisy przedmiotowe dokumentów z hasłami w jhp KABA znajdują się w KKZBUJ, gdzie istnieje możliwość zdobycia szerszej informacji rzeczowej w oparciu o indeks przedmiotowy jhp KABA i kartotekę haseł wzorcowych z terminami nadrzędnymi, skojarzonymi i podrzędnymi.
4. Słowa kluczowe w JBC są oparte na terminologii jhp KABA; dotyczy to zwłaszcza tematów jhp KABA (szczególnie nazw osobowych, nazw pospolitych).
5. Tematy – nazwy osobowe są przejmowane jako słowa kluczowe w formie występującej w jhp KABA (wraz z dopowiedzeniami – daty życia i in.).
6. Tematy – nazwy korporatywne, czyli nazwy instytucji i organizacji oraz nazwy imprez są przejmowane z jhp KABA jako słowa kluczowe, jeżeli występują w języku polskim oraz w prostej formie językowej np. *Biblioteka Jagiellońska.* Jeśli nazwa korporatywna występuje w formie obcojęzycznej (np. nazwy zakonów) do tworzenia słów kluczowych stosuje się polski odpowiednik nazwy (w jhp KABA – termin odrzucony [TO]). Jeśli nazwa korporatywna w jhp KABA zaczyna się od nazwy geograficznej lub jest zbyt skomplikowana, słowo kluczowe tworzy się także w oparciu o terminy odrzucone w jhp KABA.
7. Tematy – nazwy geograficzne. Słowa kluczowe w JBC związane z nazwami geograficznymi są oparte na nazwach geograficznych jhp KABA, jednak – generalnie – bez dopowiedzeń lokalizujących. W wyjątkowych wypadkach, gdy istnieje taka potrzeba, występuje w JBC dopowiedzenie lokalizujące. W wyjątkowych wypadkach dopuszcza się zmianę nazwy geograficznej np. przez inwersję nazwy z jhp KABA lub wykorzystanie terminów odrzuconych [TO] w jhp KABA. np. jhp KABA: *Osmańskie, Imperium;* JBC słowo kluczowe:: *Imperium Osmańskie.*
8. Określniki geograficzne [OG] mogą być stosowane jako odrębne słowa kluczowe. Jeśli OG występuje bezpośrednio po temacie, możliwe jest tworzenie fraz słów kluczowych złożonych z tematu i OG np. jhp KABA: *Polska – historia*; JBC słowo kluczowe: *Historia Polski.*
9. Określniki rzeczowe [ORZ]. Tematy z jhp KABA połączone z określnikiem rzeczowym [ORZ] nie są przejmowane w niezmienionej formie do JBC. Słowa kluczowe w JBC są tworzone osobno w oparciu o temat i określnik. W niektórych przypadkach słowa kluczowe są budowane w oparciu o frazy tworzone z terminów odrzuconych (TO) jhp KABA np. jhp KABA: *Zwierzęta -- ochrona;* JBC słowo kluczowe: *ochrona zwierząt,* jhp KABA: *polowanie –prawo;* JBC słowo kluczowe: *prawo łowiecki.* W wypadku, gdy w jhp KABA temat z określnikiem rzeczowym jest hasłem przedmiotowym rozwiniętym, nie posiadającym TO, dopuszcza się tworzenie swobodnych słów kluczowych z terminów tematów i określników jhp KABA np. jhp KABA: *Sieroty – opieka i pomoc;* JBC słowo kluczowe: *opieka nad sierotami.*

10. Określniki chronologiczne [OCH]. Jeśli w jhp KABA występuje OCH, słowa kluczowe tworzy się z OCH. W JBC stosuje się swobodne określniki chronologiczne; niektóre z nich ulegają w JBC uproszczeniu np. OCH 1870-1914 zastępuje się w JBC OCH 19 w. lub 1900-1945. Przy pewnych tematach zachowuje się OCH związane np. nazwy wojen, rewolucji itp.
11. Określniki formy istniejące w jhp KABA nie są stosowane w JBC jako słowa kluczowe. Terminy jhp KABA w tym zakresie są wykorzystywane w JBC do określenia 'typu zasobu' (w liczbie pojedynczej) np. jhp KABA: *podręczniki;* JBC typ zasobu : *podręcznik;* jhp KABA: *słowniki, glosaria;* JBC typ zasobu: *słownik.*
12. W JBC, podobnie jak w jhp KABA istnieją tematy formalne, stosowane do tekstów utworów literackich. Nazwy gatunków literackich w jhp KABA (tematy formalne – w opisie bibliograficznym dokumentu pole 655) są wykorzystywane w JBC do tworzenia słów kluczowych według zasady:
    - nazwa gatunku z przymiotnikiem językowo-etnicznym + OCH np. *Powieść francuska 19 w.*
    - szczegółowa nazwa gatunku np. *powieść przygodowa.*

