

# ZIN

Studia informacyjne  
Information studies

VOL. 54 2016 NO. 1(107)

p-ISSN 0324-8194

e-ISSN 2392-2648



---

**WDiB UW**

---



STOWARZYSZENIE  
BIBLIOTEKARZY  
POLSKICH

# ZAGADNIENIA INFORMACJI NAUKOWEJ

Studia informacyjne

## ISSUES IN INFORMATION SCIENCE

Information Studies

VOL. 54 2016 NO. 1(107)

p-ISSN 0324-8194

e-ISSN 2392-2648



**WDiB UW**



STOWARZYSZENIE  
BIBLIOTEKARZY  
POLSKICH



Warszawa 2016

## ISSUES IN INFORMATION SCIENCE – INFORMATION STUDIES

The core purpose of this journal is to provide a forum for the dissemination of scientific papers and research results in the field of information science and other disciplines which analyze social and technological aspects of various information-related activities performed by contemporary communities. Moreover, the journal is to disseminate critical reviews and summaries of new publications in the field of information science and reports from important conferences discussing contemporary information problems.

In addition to that, after fifty years *Issues in Information Science* (ZIN) becomes bilingual to reach foreign readers – papers are available either in Polish or English.

The reconstructed journal is published with slightly extended title *Issues in Information Science – Information Studies* (ZIN – *Information Studies*) – the subtitle emphasizes the interdisciplinary nature of its subject profile covering a broad spectrum of issues studied by various academic disciplines and professional activity domains related to access to resources of recorded information and knowledge and the use of these resources by contemporary man and society. Other subjects to be covered by ZIN involve: 1) theoretical ponderings on the practice of information-related activities performed by various communities, 2) the results of research on the conditions influencing those activities and ways of improving methods and tools employed for the activities in question, 3) the methodology of information science research, information science history and education concerning the information science. The subject profile of semiannual ZIN – *Information Studies* covers, among else, the issues of:

- information science in relation to library science, archival science, museology and other disciplines researching preservation and access to scientific and cultural heritage
- information and knowledge management
- traditional and online scholarly communication
- information and knowledge organization
- metadata theory and practice
- Web 2.0
- Semantic Web
- information architecture
- information websites usability
- digital humanities
- human-computer interaction
- natural language processing
- information retrieval
- use of information and behavior of the information users
- social response to modern information technologies
- digital humanities
- information and digital skills
- information policy
- information ethics.

ZIN – *Information Studies* is addressed to: 1) information science teachers and lecturers, researchers and students, 2) practitioners of information-related activities who analyze methods and tools used to implement those activities in various domains and organizational environments, 3) politicians and donors related to information activities in various domains. The journal content may also be of some interest to teachers, students and researchers in other disciplines of science which deal with various aspects of information existence and use in the contemporary world.

ZIN – *Information Studies* is included in 'B' list of journals scored by Polish Ministry of Science and Higher Education and indexed by: Central European Journal in Social Sciences and Humanities (CEJSH), Cambridge Scientific Abstracts (CSA), Library and Information Science and Technology Abstracts (LISTA), Polish Bibliography of Book Studies (PBB), Knowledge Organization Literature and Polish Scholarly Bibliography (PBN).

## ZAGADNIENIA INFORMACJI NAUKOWEJ – STUDIA INFORMACYJNE

Głównym celem niniejszego czasopisma jest zapewnienie forum dla rozpowszechniania artykułów naukowych i wyników badań z zakresu nauki o informacji (informatologii) oraz innych dyscyplin, w których podejmowane są analizy społecznych i technologicznych aspektów działalności informacyjnej prowadzonej w różnych sferach współczesnego życia społecznego. Czasopismo służyć ma również rozpowszechnianiu krytycznych recenzji i omówień publikacji z tego zakresu oraz problemowych sprawozdań z ważnych konferencji poświęconych współczesnym problemom informacyjnym.

W minionym pięćdziesięcioleciu *Zagadnienia Informacji Naukowej* (ZIN) były czasopismem publikującym teksty wyłącznie po polsku, a zatem adresowanym tylko do czytelnika polskiego.

W nowej formie czasopismo adresowane jest zarówno do czytelnika polskiego jak i zagranicznego, publikujemy artykuły zarówno w języku polskim jak i angielskim. Obecnie czasopismo ukazuje się pod rozszerzonym tytułem: ZIN – *Studia Informacyjne*. Dodany podtytuł podkreśla interdyscyplinarny charakter jego profilu tematycznego, który obejmuje szeroki zakres problemów podejmowanych przez dyscypliny akademickie i dziedziny działalności zawodowej związane z zapewnianiem dostępu do utrwalonych zasobów informacji i wiedzy oraz ich wykorzystywaniem przez współczesnego człowieka i współczesne społeczeństwo. Czasopismo publikuje też artykuły prezentujące teoretyczną refleksję o praktycznej działalności informacyjnej prowadzonej w różnych dziedzinach i obszarach życia społecznego, a także wyniki badań służących poznaniu różnych uwarunkowań tej działalności oraz doskonaleniu jej metod i narzędzi. Na łamach ZIN publikowane są także artykuły poświęcone metodologii badań informatologicznych, historii nauki o informacji oraz edukacji w zakresie nauki o informacji. Profil tematyczny półrocznika ZIN – *Studia Informacyjne* obejmuje m.in. problematykę:

- nauki o informacji w powiązaniu z bibliotekoznawstwem, archiwistyką, muzeologią innymi dyscyplinami zajmującymi się problematyką zachowania i zapewnienia dostępu do dziedzictwa nauki i kultury
- zarządzania informacją i wiedzą
- komunikacji naukowej i cyfrowej komunikacji naukowej
- organizacji informacji i wiedzy
- teorii i praktyki metadanych
- zagadnień Web 2.0
- zagadnień Sieci Semantycznej
- architektury informacji
- projektowania użytecznych serwisów informacyjnych
- humanistyki cyfrowej
- interakcji człowiek – komputer
- przetwarzania języka naturalnego
- wyszukiwania informacji
- wykorzystywania informacji i zachowań informacyjnych użytkowników
- społecznej recepcji nowoczesnych technologii informacyjnych
- kompetencji informacyjnych i cyfrowych
- polityki informacyjnej
- etyki informacyjnej.

*Zagadnienia Informacji Naukowej – Studia Informacyjne* adresowane są do wykładowców, badaczy i studentów nauki o informacji, a także praktyków działalności informacyjnej, krytycznie analizujących metody i narzędzia jej realizacji w różnych środowiskach dziedzinowych i organizacyjnych oraz polityków i donatorów działalności informacyjnej w różnych dziedzinach. Lektura czasopisma może też zainteresować wykładowców, studentów i badaczy innych dyscyplin, które zajmują się równymi aspektami funkcjonowania informacji we współczesnym świecie.

*Zagadnienia Informacji Naukowej* znajdują się na liście B czasopism punktowanych Ministerstwa Nauki i Szkolnictwa Wyższego. Czasopismo jest indeksowane w bazach: Central European Journal in Social Sciences and Humanities (CEJSH), Cambridge Scientific Abstracts (CSA), Library and Information Science and Technology Abstracts (LISTA), Polska Bibliografia Bibliologiczna (PBB), Knowledge Organization Literature, Polska Bibliografia Naukowa (PBN).



# Spis treści | Contents

ROZPRAWY. BADANIA. MATERIAŁY | THESES. RESEARCH. MATERIALS

|                                                                                                                                                                                                                                                                                                 |     |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Remigiusz Sapa<br><i>Metodologia badań indywidualnego zarządzania informacją: wybrane aspekty</i> [Research Methodology of Personal Information Management Studies: Selected Aspects] .....                                                                                                     | 7   |
| Marcin Roszkowski, Bartłomiej Włodarczyk<br><i>Cytowania zasobów sieciowych w polskich czasopismach z zakresu bibliotekoznawstwa i informatologii: analiza aktualności adresów URL</i> [Web Citations in Polish Library and Information Science Journals: the Analysis of URLs' Validity] ..... | 21  |
| Ewa Głowacka<br><i>Polskie repozytoria instytucjonalne jako miejsce dla otwartych zasobów naukowych i edukacyjnych</i> [Polish Institutional Repositories as a Place for Open Scientific and Educational Resources] .....                                                                       | 44  |
| Piotr Tańkowski<br><i>Media społecznościowe jako źródło informacji bibliograficznej w naukach humanistycznych. Komunikat z badań</i> [Social Media as a Source of Bibliographic Information in the Humanities. Research Report] .....                                                           | 55  |
| Bartłomiej Włodarczyk<br><i>Przekształcenie relacji występujących w hasłach rozwiniętych i opisach przedmiotowych KABA do postaci powiązań w mapie tematów</i> [The Transformation of Relationships in KABA Subject Heading Strings into Associations in Topic Maps] .....                      | 67  |
| Magdalena Shishenkova<br><i>Credibility of Information: an Overview of the Concept in the Digital Age</i> [Wiarygodność informacji w epoce cyfrowej: przegląd wybranych koncepcji] .....                                                                                                        | 85  |
| Jarosław Pacek<br><i>Pole semantyczne terminu „książka elektroniczna”</i> [Semantic Field of the Term “Electronic Book”] .....                                                                                                                                                                  | 95  |
| Grzegorz Gmiterek<br><i>Sieciowe sposoby społecznościowego zarządzania informacją o dokumentach</i> [Networked Approaches to Social Information Management on Documents] .....                                                                                                                  | 108 |