    Opracowania dotyczące literatur i poszczególnych gatunków literackich w jhp KABA (w opisie bibliograficznym dokumentu pole 650) są indeksowane poprzez hasła typu: *[nazwa literatury lub gatunku literackiego]* + ORZ – *historia i krytyka, -- tematy, motywy, -- aspekt polityczny,-- aspekt społeczny* itd. W odniesieniu do tego typu dokumentów konieczne jest utworzenie frazy słów kluczowych np. *Powieść historyczna polska – 19 w. – historia i krytyka*; w JBC fraza: *Historia powieści historycznej polskiej 19 w.*; *Literatura francuska –20 w. – tematy, motywy*; w JBC fraza*: Motywy powieści francuskiej 19 w.* W przeciwnym razie istnieje niebezpieczeństwo, iż przy wyszukiwaniu w JBC użytkownik otrzyma teksty literackie pomieszane z opracowaniami na temat danej literatury lub gatunku literackiego. Przy hasłach przedmiotowych dotyczących gatunków literackich do budowy słów kluczowych stosuje się hasła z jhp KABA np. słowo kluczowe Historia powieści przygodowej francuskiej 20 w. (dla opracowania na temat francuskiej 20-wiecznej powieści przygodowej). Ogólniejsze słowo kluczowe Historia powieści przygodowej oznacza, iż dokument jest opracowaniem powieści przygodowej w ogóle (z różnych języków).
13. Istnieje możliwość tworzenia w JBC dodatkowych słów kluczowych w oparciu o kartotekę wzorcową jhp KABA (terminy nadrzędne oraz skojarzone z danym terminem użytym w opracowaniu danego dokumentu w JHP KABA).
14. Najistotniejsza jest konsekwencja w tworzeniu terminologii słów kluczowych; dla utrzymania tej konsekwencji i kontroli słownictwa – dla terminów słów kluczowych różniących się od terminologii w jhp KABA, proponuje się tworzenie – na bieżąco – słownika słów kluczowych JBC, zwłaszcza dla terminów innych niż występują w jhp KABA.

15. Tagowanie. W osobnym indeksie JBC znajdują się tagi – słowa kluczowe tworzone przez użytkowników; jest to słownictwo częściowo kontrolowane przez moderatora JBC.
16. Ustalenia te dotyczą kolekcji *Książki* w JBC. Opracowanie zbiorów specjalnych wymaga osobnych ustaleń [oprac. Danuta Patkaniowska, Łukasz Mesek].

Przyjęte zasady pozwalają na tworzenie jednolitej i co najważniejsze kontrolowanej bazy słów kluczowych stosowanych w Jagiellońskiej Bibliotece Cyfrowej.

7. Przebieg realizacji projektu

Jak już wspomniano powyżej w ramach Programu Operacyjnego Infrastruktura i Środowisko na lata 2007-2013, Priorytet XI. *Kultura i Dziedzictwo Kulturowe*, Działanie 11.1. *Ochrona i zachowanie dziedzictwa kulturowego o znaczeniu ponadregionalnym* zostały pozyskane fundusze z Unii Europejskiej ze środków Europejskiego Funduszu Rozwoju Regionalnego.

19 lipca 2010 r. została uruchomiona Jagiellońska Biblioteka Cyfrowa. Po zindeksowaniu zasobów przez Federację Bibliotek Cyfrowych JBC plasowało się na 39 miejscu. Wówczas w zasobach JBC znajdowały się głównie rękopisy (247 publikacji), które były podstawą prac w tzw. wersji beta JBC.

Od 1 lipca 2010 r., kiedy zatrudniono do projektu pierwszych dwóch pracowników oraz zmieniono zakres obowiązków, przystosowując go do zadań w JBC, dla dwóch dotychczasowych pracowników Oddziału Dokumentów Audiowizualnych, prace ruszyły znacznie szybciej. W okresie lipiec-wrzesień trwały prace nad wyznaczeniem standardów procesu digitalizacji, sposobem oceny jakościowej wykonanych kopii cyfrowych, zarówno przechowywanych na dyskach optycznych jak i nowo skanowanych dokumentów, wytycznymi dotyczącymi konwersji do formatu DjVu, doborem metadanych, formie publikacji i organizacji kolekcji w JBC. W tym czasie została także przeredagowana i dostosowana graficznie do standardu Uniwersytetu Jagiellońskiego strona Jagiellońskiej Biblioteki Cyfrowej.

Od 1 września 2010 r. zatrudnione zostały w ramach projektu dwie kolejne osoby, które zajęły się zgraniem na macierz i oceną kolekcji starych druków przechowywanych dotychczas na dyskach optycznych. W wyniku tych prac zostało przegrane i ocenione pod względem kompletności i jakości 450 starych druków, ale pod koniec września proces ten został wstrzymany ze względu na zmniejszającą się powierzchnię dyskową na macierzy z jednej strony oraz rozpoczęciem skanowania tytułów czasopism objętych projektem.

W ramach comiesięcznych sprawozdań kontrolowano ilość kopiowanych dokumentów wraz z liczbą plików, ilość plików poddanych standardowej i specjalnej obróbce graficznej, ilość stworzonych opisów publikacji, ilość kontrolowanych plików TIFF otrzymywanych ze skanerów oraz ilość plików konwertowanych do formatu DjVu. Tabele dokumentujące działalność zespołu JBC za poszczególne miesiące załączone są w aneksie.

Od 1 października 2010 r. zostały przyjęte ostatnie dwie osoby tworzące zrąb zespołu Jagiellońskiej Biblioteki Cyfrowej w skład której, weszło osiem osób oraz koordynator prac w osobie kierownika Oddziału Dokumentów Audiowizualnych (ODA).

W pierwszej fazie projektu związanej z pracami nad bieżąco skanowanymi dokumentami zespół został podzielony na trzy grupy, z których każda odpowiedzialna jest za dokumenty cyfrowe powstające na każdym z trzech skanerów. Głównym założeniem organizacji pracy w projekcie jest odpowiedzialność za każdy zdigitalizowany tom jednej osoby, która prace nad nim prowadzi od początku do końca. Proces ten wygląda następująco:

1. sprawdzenie jakości i kompletności wykonanych obrazów – porównanie zapisu w rekordzie zasobu dla tytułu czasopisma w KKZBUJ, wykonanym na podstawie oglądu z autopsji oryginału przez pracownika OMA, z oglądem kopii cyfrowej wgranej na macierz przez skanujących,
2. wypełnienie tabeli z oceną jakości i kompletności wykonanych obrazów oraz na podstawie, której wysyłane są ewentualne prośby o korektę do konkretnego skanującego,
3. obróbka graficzna materiału w programach ScanTaylor i Xn View polegająca na wyczyszczeniu i wyprostowaniu uzyskanych kopii cyfrowych,
4. wgranie obrobionego graficznie materiału do specjalnego katalogu przygotowanego na potrzeby dodatkowego zabezpieczenia dokumentów w postaci mikrofilmów, z którego jest on pobierany na konwerter,
5. konwersja do formatu DjVu według wcześniej przygotowanego profilu dla danej publikacji,
6. utworzenie publikacji planowanej w programie dLibra wraz opisem bibliograficznym w formacie Dublin Core, uzyskiwanym poprzez konwersję opisu z KKZBUJ w formacie MARC21,
7. publikacja wraz z uzupełnieniem informacji dotyczących cech szczególnych numeru.

W pierwszej kolejności rozpoczęto digitalizację *Gazety Lwow*skiej jako najdłużej wychodzącego czasopisma z wyselekcjonowanych do projektu. Miało to na celu sprawdzenie w praktyce ustalonych procedur postępowania przy zabezpieczeniu i opublikowaniu największego, choć jak później okazało, nie najbardziej pracochłonnego obiektu realizowanego w tym projekcie. Również w październiku 2010 r. w porozumieniu z Instytutem Fryderyka Chopina Jagiellońska Biblioteka Cyfrowa udostępniła wybrane rękopisy Fryderyka Chopina wraz z muzycznymi aranżacjami utworów. Projekt ten miał na celu zilustrowanie opublikowanych rękopisów materiałem dźwiękowym udostępnionym w postaci plików w formacie mp3 przez Instytut Fryderyka Chopina. Technologiczne rozwiązanie polegało na umiejscowieniu w opisie bibliograficznym w formacie Dublin Core atrybutu z odnośnikiem do pliku dźwiękowego konkretnego utworu posadowionego na serwerze w Instytucie Fryderyk Chopina.

W październiku został podpisany kolejny projekt *Bezpieczne i kompleksowe udostępnienie zasobów cyfrowych Uniwersytetu Jagiellońskiego w sieci Internet. Etap* I finansowany w ramach Małopolskiego Regionalnego Programu Operacyjnego na lata

2007-2013, Oś Priorytetowa 1 *Warunki dla rozwoju społeczeństwa opartego na wiedzy*, Działanie 1.2 *Rozwój społeczeństwa informacyjnego* z Europejskiego Funduszu Rozwoju Regionalnego. Projekt ten jest realizowany w kooperacji z biblioteką Wydziału Prawa UJ oraz biblioteką Instytutu Muzykologii UJ. W ramach tego projektu zostało wyselekcjonowane kolejne 80 czasopism, z czego 33 tytuły przechowywane jedynie w Bibliotece Jagiellońskiej, oraz 1200 ulotnych druków muzycznych. Dla tego projektu wytypowane do zakupu zostały kolejne skanery: jeden o formacie A1 do skanowania oryginałów oraz skaner do mikrofilmów. Skaner do mikrofilmów pozwoli na szerokie udostępnienie zabezpieczonych już tytułów czasopism, ale dotychczas dostępnych jedynie na miejscu w BJ.

W listopadzie 2010 r. nadal trwały prace nad obróbką graficzną i przygotowaniem do publikacji Gazety Lwowskiej. W związku z ograniczoną przestrzenią dyskową na czterech komputerach lokalnych w zespole JBC służących do obróbki graficznej zaistniała potrzeba doposażenia ich w zewnętrzne 1 TB dyski twarde. Oprócz prac nad *Gazetą Lwowską* sprawdzono i opublikowano 90 tytułów książek, uprzednio skatalogowanych w KKZBUJ w formacie MARC 21.

W grudniu kontynuowane były wciąż prace nad *Gazetą Lwowską*, ale z racji zakończenia okresu 1811-1848 kiedy była wydawana w formacie III, na skaner nr 3 (formatu A2) trafiły Sprawozdania Sejmu Galicyjskiego. W związku z tym zespół w JBC odpowiedzialny za obróbkę kopii cyfrowych z tego skanera wypracował strukturę dla tej publikacji, co wymagało gruntownych badań i zapoznania się zarówno z *Alegatami Sejmu Galicyjskiego,* jak i *Protokołami Sejmu Galicyjskiego*. Ustalono, że czasopismo to będzie publikowane w dwóch wersjach:

1. z podziałem na kadencje, sesje, posiedzenia – każde posiedzenie publikowane osobno w ramach sesji, a te w ramach kadencji,
2. całościowo tak jak są oprawione, z uwzględnieniem indeksów i spisów treści.

Taka forma publikacji miała na celu ułatwienie użytkownikowi dostępu do treści zawartych w dokumencie i możliwość opcji wyboru jego przeszukiwania – w zależności czy chce korzystać z indeksów i przeszukiwać publikację niejednokrotnie przekraczającą 500 stron, czy woli korzystać z chronologicznego podziału na etapy pracy Sejmu i publikacji, jakie z danego okresu obrad się ukazały.

Prowadzono ponadto prace nad projektem Literatura dla dzieci i młodzieży wykonywanym w kooperacji z International Children's Digital Library. W ramach tego projektu zeskanowano, opracowano i opublikowano, tworząc specjalną kolekcję 25 tytułów książek skierowanych do młodych czytelników.

Po zindeksowaniu przez Federację Bibliotek Cyfrowych zasoby JBC znalazły się na 19 miejscu pod względem ilości opublikowanych dokumentów.