## RECENZJE I OMÓWIENIA | REVIEWS

|                                                                                                                                                                                                                                                                                          |     |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <i>Dorota Siwecka: Światowy model informacji bibliograficznej: programy i projekty (1950–2010) [World Model of Bibliographic Information: Programmes and Projects (1950–2010)]. Wrocław; Warszawa 2015 (Jakub Maciej Łubocki) .....</i>                                                  | 134 |
| <i>Mariusz Jarocki: Open Source w bibliotekach w świetle badań publicznych uczelni akademickich w Polsce [Open Source in the Libraries in View of the Research on Polish Public Universities]. Toruń 2015 (Grzegorz Gmiterek) .....</i>                                                  | 137 |
| <i>Małgorzata Kowalska: Crowdsourcing internetowy – pozytywny wymiar partycypacji społecznej. Konteksty – istota – uwarunkowania [Online Crowdsourcing – the Positive Dimension of Social Participation. Contexts – Nature – Determinants]. Warszawa 2016 (Katarzyna Materska) .....</i> | 141 |
| <i>Wśród zagranicznych książek [Foreign Publications] (Jacek Wojciechowski) .....</i>                                                                                                                                                                                                    | 145 |
| <i>Przegląd polskich nowości wydawniczych [New Polish Publications] (Anna Stanis) ..</i>                                                                                                                                                                                                 | 154 |

## KRONIKA | CHRONICLE

|                                                                                                                                                                                                                                |            |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| <i>„Europejska konferencja edukacji informacyjnej” (Tallinn, Estonia, 19–22 października 2015) [European Conference on Information Literacy (Tallinn, Estonia, October 19–22, 2015)] (Ewa Rozkosz, Zuza Wiorogórska) .....</i> | 159        |
| <i>„Digital Humanities 2016” (Kraków, 11–16 lipca 2016) (Anna Kamińska) .....</i>                                                                                                                                              | 163        |
| <b>WSKAZÓWKI DLA AUTORÓW   GUIDELINES FOR AUTHORS .....</b>                                                                                                                                                                    | <b>167</b> |

# Metodologia badań indywidualnego zarządzania informacją: wybrane aspekty

Remigiusz Sapa

*Instytut Informacji Naukowej i Bibliotekoznawstwa,  
Wydział Zarządzania i Komunikacji Społecznej, Uniwersytet Jagielloński*

---

## Abstrakt

**Cel/teza:** Celem badań, których wyniki przedstawione są w artykule było zidentyfikowanie sposobów rozwiązywania istotnych dylematów w projektowaniu badań z zakresu indywidualnego zarządzania informacją.

**Koncepcja/metody badań:** Przyjęto perspektywę jakościową, a badaniami objęto publikacje udostępniane w zasobach *Web of Science Core Collection: Citation Indexes* (wyszukane 21 sierpnia 2015 roku), które w tytule zawierały termin *personal information management* i omawiały badania empiryczne różnych aspektów tej problematyki. Dążono do rozpoznania sześciu elementów: ogólnego celu badań, ujęcia w czasie, środowiska badań, rodzaju danych i sposobu ich interpretacji, zastosowanych technik i sposobu doboru uczestników badań.

**Wyniki i wnioski:** Przeprowadzona analiza wskazuje na dużą różnorodność rozwiązań czerpanych z różnych perspektyw metateoretycznych. Wyniki ujawniają kluczowe rozstrzygnięcia metodologiczne podejmowane w badaniach indywidualnego zarządzania informacją.

**Oryginalność/wartość poznawcza:** Wyniki mogą być pomocne w projektowaniu tego typu badań oraz jako punkt wyjścia do pogłębionych rozważań nad ich metodologią.

## Słowa kluczowe

Indywidualne zarządzanie informacją. Metodologia. Projektowanie badań. Techniki badawcze.

Otrzymany: 12 lutego 2016. Zrecenzowany: 30 czerwca 2016. Zaakceptowany: 9 lipca 2016.

---

## 1. Wprowadzenie

Problematyka indywidualnego zarządzania informacją (dalej PIM – od *personal information management*) rozumianego jako

działania podejmowane przez jednostkę w celu pozyskania lub stworzenia, zebrania, zorganizowania, utrzymania, ponownego odszukania, wykorzystania i dystrybucji informacji potrzebnej do realizacji różnych celów (...) i wypełnienia wielu ról życiowych i zadań... (Jones, 2008, 5)

ma charakter wieloaspektowy, a badania skupiają się w dwóch nurtach: działania i zasoby informacyjne jednostek oraz rozwój narzędzi PIM (Jacques & Fastrez, 2014, 321).

Jednym z obszarów, który nadal wymaga intensywnej eksploracji, jest metodologia badań PIM. Jedynie nieliczni autorzy uczynili tę problematykę wyodrębnionym obiektem zainteresowania, a wnioski i refleksje mają charakter wycinkowy, przede wszystkim normatywny i nie są wynikiem jakiegś systematycznej (określonej co do zakresu i sposobu) analizy piśmiennictwa. Nadrzędnym celem, co wydaje się naturalne w przypadku stosunkowo młodego obszaru zainteresowania, jest poszukiwanie odpowiedzi na pytanie, jak

można i jak należy badać, a nie w jaki sposób badania są czy były prowadzone. Np. Charles M. Naumer i Karen E. Fisher podjęli refleksję na temat możliwości wykorzystania „podejścia naturalistycznego” (ang. *naturalistic approach*) w badaniach PIM, odwołując się z jednej strony do specyfiki indywidualnego zarządzania informacją, a z drugiej do korzeni zainteresowania tą problematyką w badaniach zachowań informacyjnych człowieka i interakcji człowiek-komputer oraz w tradycji jakościowych badań etnograficznych. Jednocześnie wskazali na potencjalną wartość stosowania zróżnicowanych podejść badawczych, a zatem też metodologii mieszanej, oraz łączenia różnych technik zbierania danych w jednym badaniu (Naumer & Fisher, 2007). William Jones, formułując dość ogólnikowe zalecenia, także podkreślił znaczenie metody etnograficznej z długotrwałymi studiami przypadków, w których zwraca się szczególną uwagę na konteksty, wykorzystuje się samoobserwację i wywiady (Jones, 2007, 484). Dążono też do zidentyfikowania zmiennych warunkujących zachowania w ramach PIM (Bergman, 2013) i wypracowania narzędzi pomiaru tych zachowań (Helvoort, 2012). Podjęto próbę wyznaczenia podstawowych podejść badawczych do oceny narzędzi PIM (naturalistyczne, długookresowe, laboratoryjne, studia przypadków), która jednak nie ma charakteru klasyfikacji, a raczej stanowi wykaz charakterystyk, które w różnych konfiguracjach można przypisać poszczególnym badaniom (Kelly & Teevan, 2007, 191 i nn.). Jednocześnie w tej samej publikacji znaleźć można wnikliwą i uporządkowaną analizę wyzwań, z jakimi uporać musi się ktoś, kto projektuje badania narzędzi wspierających działania w zakresie PIM, a na trudności metodologiczne badań PIM zwracali też uwagę inni badacze (np. Gwizdka & Chignell, 2007; Whittaker, 2011, 54–55). Nadal jednak słuszne pozostaje stwierdzenie Jonesa, że

rozwój metod przeznaczonych specjalnie do badań PIM jest ciągle w okresie dzieciństwa (Jones, 2007, 482).

## 2. Cel i podstawowe rozstrzygnięcia terminologiczne

Zasadniczym celem badań, których wyniki zostały przedstawione w niniejszym artykule było zidentyfikowanie sposobów rozwiązywania kluczowych dylematów pojawiających się w projektowaniu badań z zakresu PIM. Przyjęto perspektywę jakościową, heurystyczną, dążąc do indukcyjnego odkrywania i opisanie różnych rozwiązań na poziomie szczegółowym, jednocześnie rezygnując z określania powszechności ich stosowania oraz z predefiniowania jakiejś typologii metod składających się z tych rozwiązań (ewentualnie wykorzystania już istniejących), a następnie „dopasowywania” do niej zidentyfikowanych postępowań badawczych. Takie podejście, sprzyjające redukcji wpływu założeń apriorycznych na wynik badań, jest w przekonaniu autora właściwym rozwiązaniem w odniesieniu do metod PIM, co do których przyjęto, że mają charakter idiograficzny. Nie oznacza to jednak założenia całkowitego nieistnienia podobieństw, czy powtarzalności poszczególnych rozwiązań, ale przyjęcie, że ich dobór, zestawienie i wersje projektowane na użytek konkretnych badań mogą składać się każdorazowo na specyficzną i niepowtarzalną całość tworzoną przez badacza.

By uniknąć nieporozumień, konieczne jest określenie znaczeń, w jakich używane będą przynajmniej dwa kluczowe terminy. Pod pojęciem **metody** będzie rozumiany, obejmujący całość postępowania badawczego, spójny i ugruntowany teoretycznie system koncepcji poznawczych, założeń, decyzji dotyczących sposobu prowadzenia badań i interpretacji

danych oraz technik i narzędzi, pozwalający w intersubiektywnie sprawdzalny sposób dochodzić do nowej wiedzy naukowej. Tak rozumiana metoda nie musi mieć charakteru powtarzalnego i może być oryginalnym efektem splotu różnych czynników i decyzji. Z kolei **technika** badań będzie rozumiana jako w dużym stopniu powtarzalny i standardowy zestaw dyrektyw i procedur, pozwalający na zbieranie danych o rzeczywistości, ich odkrywanie lub kreowanie w zależności od przyjmowanego poglądu na istotę procesu poznawczego.

### 3. Zakres i zasięg badań oraz metoda

Przyjęto, że przedmiotem zainteresowania będą metody badań, których wyniki są upubliczniane w głównym nurcie globalnej komunikacji naukowej, a zatem te, które mają szczególnie duży potencjał wpływu na kierunki rozwoju badań. W związku z tym wyszukiwanie publikacji (materiału badawczego) przeprowadzono w języku angielskim w bazach *Web of Science* (21 sierpnia 2015 r.). W celu „wyłowienia” tylko tych tekstów, w których problematyka PIM stanowi centralny obiekt zainteresowania (a zatem metoda badań projektowana była właśnie dla badań PIM), zdecydowano się skupić wyłącznie na publikacjach, w których termin *personal information management* występuje w tytule. Choć takie zawężenie prowadzi do pominięcia części publikacji, w których tematyka PIM była poruszana (ale niekoniecznie jako podstawowy obiekt badań), to daje dużą gwarancję precyzji wyszukanego zbioru z punktu widzenia celu niniejszego artykułu, w którym nie chodzi o kompletną czy statystyczną analizę wszystkich tekstów, w których różne zagadnienia PIM się pojawiały, a o jakościowe rozpoznanie rozwiązań stosowanych wyłącznie w badaniach dedykowanych problematyce PIM. Przyjęto również, że problematyka PIM może być przedmiotem zainteresowania różnych dyscyplin i dlatego wyszukiwanie przeprowadzono w całych zasobach *Web of Science Core Collection: Citation Indexes*. Założenie to okazało się słuszne: wyszukano łącznie 89 publikacji zaklasyfikowanych w *Web of Science* do różnych dyscyplin, w tym przede wszystkim informatyki, informatologii i bibliotekoznawstwa oraz m.in. do ergonomii, nauk pedagogicznych, psychologii czy zarządzania. Nie zastosowano żadnych ograniczeń czasowych kwerendy. Najstarsza wyszukana publikacja ukazała się w 1985 r. Do pełnych tekstów siedmiu publikacji nie udało się uzyskać dostępu.

Jak wspomniano wcześniej, przyjęto podejście jakościowe, rezygnując z analiz ilościowych. Materiał został dobrany bez dążenia do osiągnięcia jego reprezentatywności, a zatem uzyskane wnioski dotyczą właśnie tego materiału i możliwe jest istnienie jeszcze innych rozwiązań metodologicznych w badaniach PIM, niezidentyfikowanych w toku przeprowadzonej analizy.

Badania opierały się na analizie i krytyce treści publikacji (Cisek, 2010) i wykonywane były w sposób systematyczny w odniesieniu do każdego tekstu. Po pierwsze, dążono do stwierdzenia, czy dana publikacja została oparta na oryginalnych badaniach empirycznych (dotyczących rzeczywistości, a nie publikacji na temat PIM) i czy rzeczywiście dotyczyła problematyki PIM. Jeśli nie, nie była już dalej brana pod uwagę. Następnie, w każdym przypadku dążono do rozpoznania sześciu elementów, co nie zawsze było działaniem oczywistym ze względu na jakość i kompletność informacji dostępnych w badanych publikacjach:

- ogólny cel badań,
- ujęcie w czasie,

- środowisko badań,
- rodzaj zbieranych danych i sposób ich interpretacji,
- zastosowane techniki badawcze,
- dobór uczestników badań.

Zarzucono natomiast planowaną analizę głębszych założeń metateoretycznych oraz wykorzystywanych narzędzi. Okazało się bowiem, że po pierwsze, autorzy nie sygnalizowali tak głębokich założeń, a próby ich „wyinterpretowania z tekstu” oznaczałyby duże odejście od materiału empirycznego i wymagałyby dodatkowego zastosowania innych technik (w tym szczególnie wywiadu), a po drugie, badacze często nie ujawniali precyzyjnych informacji o stosowanych narzędziach (treści ankiet, strukturyzowanych części wywiadów itp.).

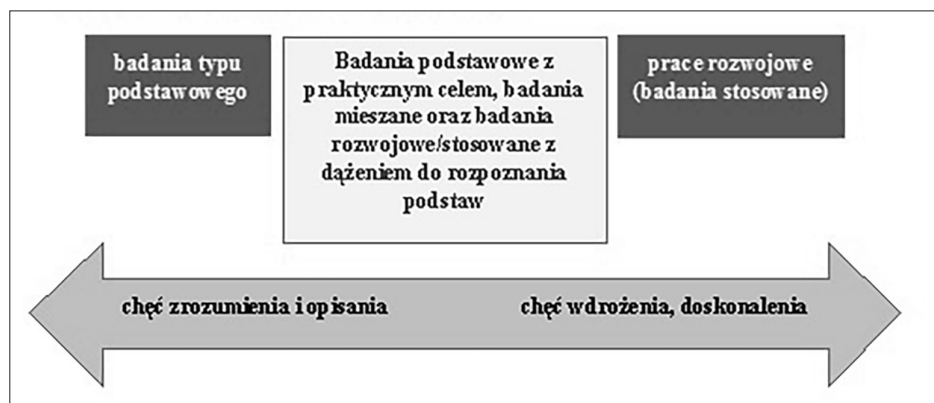
## 4. Rozwiązania metodologiczne w badaniach PIM

### 4.1. Ogólny cel badań

Część analizowanych badań nakierowana była na „czyste” poznanie (odkrywanie, opisanie) wybranych aspektów PIM (np. Barreau, 1995; Bergman, 2013; Huvila et al., 2014; McAlpine et al., 2011). Jeśli szukać dla nich określenia w istniejącej typologii badań, to należałoby je nazwać badaniami podstawowymi, choć nie były nastawione na tworzenie szerszych teorii wyjaśniających. Autorzy skupiali się raczej na opisie zjawisk, typologizacji czy próbach wykrycia prostych relacji (prawidłowości). Z drugiej strony zidentyfikowano grupę badań, których celem było wskazanie konkretnych, właściwych sposobów postępowania (np. Helvoort, 2012; Wang et al., 2012; Zhou et al., 2012), przy czym dotyczyły one wąsko określonych działań wspierających rozwój lub projektowanie konkretnych narzędzi PIM, usprawniania rozwiązań wspomagających PIM lub narzędzi badań PIM. Przybierały nawet postać prostego testowania (ang. *test*, *trial*, *evaluation*) prototypów i nowych rozwiązań w celu dowiedzenia ich przydatności lub wyeliminowania błędów (Lehmann et al., 2008; Thint et al., 2003; Van Kleek et al., 2010). Można je, w pewnym przybliżeniu, określić mianem badań stosowanych lub wręcz pracami rozwojowymi.

Te dwie grupy wyznaczają skrajne podejścia. Najciekawsze jednak wydaje się to, co jest pomiędzy nimi (Rys. 1). Zaliczyć tutaj należy te badania, w których wprawdzie stawiane są ograniczone cele podstawowe, ale jedną z kluczowych przyczyn ich podjęcia są, jak deklarują autorzy, możliwości wykorzystania nowej wiedzy w przyszłej działalności praktycznej, w tym m.in. do projektowania lepszych narzędzi i usług wspierających PIM (np. Bergman et al., 2008b; Bergman et al., 2013; Jones & Thomas, 1997; Majid et al., 2010; Otopah & Dadzie, 2013). W jeszcze innych projektach wyraźnie obecne są oba podejścia jednocześnie, a badaniami empirycznymi objęto zarówno cele czysto poznawcze, jak i praktyczne (jak jest i co zrobić, żeby było lepiej) (np. Bergman et al., 2008a; Bernstein et al., 2008; Elswiler et al., 2007; Golemati et al., 2007). Znalezione też badania, których autorzy wprawdzie wychodzą od koncepcji badań stosowanych, ale niejako przy okazji szukają także odpowiedzi na pytania podstawowe (np. Yeo & Lee, 2014).

Rys. 1. Rodzaje badań PIM ze względu na cel



#### 4.2. Ujęcie w czasie

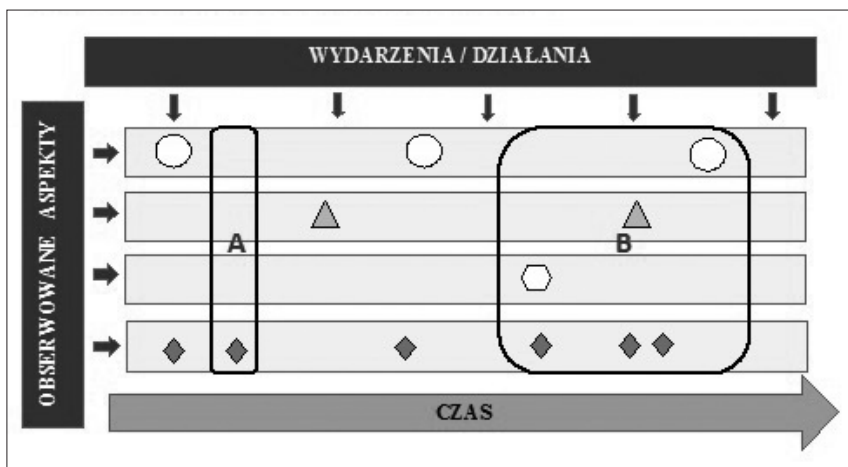
W niektórych przypadkach przedział czasowy nie był określany, ale analiza takich publikacji sugeruje, że badaczom chodziło o stan aktualny, jakkolwiek owa „aktualność” byłaby rozumiana, a tego typu badania przypominają robienie statycznej fotografii. Przykładem mogą być badania zmierzające do ustalenia, jakie narzędzia PIM są aktualnie używane, w jaki sposób są wykorzystywane, jakie kompetencje w tym zakresie reprezentuje wybrana grupa respondentów, czy jakieś działania w ramach PIM są preferowane przez konkretne osoby w określonych kontekstach (np. Bergman et al., 2003; Jacques & Fastrez, 2014; McAlpine et al., 2011; Otopah & Dadzie, 2013; Zhou et al., 2012). Takie podejście, w którym chodzi o poznanie lub uzyskanie konkretnego stanu, a nie o odkrywanie związków przyczynowo-skutkowych czy zmian w czasie, można określić mianem synchronicznego.