W ramach półrocznego podsumowania działalności Jagiellońska Biblioteka Cyfrowa, mogła poszczycić się nie tylko ilością publikacji, ale szczególnie godnym podkreślenia jest fakt, iż wraz z ilością w parze szła również jakość opracowania zbiorów. Metody pracy przyjęte w Jagiellońskiej Bibliotece Cyfrowej odznaczają się spełnieniem standardów obowiązujących w bibliotekach cyfrowych na światowym poziomie. Dotyczy to zarówno jakości publikowanych obrazów, które są czytelne, nawet

w powiększeniu, o 1000%, co stanowi duże udogodnienie dla osób słabo widzących, ale co najważniejsze, dzięki zastosowaniu techniki OCR, tekst w nich zawarty jest edytowalny i przeszukiwalny. Kolejnym ułatwieniem dla użytkowników było zastosowanie linkowania zwrotnego pomiędzy opisami bibliograficznymi w KKZBUJ i JBC. W praktyce przełożyło się to na łatwe w udostępnianiu, przeglądanie pełnego opisu bibliograficznego i bezpośrednio od niego przejście do zasobu cyfrowego. Zwrotność tego sprzężenia pozwoliła na przejście od obrazu cyfrowego do opisu w KKZBUJ, który z racji zastosowanego formatu jest pełniejszy o informacje o oryginale nie uwzględnione w metadanych dołączonych do kopii cyfrowej.

Dla zilustrowania półrocznego dorobku JBC przedstawiona jest tabela 1.

Tabela 1. Działalność JBC w okresie lipiec-grudzień 2010.

Migracja danych				
kopiowanie:	268117	plików	15752,947	GB
korekta nazw:	13370	woluminów	—	—
Opracowanie graficzne				
obróbka wsadowa:	190555	plików	687,5450	GB
obróbka specjalna:	11456	plików	622,886	GB
Stworzenie opisu publikacji				
import:	921	rekordów	—	—
opis z autopsji:	405	rekordów	—	—
Publikacja w dLibrze:				
opublikowanych:	2313	publikacji	—	—
Kontrola jakości TIFF:				
skontrolowanych:	668	woluminów	—	—
Konwersja do DjVu:				
wykonanych:	13780	publikacji	67022	MB
	148107	plików		

Źródło: Opracowanie własne na podstawie materiałów Oddziału Dokumentów Audiowizualnych Biblioteki Jagiellońskiej

Objaśniając tę tabelę należy podkreślić, że w ramach prac nad JBC w 2010 r. zostało przekonwertowane do formatu DjVu 13780 pozycji zeskanowanych w ramach projektu JBC, o łącznej wadze 15752 GB. Waga plików wynikowych w formacie DjVu wynosi 67 GB. 2313 z tych pozycji, przy pomocy oprogramowania dLibra zostało opublikowanych w JBC do końca 2010 r.

Biorąc pod uwagę formalny podział obiektów publikowanych w JBC to w głównej mierze udostępnione zostały czasopisma, z czego zasoby dla czterech tytułów: *Gazeta Lwowska, Sprawozdania Sejmu Galicyjskiego, Alegata Sejmu Galicyjskiego* oraz *Protokoły Sejmu Galicyjskiego* zostały zeskanowane mniej więcej w połowie.

W styczniu 2011 r. w wyniku podjęcia działań nad usprawnieniem procesu publikacji, liczba dokumentów widocznych na stronie JBC zaczęła gwałtownie przyrastać. Zmiana ta nastąpiła po zastosowaniu kreatora masowego ładowania publikacji, co pozwoliło przesłać na serwer wiele publikacji na raz, w tym publikacje grupowe. Funkcja ta ma zastosowanie w przypadku publikacji czasopism, co wiąże się z niewielką reorganizacją stylu pracy zespołu JBC. Od tego momentu każdy z pracowników jest odpowiedzialny za dany wolumin czasopisma od samego początku procesu jego cyfryzacji, czyli przeglądania uzyskanych kopii cyfrowych pod względem kompletności i jakości obrazu, poprzez obróbkę graficzną, konwersję do formatu DjVu, aż do masowego załadowania przygotowanych numerów do programu dLibra, co skutkowało dużym skokowym przyrostem publikacji na stronie JBC. Dla porównania w 2010 r. zostało w JBC opublikowanych 2313 pozycji, z czego niewielka część obejmowała czasopisma, natomiast po zastosowaniu kreatora masowego ładowania publikacji ich liczba w styczniu wyniosła 9779 pozycji. Dla uściślenia należy dodać, że publikacją w programie dLibra, a więc także w Federacji Bibliotek Cyfrowych nie jest tytuł czasopisma, ale jego poszczególny numer. Podyktowane jest to specyfiką programu, według założeń, którego tytuły czasopism są rozbijane na poszczególne numery (łączone jednak w postaci struktury lat, w których czasopismo się ukazywało), co ułatwia z jednej strony publikację, a z drugiej przeszukiwanie zasobu. W przypadku gdyby cały wolumin (a woluminy mogą być odmiennie oprawne w zależności od biblioteki) został zeskanowany i wprowadzony do publicznego dostępu jako całość, jego przeglądanie byłoby praktycznie niemożliwe, a to za sprawą zbyt dużego rozmiaru pliku wygenerowanego dla dzieła obejmującego średnio 750 stron. Co więcej, udostępnianie pojedynczych numerów pozwala na odtworzenie struktury czasopisma w takiej formie, w jakiej się ono ukazywało i dodatkowo umożliwia wyłuskiwanie braków, możliwych do opublikowania przez inne biblioteki cyfrowe. Przykładem takiej koordynacji działań na rzecz udostępnienia czytelnikowi całości dzieła jest tytuł *Rzeszowianin*, opublikowany w Podkarpackiej Bibliotece Cyfrowej (PBC), a uzupełniony w ramach prac JBC. Równolegle informacja o dostępie do tego tytułu z poziomu dwóch bibliotek cyfrowych jest uwzględniana w opisie bibliograficznym w wprowadzonym do KKZBUJ w postaci dwóch odnośników kierujących użytkownika do repozytoriów, gdzie opublikowane zostały poszczególne fragmenty zasobu.

Zastosowanie kreatora masowego ładowania skutkowało wynikiem 12168 publikacji w dniu 29.01.2011, co uplasowało JBC na 12 miejscu w Federacji Bibliotek Cyfrowych.