W innych przypadkach badano wybrane zagadnienia w dłuższym okresie, w trakcie ich „dziania się”, a przedmiotem były procesy, zmiany zachodzące w czasie, relacje następstwa itp. Przykładem może być badanie zmian PIM pod wpływem rozwoju technologii czy analiza działań w zakresie PIM w trakcie realizowania określonych w czasie projektów (np. Tungare, 2007; Bergman et al., 2008b). W przeciwieństwie do badań synchronicznych, można je określić jako diachroniczne. Specyficznym przypadkiem są badania, w których wprawdzie autorzy wykonują typowo synchroniczną analizę stanu wybranego aspektu PIM, ale z zamiarem porównania do wcześniejszych wyników (Barreau, 1995).

Granica między tymi dwoma podejściami bywa jednak czasem trudno uchwytna ze względu na „rozciąganie” badań synchronicznych w czasie, co jest związane przede wszystkim z nieprzewidywalnością i nieregularnością części zdarzeń w zakresie PIM i wydłużaniem obserwacji w celu zarejestrowania szerszej gamy zjawisk czy zachowań (Rys. 2). Dane do tego typu badań synchronicznych mogą być zbierane przez tydzień (np. Elswiler et al., 2007), a nawet miesiąc i dłużej (np. Abeta & Kakizaki, 1998; Bergman et al., 2013; Lehmann et al., 2008; Yeo & Lee, 2014), a tym samym pozwalają uchwycić zjawiska i zdarzenia, tak jak one pojawiają się w czasie, co przybliża je do ujęcia diachronicznego. Mimo to, tak długo jak przedmiotem rozważań nie są relacje następstwa czy zmiana, a jedynie

sam fakt wystąpienia pewnych zjawisk lub ich charakterystyka, badania takie zachowują synchroniczny charakter.

Rys. 2. Okres obserwacji w podejściu synchronicznym



A – krótki okres obserwacji uniemożliwia uchwycenie wszystkich zdarzeń

B – wydłużony okres obserwacji zwiększa szansę na dostrzeżenie szerszego spektrum zdarzeń

#### 4.3. Środowisko badań

Niestety, w analizowanym piśmiennictwie brakuje refleksji na temat środowiska badań i konsekwencji jego wyboru do uzyskiwanych wyników. Wyjątkiem okazał się krótki tekst na temat badań osób, których praca wymaga częstego przemieszczania się i jest realizowana w różnych miejscach (Thomson & Jarrahi, 2014) oraz badania, w których te same zachowania były przedmiotem obserwacji w dwóch różnych środowiskach rozumianych w kategoriach wykorzystywanego oprogramowania, by w jakimś stopniu wyeliminować jego wpływ na badane preferencje i umożliwić wyższy poziom generalizacji wyników (Bergman et al., 2013). Zidentyfikowano następujące środowiska badań:

- środowisko zawodowe lub edukacyjne uczestników badań (informatorów, ang. *informants*) rozumiane jako fragment realnego świata, wyposażone w komputer i inne narzędzia zarządzania informacją,
- środowisko prywatne respondentów rozumiane jak wyżej,
- fragment wirtualnej przestrzeni cyfrowej ograniczany do wybranego środowiska softwarowego, konkretnych programów, serwisów i zasobów cyfrowych (np. Bergman, 2013; Yeo & Lee, 2014),
- laboratorium – kontrolowane, zaprojektowane przez badaczy środowisko, w którym prowadzono testy aplikacji komputerowych (np. Zhou et al., 2012) lub, w połączeniu z badaniami terenowymi, badania zachowań ludzi (np. Bernstein et al., 2008).

Bywa, że badacze traktują wycinkowo zdefiniowane środowisko badań (np. biuro, program do zarządzania pocztą elektroniczną) jako środowisko naturalne informatorów i rezygnują z kontroli wpływu czynników środowiskowych na uzyskiwane wyniki. Tymczasem

PIM osadzone jest w różnych przestrzeniach i kontekstach życia, a ich ograniczanie (wybór) do celów badawczych stanowi istotną ingerencję w zachowania jednostki, szczególnie gdy badania dotyczą właśnie tych zachowań. Z jednej strony, różne funkcje PIM mogą być realizowane przez jedną osobę w różnych miejscach (np. skanuje materiały w bibliotece, a opracowuje w domu), z drugiej, jedna osoba może w różnych miejscach prowadzić odrębne systemy PIM (np. w pracy sprawy zawodowe, w domu sprawy rodziny), a z jeszcze innej – część jednego środowiska może stanowić również część innego (np. laptop na biurku w domu i ten sam laptop po przeniesieniu go na biurko w pracy).

#### **4.4. Rodzaj zbieranych danych i sposób ich interpretacji**

Z jednej strony, napotkano badania bazujące na danych liczbowych i ich statystycznej obróbce, których celem było np. określenie powszechności różnych zachowań w ramach PIM, popularności i sposobów wykorzystywania określonych narzędzi (np. Majid et al., 2010; Otopah & Dadzie, 2013) czy częstotliwości występowania różnych kompetencji informacyjnych (np. Truyen, 2010). Z drugiej strony, odnaleziono badania oparte przede wszystkim na danych opisowych, w tym zmierzające do ustalenia zmiennych, które powinny być przedmiotem badań w zakresie PIM (Bergman, 2013), mające na celu zidentyfikowanie i opisanie profili zachowań w ramach PIM (Ferran-Ferrer & Pérez-Montoro, 2009) czy zdefiniowanie części kompetencji niezbędnych do sprawnego zarządzania informacją (Jacques & Fastrez, 2014).

Trudno byłoby jednak dokonać ostrego podziału analizowanych badań na ilościowe i jakościowe. Okazało się, że niektórzy autorzy posługiwali się w ramach jednego badania różnymi rodzajami danych (np. wrażenia i oceny respondentów wyrażone w języku naturalnym łączono z ich sparametryzowaną analizą ilościową – np. Golemati et al., 2007), a ich interpretacje ilościowe oraz jakościowe wzajemnie się uzupełniały, prowadząc do uzyskania pełniejszego obrazu badanego zjawiska. Przykładem mogą być badania, w których wykorzystano zarówno predefiniowane mapy uzupełniane przez respondentów subiektywnymi treściami oraz ich indywidualne wypowiedzi pozyskiwane w trakcie wywiadów, jak i dane ilościowe dotyczące np. popularności różnych narzędzi wspomagających PIM (McAlpine et al., 2011). W innych badaniach połączono jakościowe analizy praktyk realizowanych w ramach PIM z ilościowymi wynikami ankiety na temat stosowanych sposobów indywidualnego zarządzania cyfrowymi artykułami naukowymi (Huvila et al., 2014). W niektórych badaniach łączono wstępne jakościowe rozpoznanie problemu z późniejszymi analizami ilościowymi występowania określonych cech czy zachowań (w pewnym stopniu np. Van Kleek et al., 2010), a w innych dane ilościowe i jakościowe zbierano łącznie w ramach jednej techniki – np. w technice dzienniczkowej, gdzie proszono nie tylko o swobodne wypowiedzi, ale także o odnotowywanie wystąpień policzalnych faktów i określanie wielkości uprzednio zdefiniowanych parametrów liczbowych (Elsweiler et al., 2007).

#### **4.5. Techniki badawcze**

W badaniach przedstawionych w analizowanych tutaj publikacjach zastosowano następujące techniki (w kolejności alfabetycznej):

- analiza artefaktów rozumiana tutaj jako systematyczny ogląd i analiza notatek, uwag i innych drobnych porcji utrwalonej informacji (ang. *information scraps*)

sporządzanych przez człowieka w celu zanotowania jakichś treści, które nie są przedmiotem zarządzania w ramach typowych narzędzi wspomagających PIM (Bernstein et al., 2008). Technika częściowo zbieżna z techniką analizy zawartości lub tekstu (w jej jakościowej wersji), ale poszerzona o badania formy i postaci nośnika informacji i jego powiązań z infrastrukturą informacyjną, tutaj wykonywana w połączeniu z wywiadem;

- analiza logów rozumiana jako analiza śladów zdarzeń i działań wykonywanych przez konkretne jednostki, zarejestrowanych i zachowanych w pamięci komputera za pomocą narzędzi softwarowych (Bergman et al., 2008b);
- analiza zawartości (ang. *content analysis*) rozumiana jako ilościowa lub jakościowa, kontekstowa analiza szeroko rozumianych komunikatów, dotyczyła nie tyle samej zawartości dokumentów, co aparatu ich organizacji (nazwy i zawartość folderów, tagi) (np. Bergman et al., 2013, zob. też opis techniki wizji lokalnej i techniki oprowadzania);
- ankieta online lub bezpośrednia, stosowana:
  - jako technika autonomiczna, także w formie dwóch kwestionariuszy mierzących zmianę (stan przed i po doświadczeniu – Tungare, 2007),
  - technika dodatkowa, kierowana do uczestników badań prowadzonych innymi technikami w celu zebrania informacji o samych respondentach (np. Elswiler et al., 2007) lub w celu uzupełnienia danych zbieranych w inny sposób, np. w ramach eksperymentów (np. Golemati et al., 2007; Van Kleek et al., 2010; Zhou et al., 2012);
- eksperyment rozumiany jako technika zmierzająca do zebrania danych poprzez odpowiednie manipulowanie zmiennymi w doświadczeniu naukowym wykonywanym w warunkach kontrolowanych (Yeo & Lee, 2014);
- obserwacja ustrukturyzowana (standaryzowana) lub częściowo ustrukturyzowana, w której badacz obserwuje działania jednostki i skupia się na rejestrowaniu wyłącznie lub między innymi zdarzeń przewidzianych przed podjęciem badań. Stosowana jako jedna z wielu technik użytych w jednym badaniu (np. Bergman et al., 2013; Golemati et al., 2007). W praktyce zbliżona do eksperymentu, choć w odróżnieniu od niego ma na celu przede wszystkim poznanie samego zjawiska lub przedmiotu (traktowanego jako danego, bez jego modyfikowania), a nie relacji między poszczególnymi zmiennymi, czy skutków ich zastosowania;
- oprowadzanie, zwiedzanie z przewodnikiem (ang. *guided tour*) – technika polegająca na jednoczesnym wywiadzie i oglądzie indywidualnego warsztatu informacyjnego danej jednostki (tutaj przede wszystkim w środowisku cyfrowym, ale nie tylko) oczami samego informatora (uczestnika badań), który po swoich zasobach i procedurach „oprowadza” badacza. Technika traktowana też jako forma wywiadu (Bergman, 2013, 466), czasem połączona z utrwalaniem obserwacji w celu późniejszej analizy zawartości;
- przegląd systematyczny (ang. *walkthrough*) narzędzi softwarowych rozumiany jako testowanie i ocena funkcji programów użytkowych wykonywanych z perspektywy ekspertów według ustalonego planu (Chengxia & Minling, 2013; Tramullas et al., 2009). Technika zbliżona do „wizji lokalnej”, ale wykonywana w odniesieniu do wybranych funkcjonalności określonego oprogramowania;

- technika dzienniczkowa polegająca na samoobserwacji badanego i indywidualnym zapisywaniu (w dzienniczku) na bieżąco określonych zdarzeń w przyjętym z góry okresie. Technika ta wystąpiła w postaci częściowo standaryzowanej, w której ową standaryzację narzucała świadomie predefiniowana przez badaczy forma dzienniczka (Elsweiler et al., 2007);
- testy (ang. *tests, trials*) traktowane jako technika oceny różnych cech (np. skuteczności, funkcjonalności itp.) i działania narzędzi softwarowych (np. Abeta & Kakizaki, 1998; Dengel & Adrian, 2011; Thint et al., 2003; Van Kleek et al., 2010; Wang et al., 2012; Zhou et al., 2012). W niektórych przypadkach testy były poprzedzane instruktażem czy zapoznawczym okresem użytkowania prototypu przez uczestników badań (np. Bergman et al., 2008b; Elsweiler et al., 2007), a same testy bazowały na standaryzowanych lub indywidualnych zbiorach danych i określonych z góry scenariuszach działań albo prowadzone były w toku typowego użytkowania oprogramowania (np. Lehmann et al., 2008);
- testy psychologiczne (test ukrytych figur – ang. *embedded figure test*, test pamięci – ang. *memory test*) jako techniki wspomagające badania zachowań jednostek w ramach PIM (Tungare, 2007);
- wizja lokalna (różnie określana w analizowanych badaniach: ang. *natural observations, field study*) rozumiana jako ogląd indywidualnej przestrzeni informacyjnej (realnej lub cyfrowej) lub jej fragmentów, w niektórych przypadkach połączona z różnymi sposobami jej utrwalenia (zrzuty ekranowe, filmy itp.) w celu późniejszej analizy, co w dalszej kolejności może prowadzić do analizy zawartości (treści). Stosowana tutaj w połączeniu z wywiadem i różniąca się od techniki oprowadzania perspektywą oglądu – w przypadku wizji lokalnej jest to perspektywa badacza (np. Bergman et al., 2008b; Bergman et al., 2013);
- wywiad jawny, bezpośredni, indywidualny (w jednym przypadku zidentyfikowano nieformalny wywiad grupowy w pracach nad rozwojem prototypu – Van Kleek et al., 2010), w postaci częściowo ustrukturyzowanej, w której respondenci odpowiadali na stosunkowo nieliczne pytania przygotowane przez badacza, ale poza tym mieli dużą swobodę w zakresie formułowania wypowiedzi, albo nieustrukturyzowany, zogniskowany lub pogłębiony, gdzie zapytania stawiane przez badacza miały służyć nakierowaniu rozmowy na tematy właściwe z punktu widzenia celu badań. Wywiad występował jako samodzielna technika lub jako technika zespolona, w tym z obserwacją w technice oprowadzania (np. Barreau, 1995; Bergman, 2013; Jacques & Fastrez, 2014; Jones et al., 2005), techniką mapowania myśli (McAlpine et al., 2011) i techniką dzienniczkową (Elsweiler et al., 2007).

#### 4.6. Dobór uczestników badań

Analiza tego aspektu badań była utrudniona z powodu często nieprecyzyjnych informacji na ten temat lub wręcz ich braku (np. Abeta & Kakizaki, 1998). Stwierdzono wyłącznie nielosowe (nieprobabilistyczne) sposoby doboru uczestników badań. Badaczom raczej nie chodziło o próbę rozumianą jako reprezentatywna część większej populacji i statystyczne wnioskowanie na jej temat (może poza badaniami prowadzonymi z udziałem wyjątkowo licznych grup: Otopah & Dadzie, 2013 – ponad 200 respondentów; Truyen, 2010 – ponad

7000), co raczej o grupę uczestników niezbędną do przetestowania oprogramowania, przeanalizowania jakiegoś zjawiska czy obserwacji możliwych działań w ramach PIM. Takie podejście wydaje się odzwierciedlać idiograficzny charakter problematyki PIM. Poza wspomnianymi wyjątkami, w badaniach uczestniczyło od kilku do kilkudziesięciu uczestników. Stosowano dobór celowy lub tzw. dobór wygodny (ang. *convenience sampling*).

W pierwszym przypadku badacze dobierali uczestników według wybranych cech (np. zawodu, roli życiowej, znajomości określonych technologii), uciekając się czasem do formowania grup maksymalnie zróżnicowanych, by bardziej kompleksowo uchwycić badany problem. Jednak kryteria owego różnicowania miały raczej intuicyjny charakter, a przynajmniej nie były przedstawiane jako zmienne w istotny sposób kształtujące dane zjawisko. Były dobierane specyficznie dla każdego z tego typu badań (np. Barreau, 1995; Bernstein et al., 2008; Elswiler et al., 2007; McAlpine et al., 2011). Szczególnym przypadkiem celowego różnicowania uczestników badań była próba uchwycenia analizowanego aspektu PIM z wielu perspektyw poprzez przeprowadzenie badań na różnych grupach w ramach trzech studiów przypadków (Huvila et al., 2014).

Z kolei dobór wygodny pojawiał się w badaniach, w których uczestnicy traktowani byli nie jako przedmiot badań, ale jako ich narzędzie, szczególnie w przypadku niektórych testów konkretnych rozwiązań czy oprogramowania. Choć autorzy nie informowali precyzyjnie o zasadach doboru, wydaje się, że decydowała właśnie dostępność uczestników i łatwość przeprowadzenia badań z ich udziałem (np. Cutrell et al., 2006; Golemati et al., 2007; Zhou et al., 2012). Bywało, że podawano szczegółową charakterystykę uczestników, choć nie wiadomo, dlaczego akurat według cech (np. płeć), których wpływ na uzyskiwane wyniki nie jest ani znany, ani nie był uwzględniany w interpretacji wyników.

Stosowano też oba rozwiązania jednocześnie. Na przykład w badaniach relacji między stosowaniem folderów i tagowaniem w PIM, w pierwszym etapie zdecydowano się na osoby używające narzędzi, które były wykorzystane w badaniach. Jednak już nie wiadomo dlaczego z tak określonej populacji wybrano następnie do badań izraelskich studentów o określonej płci i wieku. Można się domyślać, że zastosowano tzw. dobór wygodny ze względu na miejsce pracy autorów (Bergman et al., 2013). Przykładów domniemanego zastosowania doboru wygodnego na drugim etapie selekcji, i to także wówczas, gdy chodziło o badania cech lub zachowań jakiejś populacji, można wskazać więcej: np. badanie możliwości dzieci na przykładzie grupy z konkretnej szkoły (Yeo & Lee, 2014), badanie kompetencji studentów jednej uczelni na podstawie wybranych osób (Helvoort, 2012).

Zauważyć można niepowtarzalność wybieranych grup, których skład był zawsze przedmiotem indywidualnych decyzji prowadzących badania. Uwagę zwraca wyraźna intuicyjność kryteriów doboru i nietraktowanie ich jako zmiennych w badaniach, przy jednoczesnym dążeniu do informowania czytelnika (precyzyjnie, ale znowu według arbitralnie dobieranych cech), kim byli uczestnicy. Być może powodem takiej sytuacji jest ciągły niedostatek wiedzy na temat roli różnych zmiennych w PIM oraz specyfika badań, które często wymagają bardzo dobrego, bliskiego kontaktu z informatorami, a także głębokiego wniknięcia w ich prywatność.