Tak jak w miesiącu poprzednim, także w lutym nadal trwały prace nad publikacją digitalizowanych na bieżąco czasopism. Najdłużej wychodzące wśród wytypowanych czasopism: *Gazeta Lwowska* była już prawie na ukończeniu. *Sprawozdania Sejmu Galicyjskiego, Alegata Sejmu Galicyjskiego* oraz *Protokoły Sejmu Galicyjskiego* zostały ukończone. Do cyfryzacji przygotowane zostały kolejne duże tytuły: *Dziennik Polski, Gazeta Narodowa, Głos Narodu, Nowa Reforma, Nowości Ilustrowane* i *Słowo Polskie*. Oprócz tego do digitalizacji pozostawało jeszcze ok. 200 krócej ukazujących się tytułów, dopełniających listę przewidzianych do zabezpieczenia w projekcie obiektów.

Ilość obiektów opublikowana w lutym 2011 r. odbiegała od imponującej ilości z poprzedniego miesiąca. W poprzednich miesiącach materiał był głównie przeglądany, obrabiany graficznie i konwertowany do formatu DjVu, z myślą o zastosowaniu kreatora masowego ładowania publikacji. W styczniu nastąpiła kulminacja w postaci ogromnej ilości wpuszczenia do JBC przygotowanego wcześniej materiału. Wyniki zawarte w tabeli z lutego pokazują jak kształtuje się praca zespołu z przetestowanym i wdrożonym narzędziem.

W dniu 14.03.2011 r. JBC plasowała się na 8 pozycji z 25402 publikacjami, co w oparciu o szacunkową ilość 7 tys. publikacji w ciągu miesiąca, zrodziło ambicje wejścia do pierwszej trójki największych bibliotek cyfrowych w Polsce na pierwszą rocznicę działalności JBC w FBC.

W kwietniu i w maju trwały prace nad wyłuskiwaniem wszystkich dodatków do *Gazety Lwowskiej*, wśród których doliczono się 10 dodatków samoistnych. Zadecydowano, że w strukturze publikacji w JBC dodatki samoistne nie są dodawane do każdego poszczególnego numeru gazety, ale zgromadzone razem i publikowane pod strukturą czasopisma głównego. Takie podejście spowodowane zostało sposobem opracowania czasopism w KKZBUJ, gdzie dla każdego dodatku samoistnego tworzy się osobny rekord bibliograficzny i łączy z czasopismem głównym. Każdy z dodatków ma swój odrębny zasób, co niejednokrotnie przekłada się również na zastaną praktykę biblioteczną, gdzie wydzielone dodatki były oprawiane osobno, a czasami nawet przechowywane pod innymi sygnaturami.

Od czerwca rozpoczęła się praca nad krócej wychodzącymi tytułami, jak również nad zbiorami dotychczas przechowywanymi na dyskach optycznych. Równolegle prowadzone były prace nad przygotowaniem materiału do skanowania w ramach projektu *Bezpieczne i kompleksowe udostępnienie zasobów cyfrowych Uniwersytetu Jagiellońskiego w sieci Internet. Etap I*. Tak jak w realizowanym już projekcie, skupiono się nad opracowaniem czasopism wraz z dokładną informacją o ich zasobie. Ten sposób organizacji pracy sprawdził się już w praktyce i potwierdza hasło społeczeństwa informacyjnego: najważniejsza jest informacja.

19 lipca 2011 r. minął rok od uruchomienia Jagiellońskiej Biblioteki Cyfrowej. Dla ilustracji rocznej działalności JBC przedstawiona jest tabela 2:

Tabela 2. Działalność JBC od lipca 2010 do czerwca 2011.

Migracja danych				
kopiowanie:	268117	plików	15752,947	GB
korekta nazw:	13370	woluminów	-	-
Opracowanie graficzne				
obróbka wsadowa:	190555	plików	687,5450	GB
obróbka specjalna:	11456	plików	622,886	GB
Stworzenie opisu publikacji				
import:	921	rekordów	-	-
opis z autopsji:	405	rekordów	-	-

Publikacja w dLibrze:				
opublikowanych:	2313	publikacji	-	-
Kontrola jakości TIFF:				
skontrolowanych:	668	woluminów	-	-
Konwersja do DjVu:				
wykonanych	13780	publikacji	67022	MB
	148107	plików		

Źródło:Opracowanie własne na podstawie materiałów Oddziału Dokumentów Audiowizualnych Biblioteki Jagiellońskiej.

Kolejne miesiące 2011 r. to skupienie się głównie na zawartości dysków optycznych, ich weryfikacji pod względem kompletności i jakości skanów. Szczególnie ten drugi aspekt nastręcza wielu problemów, ponieważ kryteria oceny jakości nie są ostre. Zespół JBC pod tym względem stoi na straży jakości, co przekłada się na konieczność ponownego skanowania większości dokumentów zdigitalizowanych w Bibliotece Jagiellońskiej w latach 2005-2008.

Każdy z rodzajów dokumentów zgromadzonych na dyskach optycznych opracowywany jest przez jednego pracownika, co pozwala na zachowanie pełnej kontroli nad zbiorem. Po sprawdzeniu kompletności i jakości kopii cyfrowej dokumentu jest on wgrywany na macierz poprzez system SOJBC co gwarantuje utworzenie i zachowanie kolejnej kopii cyfrowej w bibliotece taśmowej. Po etapie obróbki graficznej i wygenerowaniu podstawy dla stworzenia kopii mikrofilmowej dokument jest gotowy do publikacji. Ta jednak następuje dopiero wówczas gdy w KKZBUJ pojawi się rekord bibliograficzny dla danego dokumentu. Jest to ostatni warunek, którego spełnienie pozwala na dopełnienie procedury publikacji wraz z wprowadzeniem linkowania zwrotnego z poziomu opisu bibliograficznego w KKZBUJ i metadanych w JBC.

Oczywiście nadal digitalizowane są czasopisma i wśród znaczących tytułów pojawił się *Głos Narodu* czy *Gazeta Narodowa*. W ramach uzupełniania tytułów galicyjskich o tytuły krakowskie w projekcie *Bezpieczne i kompleksowe udostępnienie zasobów cyfrowych Uniwersytetu Jagiellońskiego w sieci Internet. Etap I* zabezpieczona i udostępniona została między innymi *Gazeta Krakowska* ukazująca się w latach 1796-1831, *Piast* ukazujący się w latach 1913-1949, czy *Prawda* ukazująca się w latach 1896-1921.

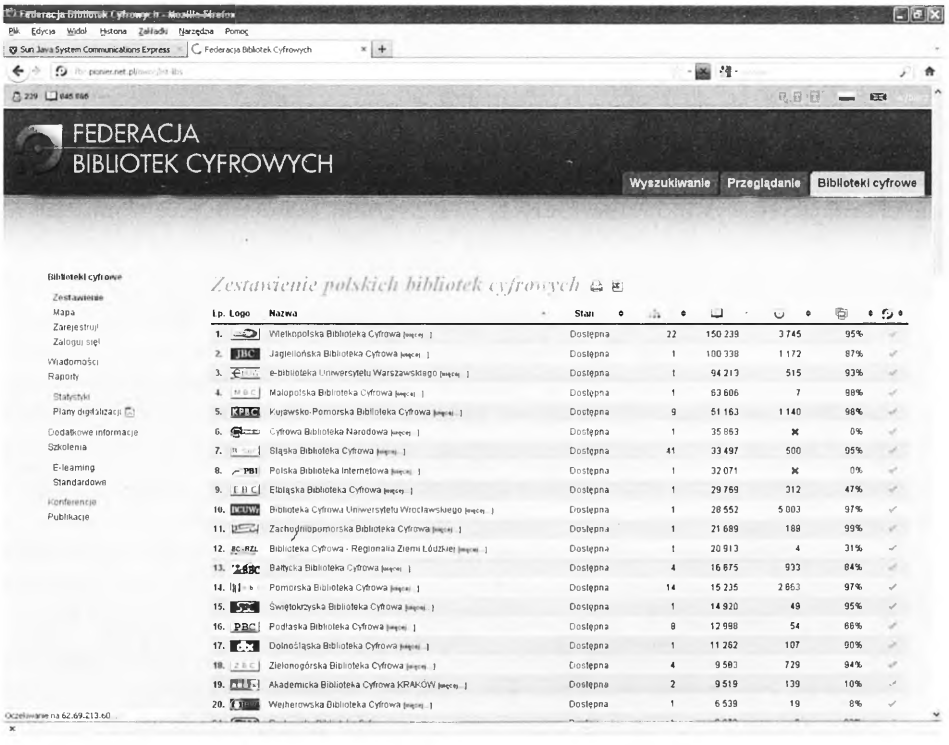
Jednocześnie prowadzone są cały czas prace nad dostosowaniem schematu metadanych do formatu PLMET opublikowanego i rozwijanego przez PCSS z jednej strony oraz realizowanie potrzeb instytucji, które zgłaszają swój akces do JBC z drugiej. Pierwszą z jednostek Uniwersytetu Jagiellońskiego, która zdecydowała się na publikację digitalizowanych dokumentów przechowywanych we własnych zbiorach jest Archiwum UJ. W ramach projektu SYNAT udostępnione zostaną na platformie JBC prace doktorskie obronione na Uniwersytecie Jagiellońskim w latach 2004-2011. Następne w kolejności będą pozostałe prace dyplomowe: magisterskie i licencjackie. Również Instytut Informacji Naukowej i Bibliotekoznawstwa UJ w ramach

prac prowadzonych w projekcie SYNAT wprowadzi do JBC opisy bibliograficzne dla dokumentów elektronicznych o dostępie zdalnym.

Kolejnym, bardzo ciekawym projektem, jeszcze realizowanym w momencie pisania tego tekstu, jest zabezpieczenie i udostępnienie w JBC ok. 100 starodruków i druków XIX-wiecznych ze zbiorów prywatnych profesora Janusza Skalskiego. Jest to kolekcja dzieł medycznych jedyna w swoim rodzaju, a do digitalizacji została wybrana ta jej część, której Biblioteka Jagiellońska nie posiada w swoich zbiorach.

Wśród wykonanych w drugim półroczu 2011 r. prac najważniejszą była migracja do wersji 5.1 dLibry. Pociągnęło to za sobą migrację danych, co skutkowało przerwami w dostępie do zasobów JBC, ale w ogólnym rozrachunku przyniosło pozytywne efekty w postaci zwiększenia szybkości dostępu do poszczególnych publikacji, a także inny sposób ich prezentacji. Główna różnica polega na otwieraniu zawartości dokumentu w nowym oknie oraz rozbudowanymi możliwościami nawigacyjnymi.

Najlepszą ilustracją działalność JBC w roku 2011 będzie zestawienie ukazane na rysunku 1.



Rys. 1. Ranking bibliotek cyfrowych zrzeszonych w FBC na dzień 2.01.2012 r.

Osiągnięte wyniki pozwalają być dobrej myśli, co do czasu realizacji całego projektu i jego wykonania w terminie, ponieważ w ciągu półtora roku, a więc w połowie terminu jego realizacji udało się zrealizować 60% planu, a do zakończenia projektu w kwietniu 2013 r. pozostaje jeszcze ponad rok.

Wyzwania i problemy

W dotychczasowej pracy zespół JBC natknął się na wiele problemów, które wspólnymi siłami udało się rozwiązać. Wprowadzone zostały nowe skrypty, czy nawet programy ułatwiające proces obróbki, dodawania metadanych czy samą publikację w dLibrze.