## 5. Zakończenie

Przeprowadzona analiza wskazuje na dużą swobodę badaczy w rozwiązywaniu kluczowych dylematów metodologicznych i poszukiwanie indywidualnych rozstrzygnięć na potrzeby konkretnych badań. Trudno byłoby mówić o „czystości” takich rozwiązań z perspektywy dychotomicznych podziałów stosowanych w typologii badań (podstawowe – stosowane, ilościowe – jakościowe, synchroniczne – diachroniczne). Z kolei różnorodność stosowanych technik zbierania danych odzwierciedla złożoność problemu: różne techniki pozwalają na zbieranie różnych danych i wnioskowanie na temat odmiennych aspektów PIM. Dążenie do uchwycenia owej złożoności znajduje też potwierdzenie w łączeniu różnych technik w jednym badaniu, co w przypadku badań PIM służy wzajemnemu uzupełnianiu się danych w celu osiągnięcia pogłębionego, holistycznego opisu i zrozumienia zjawisk, a nie wzajemnemu potwierdzaniu się uzyskiwanych wyników czy testowania tej samej hipotezy, jak w przypadku klasycznej triangulacji.

Różnorodne i specyficzne wybory i rozwiązania przekładają się na niepowtarzalność stosowanych metod (w rozumieniu przyjętym na początku tego artykułu). Ze względu na zastosowany tutaj dobór materiału badawczego, przedstawiony katalog rozwiązań poszczególnych dylematów z pewnością nie jest kompletny, ale nie taki też był cel badań. Na ich podstawie można stwierdzić, że o specyfice przyjmowanej metody każdorazowo decyduje badacz, projektując ją i biorąc za nią odpowiedzialność, odnosząc się do swoistości problemu będącego przedmiotem badań oraz możliwości ich przeprowadzenia. Przy czym wydaje się, że owo tworzenie metod nie tyle jest determinowane przez jakiś wybrany paradygmat badawczy, co są one projektowane na bazie szerokiego spektrum możliwości czerpanych z różnych podejść i perspektyw metateoretycznych. W efekcie stosowane bywa podejście ilościowe i jakościowe, interpretatywne i statystyczne, a dane zbiera się w różnych środowiskach i na różnych grupach respondentów, jednocześnie dla zrealizowania celów podstawowych i stosowanych. Taka sytuacja wymaga precyzyjnego i wyczerpującego przedstawiania metod w każdym przypadku w celu zapewnienia intersubiektywnej sprawdzalności wyników.

## Bibliografia

- Abeta, A.; Kakizaki, K. (1998). Operation record based workflow extracting method for personal information management system. In: *Proceedings. 3rd Asia Pacific Computer Human Interaction*. Los Alamitos, California: IEEE, 224–231.
- Barreau, D. (1995). Context as a factor in personal information management systems. *Journal of the American Society for Information Science*, 46(5), 327–339.
- Bergman, O. (2013). Variables for personal information management research. *ASLIB Proceedings: New Information Perspectives*, 65(5), 464–483.
- Bergman, O.; Beyth-Marom, R.; Nachmias, R. (2008a). The user-subjective approach to personal information management systems design: evidence and implementations. *Journal of the American Society for Information Science and Technology*, 59(2), 235–246.
- Bergman, O.; Beyth-Marom, R.; Nachmias, R.; Gradovitch, N.; Whittaker, S. (2008b). Improved search engines and navigation preference in personal information management. *ACM Transactions on Information Systems (TOIS)*, 26(4), article no. 20.

- Bergman, O.; Gradovitch, N.; Bar-Ilan, J.; Beyth-Marom, R. (2013). Folder versus tag preference in personal information management. *Journal of the American Society for Information Science and Technology*, 64(10), 1995–2012.
- Bergman, O.; Nachmias, R.; Beyth-Marom, R. (2003). The use of subjective attributes in personal information management systems – initial results. *Proceedings of the American Society for Information Science and Technology*, 40(1), 509–510.
- Bernstein, M.; Van Kleek, M.; Karger, D.; Schraefel, M.C. (2008). Information scraps: how and why information eludes our personal information management tools. *ACM Transactions on Information Systems (TOIS)*, 26(4), article no. 24.
- Chengxia, L.; Minling, Z. (2013). Design and implementation of personal information management system in Internet. In: L. Zhang, X. Li, J. Chen (eds.) *Proceedings of the 2013 international conference on information, business and education technology (ICIBET 2013)* [online], [08.02.2016], [http://www.atlantis-press.com/php/download\\_paper.php?id=6170](http://www.atlantis-press.com/php/download_paper.php?id=6170)
- Cisek, S. (2010). Metoda analizy i krytyki piśmiennictwa w nauce o informacji i bibliotekoznawstwie w XXI wieku. *Przegląd Biblioteczny*, 3, 273–284.
- Cutrell, E.; Dumais, S.T.; Teevan, J. (2006). Searching to eliminate personal information management. *Communications of the ACM*, 49(1), 58–64.
- Dengel, A.; Adrian, B. (2011). Helping people remember: coactive assistance for personal information management on a semantic desktop. In: A. Fred, J.L.G. Dietz, K. Liu, J. Filipe (eds.) *Knowledge discovery, knowledge engineering and knowledge management*. Berlin, Heidelberg: Springer, 3–16.
- Elsweiler, D.; Ruthven, I.; Jones, Ch. (2007). Towards memory supporting personal information management tools. *Journal of the American Society for Information Science and Technology*, 58(7), 924–946.
- Ferran-Ferrer, N.; Pérez-Montoro, M. (2009). Gestión de la información personal en usuarios avanzados en TIC. *El Profesional de la Información*, 18(4), 365–373.
- Golemati, M.; Katifori, A.; Giannopoulou, E.G.; Daradimos, I.; Vassilakis, C. (2007). Evaluating the significance of the windows explorer visualization in personal information management browsing tasks [online]. In: E. Banissi et al. (eds.) *11th International Conference Information Visualization: IV 2007*. Los Alamitos, Calif.: IEEE Computer Society, 93–100, [09.02.2016], [https://www.researchgate.net/profile/Costas\\_Vassilakis/publication/4260997\\_Evaluating\\_the\\_Significance\\_of\\_the\\_Windows\\_Explorer\\_Visualization\\_in\\_Personal\\_Information\\_Management\\_Browsing\\_tasks/links/00b4951cb36feb0f20000000.pdf](https://www.researchgate.net/profile/Costas_Vassilakis/publication/4260997_Evaluating_the_Significance_of_the_Windows_Explorer_Visualization_in_Personal_Information_Management_Browsing_tasks/links/00b4951cb36feb0f20000000.pdf)
- Gwizdka, J.; Chignell, M. (2007). Individual differences. In: W. Jones & J. Teevan (eds.) *Personal information management*. Seattle: University of Washington Press, 206–220.
- Helvoort, A.A.J. van (2012). A questionnaire for the institutional assessment of personal information management. In: S. Kurbanoglu, U. Al, Y. Tonta, N. Ucak (eds.) *E-science and information management. Third International Symposium on Information Management in a Changing World, IMCW 2012, Ankara, Turkey, September 19–21, 2012. Proceedings*. Berlin, Heidelberg: Springer, 138–149.
- Huvila, I.; Eriksen, J.; Häusner, E.M.; Jansson, I.M. (2014). Continuum thinking and the contexts of personal information management. *Information Research* [online], 19(1), [08.02.2016], [http://www.informationr.net/ir/19-1/paper604.html#VroD\\_1IhE2Y](http://www.informationr.net/ir/19-1/paper604.html#VroD_1IhE2Y)
- Jacques, J.; Fastrez, P. (2014). Personal information management competences: a case study of future college students. In: S. Yamamoto (ed.) *Human interface and the management of information: information and knowledge design and evaluation*. Cham: Springer International Publishing, 320–331.
- Jones, S.R.; Thomas, P.J. (1997). Empirical assessment of individuals' "personal information management systems". *Behaviour & Information Technology*, 16(3), 158–160.
- Jones, W. (2007). Personal information management. *Annual Review of Information Science and Technology*, 41(1), 453–504.
- Jones, W. (2008). *Keeping found things found: the study and practice of personal information management*. Amsterdam, Boston: Morgan Kaufmann Publishers.