Jednak aby nadal usprawniać proces digitalizacji w Bibliotece Jagiellońskiej konieczne jest, aby w każdym z oddziałów przekazujących dokumenty do digitalizacji powstał opis bibliograficzny w KKZBUJ przekazywanego dokumentu, oraz wprowadzenie opisu tego dokumentu do SOJBC, czyli systemu kontrolującego proces digitalizacji w BJ. Ma to na celu lepsze zarządzanie digitalizowanym dokumentem, a zarazem jest pomocne przy imporcie opisów bibliograficznych i konwertowaniu ich z formatu MARC21 do formatu Dublin Core. Posłuży także scaleniu bazy danych katalogu komputerowego z bazą publikacji w JBC i umożliwi umieszczenie odnośników w każdym z opisów, tak, aby czytelnik miał dostęp do komplementarnej informacji. Co więcej, przy opracowaniu czasopism konieczne jest szczegółowe odzwierciedlenie zasobu, a nie tylko pobieżne wpisy bez wnikliwego obadania materiału z autopsji. Brak szczegółowego zasobu skutkuje wątpliwościami, które pojawiają się w zespole JBC i ponownej, zbędnej kontroli oryginałów w pracowni Foto.

Nadal odczuwany jest brak opracowania haseł przedmiotowych, które pozwolą na przekonwertowanie ich do postaci swobodnych słów kluczowych czy fraz, zdecydowanie ułatwiających wyszukiwanie. Dotyczy to szczególnie czasopism oraz w mniejszym zakresie zbiorów specjalnych. Stąd wniosek, aby Oddział Opracowania Przedmiotowego Zbiorów Nowych włączył się w prace nad tworzeniem haseł przedmiotowych dla czasopism, a także, aby w miarę możliwości Oddziały Zbiorów Specjalnych także włączyły się w tworzenie haseł przedmiotowych dla swoich zbiorów.

Aby sprostać wymogom, wedle których Jagiellońska Biblioteka Cyfrowa ma gromadzić zasoby całego Uniwersytetu Jagiellońskiego, przeorganizowana zostanie struktura JBC tak, aby w przyszłości wydzielić specjalne kolekcje dla poszczególnych jednostek uniwersytetu, gdzie będą udostępniane dokumenty w tych jednostkach tworzone.

W następnych programie ramowym na lata 2014-2020 Biblioteka Jagiellońska będzie szukała nowych środków finansowych na dalszą digitalizację, a w szczególności uważane są za konieczne do przeprowadzenia prace mające na celu poprawę OCR.

Podsumowując, Jagiellońska Biblioteka Cyfrowa dzięki wsparciu finansowemu Uniwersytetu Jagiellońskiego i Unii Europejskiej, sprawnej organizacji i ogromnemu zaangażowaniu zespołu, jest najdynamiczniej rozwijającą się biblioteką cyfrową w Polsce.

Summary

This monograph deals with the comprehensively analysed and presented digital libraries' issues and it is written mainly by the representatives of Polish researchers, theoreticians, librarians, and practitioners environments. Their response to the issues of the history and beginnings, terminological distinction, legal aspect of digital libraries creation, organisation and preservation, description of processes, methodology, different models and standards, digital libraries' users, case study, consists the main subject of this book. The chapters concerning different articles are written by various authors and are grouped into four main sections providing a diverse, but heterogeneous view of the main themes discussed by the authors. Because of importance of the digital libraries' issues in Poland and worldwide, and holistic approach to the problem with relevance to the international trends and digital repositories policy, the book is divided into four parts.

The first part, the most general is devoted to theoretical, basic and financial issues concerning creation, implementation and preservation of digital libraries. This chapter is composed of articles about: 1) terminological description and analysis of electronic, virtual and digital libraries definitions' distinction (*Digital library, electronic library, virtual library* by Małgorzata Janiak), 2) history of the creation, organisation and management of Digital repositories' Project and digital libraries in Poland and worldwide (*Digital libraries worldwide – the beginning and development* by Władysław Marek Kolasa) with the issues concerning 3) the most important international and national conferences devoted to digital libraries issues (*Digital libraries in the scope of the international and national conferences in 1993-2011* by Monika Krakowska), 4) bibliometric analysis of the literature on digital libraries (*Publications on Digital libraries – bibliometric analysis* by Magdalena Wójcik), 5) legal issues in the process of creation, organisation, management and preservation of digital libraries (*Legal issues* by Sybilla Stanisławska-Kloc), 6) standardisation in the process of organising and managing the digital libraries based on international guidelines and *A Framework of Guidance for Building Good Digital Collections of the National Information Standards Organization* (NISO STANDARD) *Framework Working Group, Institute of Museum and Library Services* (IMLS) (*Standards* by Bożena Bednarek-Michalska), 7) the criteria of the quality evaluation of digital libraries (*Quality evaluation of digital libraries* by Małgorzata Janiak, Monika Krakowska), 8) aesthetics issues of digital libraries (*Aesthetic issues and digital libraries* by Małgorzata Janiak), 9) digital libraries' financing and economic issues (*The costs and financing of the digital libraries* by Małgorzata Kowalska) and the use of digital content by researchers (*Digital libraries in the humanities' researcher work* by Mirosław Górny).

The second part is devoted to users, developers and employees issues related to digital libraries and is consists of the articles related to 1) digital libraries' users (*Users of the Polish digital libraries* by Jolanta Mazurek, 2) digital libraries users' mental models (*Mental models of the users of digital libraries* by Monika Krakowska), 3) professional competences of the digital librarian (*Model of the professional competences* by Łukasz Mesek).