- Jones, W.; Phuwanartnurak, A.; Gill, R.; Bruce, H. (2005). Don't take my folders away! Organizing personal information to get things done. In: *CHI '05 Extended Abstracts on Human Factors in Computing Systems*. New York: ACM, 1505–1508.
- Kelly, D.; Teevan, J. (2007). Understanding what works: evaluating PIM tools. In: W. Jones & J. Teevan (eds.) *Personal information management*. Seattle: University of Washington Press, 190–205.
- Lehmann, L.; Rensing, Ch.; Steinmetz, R. (2008). Capture of lifecycle information to support personal information management. In: P. Dillenbourg & M. Specht (eds.) *Times of convergence. Technologies across learning contexts*. Berlin, Heidelberg: Springer, 216–221.
- Majid, S.; San, M.M.; Tun, S.T.N.; Zar, T. (2010). Using Internet services for personal information management. In: S. Kurbanoglu, U. Al, P.L. Erdogan; Y. Tonta; N. Ucak (eds.) *Technological convergence and social networks in information management*. Berlin, Heidelberg: Springer, 110–119.
- McAlpine, H.Ch.; Hicks, B.; Tiryakioglu, C. (2011). The digital divide: investigating the personal information management practices of engineers. In: S.J. Culley, B.J. Hicks, T.C. McAloone, T.J. Howard, W. Chen (eds.) *DS 68–6: Proceedings of the 18th International Conference on Engineering Design (ICED 11), Impacting Society through Engineering Design, Vol. 6: Design Information and Knowledge, Lyngby/Copenhagen, Denmark, 15.-19.08. 2011* [online], 31–42, [09.02.2016], [https://www.designsociety.org/publication/30611/the\\_digital\\_divide\\_investigating\\_the\\_personal\\_information\\_management\\_practices\\_of\\_engineers](https://www.designsociety.org/publication/30611/the_digital_divide_investigating_the_personal_information_management_practices_of_engineers)
- Naurer, Ch.M.; Fisher, K.E. (2007). Naturalistic approaches for understanding PIM. In: W. Jones & J. Teevan (eds.) *Personal information management*. Seattle: University of Washington Press, 76–88.
- Otopah, F.O.; Dadzie, P. (2013). Personal information management practices of students and its implications for library services. *Aslib Proceedings*, 65(2), 143–160.
- Thint, M.; Case, S.; Azarmi, N.; Azvine, B.; Hare, S. (2003). Personal information management assistants – from research to commercialisation. *BT Technology Journal*, 21(4), 39–43.
- Thomson, L.; Jarrahi, M. (2014). Contextualising information practices and personal information management in mobile work. Information Research [online], 19(4), *Proceedings of ISIC, the Information Behaviour Conference, Leeds, 2–5 September, 2014: Part 1*, [09.02.2016], <http://InformationR.net/ir/19-4/isic/isicsp4.html>
- Tramullas, J.; Sanchez-Casabon, A.I.; Garrido-Picazo, P. (2009). Personal information management using concept mapping software. *El Profesional de la Informacion*, 18(6), 601–612.
- Truyen, F. (2010). Personal information management: the information companion. In: *EDULEARN 10 Conference Proceedings* [online], 5399–5406, [09.02.2016], [https://www.researchgate.net/profile/Frederik\\_Truyen/publication/265694460\\_Personal\\_Information\\_Management\\_The\\_Information\\_Companion/links/5592ac8308aed7453d4635fb.pdf](https://www.researchgate.net/profile/Frederik_Truyen/publication/265694460_Personal_Information_Management_The_Information_Companion/links/5592ac8308aed7453d4635fb.pdf)
- Tungare, M. (2007). Understanding the evolution of users' personal information management practices. In: C. Baranauskas, P. Palanque, J. Abascal, S. Diniz, J. Barbosa (eds.) *Human-Computer Interaction – INTERACT 2007. 11th IFIP TC 13 International Conference, Rio de Janeiro, Brazil, September 10–14, 2007, Proceedings, Part II* [online]. Berlin, Heidelberg: Springer, 586–591, [09.02.2016], [http://link.springer.com/content/pdf/10.1007%2F978-3-540-74800-7\\_56.pdf](http://link.springer.com/content/pdf/10.1007%2F978-3-540-74800-7_56.pdf)
- Van Kleek, M.; André, P.; Moore, B.; Karger, D. (2010). Mixing the reactive with the personal: Opportunities for end-user programming in personal information management. In: A. Cypher, M. Dontcheva, T. Lau, J. Nichols (eds.) *No code required: giving users tools to transform the web*. Amsterdam, Boston: Elsevier, Morgan Kaufmann, 127–152.
- Wang, W.; Peery, C.; Marian, A.; Nguyen, T.D. (2012). Efficient multidimensional fuzzy search for personal information management systems. *IEEE Transactions on Knowledge and Data Engineering*, 24(9), 1584–1597.
- Whittaker, S. (2011). Personal information management: from information consumption to curation. *Annual Review of Information Science and Technology*, 45(1), 3–62.

- Yeo, H.I.; Lee, Y.L. (2014). Exploring new potentials of blogs for learning: can children use blogs for personal information management (PIM)? *British Journal of Educational Technology*, 45(5), 916–925.
- Zhou, L.; Mohammed, A.S.; Zhang, D. (2012). Mobile personal information management agent: supporting natural language interface and application integration. *Information Processing & Management*, 48(1), 23–31.
- 

## Research Methodology of Personal Information Management Studies: Selected Aspects

### Abstract

**Purpose/Thesis:** The aim of the research the results of which are presented in this article was to identify the ways of solving essential dilemmas in the development of research in the field of personal information management (PIM).

**Approach/Research methods:** The author applied a qualitative approach. The analysis involved publications available in Web of Science Core Collection: Citation Indexes (retrieved on August 21, 2015) the titles of which contained the term personal information management and were a discussion of any empirical studies on PIM. The main objective was to identify six elements: general research goals, time coverage, research circumstances, types of data gathered and the ways of their interpretation, research techniques applied and the ways of participant selection for such studies.

**Results and conclusions:** The analysis shows a broad variety of solutions derived from different metatheoretical perspectives and reveals key methodological decisions made to design PIM studies.

**Originality/Value:** The results of this research may help to develop future research and inspire an in-depth discussion on various methodological aspects of PIM research.

### Keywords

Personal information management. Methodology. Research design. Research techniques.

---

*Dr hab. REMIGIUSZ SAPA jest dyrektorem Instytutu Informacji Naukowej i Bibliotekoznawstwa Uniwersytetu Jagiellońskiego i kierownikiem działającego w nim Zakładu Metodologii Bibliotekoznawstwa i Informacji Naukowej. Prowadzi badania na temat współczesnej komunikacji naukowej, zachowań informacyjnych i zarządzania informacją. Podejmuje też problematykę metodologii badań informatologicznych i funkcjonowania bibliotek akademickich. Ostatnio opublikował: *Information seeking behaviour of mathematicians: scientists and students*. *Information Research*, 2014, 19(4), paper 644 (współaut. M. Krakowska, M. Janiak); *Wsparcie dla indywidualnego zarządzania informacją w polskich repozytoriach uczelnianych*. W: A. Korycińska-Huras, M. Janiak (red.). *Komunikacja naukowa w środowisku cyfrowym: badania, zasoby, użytkownicy*. Warszawa: Wydaw. SBP, 2014, s. 81–96.*

*Kontakt z autorem:*

*remigiusz.sapa@uj.edu.pl*

*Instytut Informacji Naukowej i Bibliotekoznawstwa*

*Uniwersytet Jagielloński*

*ul. prof. Stanisława Łojasiewicza 4*

*30–348 Kraków*

# Cytowania zasobów sieciowych w polskich czasopismach z zakresu bibliotekoznawstwa i informatologii: analiza aktualności adresów URL

Marcin Roszkowski

*Katedra Informatologii, Wydział Dziennikarstwa, Informacji i Bibliologii,  
Uniwersytet Warszawski*

Bartłomiej Włodarczyk

*Katedra Bibliografii i Dokumentacji, Wydział Dziennikarstwa, Informacji i Bibliologii,  
Uniwersytet Warszawski*

---

## Abstrakt

**Cel/teza:** Celem artykułu jest charakterystyka stanu badań nad problemem dostępności odwołań do zasobów sieciowych obecnych w aparacie pomocniczym tekstów naukowych oraz przedstawienie wyników badań własnych związanych z tym zagadnieniem w odniesieniu do artykułów naukowych z czterech polskich czasopism z zakresu bibliotekoznawstwa i informatologii.

**Koncepcja/metody badań:** Wykorzystując krytyczną analizę piśmiennictwa naukowego zidentyfikowano grupy zagadnień szczegółowych będących przedmiotem zainteresowania badaczy. Wskazano również na metodologiczne aspekty badania tego zjawiska oraz nakreślono jego obraz w odniesieniu do poszczególnych dyscyplin naukowych. Przedmiotem badań empirycznych były artykuły naukowe opublikowane w czterech polskich czasopismach z zakresu bibliotekoznawstwa i informatologii w latach 2009–2015: *Bibliotheca Nostra*, *Śląski Kwartalnik Naukowy*, *Biuletyn EBIB*, *Przegląd Biblioteczny*, *Zagadnienia Informacji Naukowej – Studia Informacyjne*. Dostępność adresów URL w cytowaniach zbadano metodą automatyczną za pośrednictwem kodów odpowiedzi serwerów HTTP.

**Wyniki i wnioski:** W czasopismach uwzględnionych w badaniu ok. 37% wszystkich pozycji w bibliografiach załącznikowych i przypisach zawierało łącza do zasobów internetowych. Wyniki badań nad dostępnością zasobów sieciowych cytowanych w korpusie analizowanych publikacji wskazują, że ponad 70% odwołań jest nadal aktywnych w momencie przeprowadzenia badania. Wartość tej zmiennej jest zróżnicowana dla poszczególnych tytułów. Wskaźnik połowicznej dezaktywacji łączy dla czterech zbadanych tytułów wynosi ok. siedmiu lat.

**Oryginalność/wartość poznawcza:** Przedstawiony stan badań związanych z aktualnością odwołań do zasobów sieciowych w aparacie pomocniczym publikacji naukowych może być punktem odniesienia dla badań szczegółowych związanych z tym zagadnieniem i być źródłem rozwiązań metodologicznych. Problem ten nie był według wiedzy autorów eksplorowany w odniesieniu do polskich czasopism z zakresu bibliotekoznawstwa i informatologii. Tym samym przedstawione wyniki są pierwszą próbą charakterystyki skali tego zjawiska.

## Słowa kluczowe

Analiza cytowań. Polskie czasopisma naukowe. Czasopisma z zakresu bibliotekoznawstwa i informatologii. Wskaźnik połowicznej dezaktywacji łączy. Cytowania zasobów sieciowych.

Otrzymano: 3 lipca 2016. Zrecenzowano: 12 lipca 2016. Zaakceptowano: 18 lipca 2016.