The third section of the monograph is dedicated to content and processes issues and is concerned by the articles devoted to: 1) the projects and processes, functionality of digital libraries and digital collections creation (*Creation and development of digital libraries* by Cezary Mazurek, Tomasz Parkoła, Marcin Werla), 2) digitisation processes and digital libraries organisation (*Organisation and logistic of digital libraries* by Tomasz Kalota, Marcin Szala), 3) the formal description of digital collection (*Data structure* by Marek Nahotko, *Metadata in the Polish Digital libraries* by Piotr Myszkowski, *Mapping the electronic documents' content* by

Marek Nahotko, *Formal description of document* by Danuta Patkaniowska and *Correlations between descriptions in library catalogues and digital libraries* by Łukasz Mesek), 4) the analysis of documents' formats (*Formats of the documents in the digital libraries* by Władysław Marek Kolasa, *DjVu format* by Bartłomiej Siedlarz).

The last, fourth section of the monograph is dedicated to the existing, fundamental Polish digital libraries examples (case study) and there are presented: 1) the Federation of Digital Libraries (*Federation of Digital Libraries – case study* by Cezary Mazurek, Marcin Werla), 2) the Digital Library of Wielkopolska (*Digital Library of Wielkopolska – case study* by Ewa Wichlińska), 3) the National Digital Library Polona (*National Digital Library Polona (NDL Polona)* by Joanna Potęga), and the regional digital libraries such as: 4) the Kujawsko-Pomorska Digital Library (*Kujawsko-Pomorska Digital Library – case study* by Bożena Bednarek-Michalska), the Malopolska Digital Library (*Malopolska Digital Library – regional digital repository* by Wojciech Kowalewski, Władysław Marek Kolasa), the Silesian Digital Library (*Silesian Digital Library – case study* by Remigius Lis) and the Jagiellonian Digital Library (*Jagiellonian Digital Library – case study* by Łukasz Mesek).

The monograph is placed within the group of the 21st century digital libraries' books published in Poland. Previous publications devoted to the less or more detailed to the digital libraries' issues was: *Modern Library Digital Environment* (Jurand B. Czermiński, 2002), *Digitisation in communication processes* (edited by Wanda Krzemińska and Piotr Nowak, 2004), *Digitisation of library collections: theory and practice* (Aleksander Trembowiecki, 2006), *Digitisation of library collections* (edited by Elżbieta Stefańczyk, 2006), conference proceedings published in 2006 by the Library of Lodz *Headed the library of the 21st century: Will we survive?* and the *Standards in the process of digitisation of objects of cultural heritage* (edited by Grzegorz Płoszajski (2008). The attention should be paid on: the handbook developed and edited by Dariusz Paradowski titled *Digitisation of literature*, released by the Polish National Library in 2010, and collective work edited by Jadwiga Woźniak-Kasperek and Grzegorz Franke from 2007: *Digital libraries: projects, case studies, technologies*, a book of Małgorzata Kowalska titled *Digitisation of the Polish libraries collections* from 2008, as well as *Digital libraries platform dLibra* by Anna Wałek (2009). Themes surrounding digital libraries' issues appeared also in several conference proceedings *Information for science and the world of digital content* edited by Halina Ganińska (2008), then in the publication titled *Publications universum and the electronic communication* edited by Krzysztof Migoń and Marta Skalska-Zlat from 2010, *Digital World of document: publishing houses, libraries, museums, archives*, edited by Henryk Hollender. The collection of articles presented at the conference *Polish Digital Libraries* organized by the library of Polish Academy of Science in Kórnik, Foundation of Scientific Libraries in Poznan and Poznan Supercomputing and Networking Center (PSNC) should be mentioned also as the publication concerning most important and actual digital libraries problems published annual from 2008 to 2011 after each national conference in Poznan.

These four parts are based on solid review of literature, research and practical work experience, and a variety of sources reflecting research and practice of modern digital repositories. Thanks to this broaden, holistic approach book could be one of the theoretical and research literature on various aspects of digital libraries. The authors manage to touch upon most of the problems regarding digital libraries in 21st century presenting the results of their research and knowledge based on practical experience in individually, various but complementary prepared articles. That is why the monograph could be helpful for most digital environment users and destined to broaden audience, from digital libraries practitioners – librarians, administrators, developers, to end users, students interested in digital repositories and collections' research area and all people who want to expand their knowledge about digital libraries.

Keywords: digital libraries; digital libraries – projects, digital libraries – case studies, digital libraries – technologies, digital libraries – users; digitisation processes.

28556 + CD
M.2

Monografia obejmuje całokształt zagadnień związanych z bibliotekami cyfrowymi, poczynając od ich genezy i dalszej historii, problemów terminologicznych, prawnych, poprzez opis procesów, metod, modeli, aż po omówienie istniejących bibliotek (case studies). Do opracowania poszczególnych rozdziałów zostali zaproszeni teoretycy i praktycy zajmujący się tworzeniem bibliotek cyfrowych z całej Polski.

Książka jest podzielona na cztery części. Pierwsza ogólna, w której znalazły się artykuły związane z zagadnieniami teoretycznymi, historycznymi, prawnymi i finansowymi dotyczącymi bibliotek cyfrowych, druga poświęcona ludziom związanym z bibliotekami cyfrowymi użytkownikom, twórcom i pracownikom bibliotek, trzecia omawiająca procesy i projekty tworzenia kolekcji cyfrowych oraz czwarta przedstawiająca opisy bibliotek cyfrowych.

* * *

Odbiorcami książki mogą być pracownicy naukowi zainteresowani tworzeniem i wykorzystywaniem zasobów cyfrowych, praktycy budujący i planujący utworzenie biblioteki cyfrowej, studenci oraz obecni i przyszli użytkownicy kolekcji cyfrowych.

Seria wydawana przez Wydawnictwo
STOWARZYSZENIA BIBLIOTEKARZY POLSKICH
we współpracy
Z INSTYTUTEM INFORMACJI NAUKOWEJ
I STUDIÓW BIBLIOLOGICZNYCH
UNIWERSYTETU WARSZAWSKIEGO

9 788361 464709

Cena 65 zł

ISBN: 978-83-61464-70-9