---

## 1. Wprowadzenie

Przypisy bibliograficzne oraz bibliografie załącznikowe stanowią istotny element pracy naukowej. Służą one dokumentacji postawionych tez i powinny umożliwić innym badaczom sprawdzenie poprawności powołań i całego wyводу naukowego. W związku ze wzrostem znaczenia zasobów sieci World Wide Web (WWW) w obiegu informacji w nauce, odwołania do zasobów sieciowych są niemal stałym elementem w bibliografiach załącznikowych publikacji naukowych. Istotnym problemem w przypadku tych materiałów pozostaje kwestia ich dostępności i stabilności. Badania nad aktywnością hiperłączy i trwałością zasobów internetowych rozpoczęły się wkrótce po powstaniu WWW, w połowie lat 90. XX w. Wiązały się one ze świadomością, że WWW nie tylko rozrasta się, ale także część jego zasobów znika, przestaje być dostępna (tzw. zjawisko *link rot*) pod podanym adresem Uniform Resource Locator (URL). W części zaś przypadków zmianie ulega treść stron internetowych (ang. *content drift*). Badania prowadzone na przełomie XX i XXI w. wskazywały na dużą skalę zjawiska zanikania zasobów WWW. Wallace Koehler pokazał na losowej próbie 361 stron WWW, zebranej w grudniu 1996 r., że problem zwiększa się wraz z upływem czasu. W styczniu 2001 r. aktywne były tylko 124 strony (34.4% stanu wyjściowego), a w maju 2003 r. – 122 (33.8%). Badacz zaobserwował zjawisko wznawiania działalności stron, które przez pewien okres były nieaktywne (Koehler, 2004). Najnowszy przegląd wspomnianej próbki stron WWW pokazał, że pod koniec 2013 r. aktywnych było tylko 6 łączy (1.6% stanu wyjściowego), a w lutym 2015 r. na zapytanie odpowiedziały tylko dwie strony (Oguz & Koehler, 2016, 477).

Od początku tego wieku przeprowadzono liczne badania dotyczące łączy do zasobów internetowych zamieszczanych w przypisach bibliograficznych i bibliografiach załącznikowych publikacji naukowych. W polskiej literaturze z zakresu bibliotekoznawstwa i informatologii (BIN) brakuje badań poświęconych tej problematyce. Artykuł ma na celu uzupełnienie tej luki w polskim piśmiennictwie naukowym. Postawiono sobie w nim dwa cele, zaprezentowanie stanu badań oraz przedstawienie wyników badania cytowania zasobów sieciowych w wybranych czasopismach ze wspomnianej dziedziny.

Artykuł został podzielony na dwie części. W części pierwszej przedstawiono stan badań nad cytowaniem zasobów sieciowych oraz propozycje rozwiązań problemów dotyczących trwałości cytowanych zasobów internetowych. Skoncentrowano się na badaniach dotyczących literatury z zakresu BIN. Następnie w drugiej części artykułu zaprezentowano metodologię badania oraz uzyskane rezultaty. W zakończeniu wskazano na propozycje rozwiązań, które warto poddać pod dyskusję w kontekście wdrożeń w polskim środowisku wydawców i autorów prac naukowych.

## 2. Stan badań

W ramach badań nad łącami internetowymi będącymi elementami bibliografii załącznikowych można wyróżnić dwie podstawowe grupy. Pierwsza z nich zawiera badania ilościowe przypisów i bibliografii załącznikowych zawierających łącza do zasobów internetowych, druga zaś koncentruje się na propozycjach rozwiązań zapewniających większą trwałość cytowanych zasobów.

Jednym z celów badań sytuujących się w ramach grupy pierwszej jest wskazanie liczby aktywnych i nieaktywnych łączy oraz przyczyn tego stanu. Badania dotyczą linków zamieszczonych w jednym czasopiśmie, kilku czasopismach z jednej dziedziny lub porównania charakterystyki łączy w czasopismach z różnych dziedzin. W ramach tej grupy badań prowadzone są również prace dotyczące zmiany treści zasobów w czasie. Przywoływane strony WWW mogą być dostępne, jednak ich zawartość może różnić się od prezentowanej w momencie sporządzania przypisu. Wallace Koehler i Fatih Oguz, którzy odróżnili badania dotyczące aktywności łączy od tych skupiających się na zmianie treści zasobów, podkreślili, że dużo więcej jest prac koncentrujących się na pierwszym zagadnieniu (Oguz & Koehler, 2016, 477).

W pierwszej grupie badań można wyróżnić następujące podstawowe zagadnienia badawcze:

- (1) Liczba cytowań zasobów internetowych.
- (2) Stosunek liczby łączy nieaktywnych do aktywnych (w tym zmiana treści przywoływanych zasobów).
- (3) Rodzaje odpowiedzi serwera.
- (4) Wpływ długości adresu URL na czas aktywności łączy.
- (5) Wpływ rodzaju domeny na czas aktywności łączy.

Wyniki badań wskazują na wzrost liczby powołań na zasoby internetowe w czasie. Dotyczy to zarówno ogólnej liczby cytowań materiałów sieciowych, jak i średniej liczby cytowań tych zasobów na artykuł. Nie jest to jednak wzrost liniowy. Niektóre badania wskazują na zahamowanie szybkiego trendu wzrostowego zaobserwowane w ostatnich latach. Podobne wyniki uzyskano badając publikacje z różnych dziedzin nauki. W porównywaniu wyników badań istotnym czynnikiem jest czas, z którego pochodzą analizowane dane. W związku z tym obok wyników zamieszczono w nawiasach okrągłych okres, z którego pochodzą badane zasoby. Wspomniany trend jest widoczny w badaniu przeprowadzonym przez Siluo Yanga, Ruizhen Hana, Jingda Dinga oraz Yanfei Songa. Chińscy naukowcy analizowali trzy zbiory danych dotyczące cytowań z zakresu humanistyki i nauk społecznych (1998–2009), informatyki (1995–1999) oraz medycyny (1994–2006). Rozkład cytowań zasobów internetowych przypomina we wszystkich dziedzinach literę „S”. Po okresie szybkiego wzrostu następuje jego wyhamowanie. Różnice między zbiorami dotyczą m.in. liczby cytowanych materiałów internetowych na artykuł. Jedną z przyczyn zaobserwowanej odrębności może być pochodzenie materiału badawczego z różnych dziedzin wiedzy. Zdaniem badaczy, zaobserwowane zahamowanie wzrostu liczby analizowanych powołań wynika z ograniczeń dotyczących liczby cytowanych pozycji bibliograficznych w jednym artykule i liczby artykułów w pojedynczym czasopiśmie (Yang et al., 2012, 782–783). Dion Hoe-Lian Goh oraz Peng Kin Ng również wskazali na wzrost liczby cytowanych materiałów internetowych w trzech czasopismach z zakresu BIN (1997–2003). W 1997 r. artykuł zawierał średnio 0.89 cytowania zasobu internetowego, w 2002 r. – 6.42, jednak w 2003 r. było tych powołań średnio tylko 3.93. Badacze nie byli w stanie przedstawić pewnej przyczyny spadku liczby powołań na zasoby internetowe w ostatnim analizowanym roku (Goh & Ng, 2007, 18–19).

Zbieżne wyniki uzyskiwano w odniesieniu do stosunku między liczbą łączy aktywnych a liczbą łączy nieaktywnych w przypisach i bibliografiach załącznikowych. Liczba aktywnych łączy była odwrotnie proporcjonalna do ich wieku. Zależność taką zaobserwowano w przypadku czasopism z wielu dziedzin nauki, m.in. z prawa (Zittrain et al., 2014, 185),

medycyny (Habibzadeh, 2013, 457), BIN (Kumar et al., 2015, 61; Wu, 2008, 487) oraz informatyki biomedycznej (Carnevale & Aronsky, 2007, 271). Ten generalny wzrostowy trend nie oznacza identyczności wyników wszystkich badań. Różnią się one między sobą nie tylko w poszczególnych dziedzinach, ale także w ramach jednej dyscypliny naukowej. Również ranga czasopisma może mieć wpływ na procent aktywnych i dostępnych URL w cytatach. P. Habibzadeh pokazał, że w przypadku czasopism medycznych (2006–2013), trwałość łączy jest większa w międzynarodowych, prestiżowych publikacjach z tej dziedziny. Zdaniem badacza niższa trwałość łączy w czasopismach regionalnych może wynikać z ograniczonego dostępu naukowców w krajach rozwijających się do naukowych baz danych i innych zasobów (Habibzadeh, 2013, 458). Poszczególne badania różnią się między sobą przyjętą metodologią. Niektóre z nich uwzględniają zmianę treści zasobów poprzez zaliczanie linków prowadzących do zmienionych treści do grupy łączy nieaktywnych (zob. np. Wu, 2008), inne zaś ignorują tę kwestię (zob. np. Goh & Ng, 2007). Na przykład Zhiqiang Wu, badający cytowania materiałów internetowych w dwóch chińskich czasopismach z zakresu informatyki i BIN (1999–2003), podkreślił, że krzywa wskazująca liczbę łączy kierujących do zasobów o niezmienionej treści jest zawsze poniżej krzywej pokazującej liczbę aktywnych linków. W 2003 r. stosunek ten wynosił 321 do 449 URL (Wu, 2008, 486–487). W badaniach Jonathana Zittraina, Kendry Alberta, Lawrence’a Lessiga z 353 wybranych aktywnych łączy z korpusu opinii Sądu Najwyższego Stanu Zjednoczonych (1996–2011) tylko 76% prowadziło do zasobów niezmienionych treściowo (Zittrain et al., 2014, 182). Zjawisko rozbieżności między liczbą aktywnych łączy a liczbą łączy prowadzących do cytowanych przez autorów zasobów było analizowane również odnośnie czasopism z zakresu BIN (zob. Mardani & Sangari, 2013, 35–36).

Badacze starali się również odnajdywać zasoby niedostępne pod adresem umieszczonym w przypisie bądź bibliografii załącznikowej. W tym celu wykorzystywali najczęściej wyszukiwarki internetowe, serwis Internet Archive (Wayback Machine) oraz samodzielnie szukali cytowanych materiałów w serwisach internetowych. Mary F. Casserly oraz James E. Bird analizujący próbkę 500 przypisów z czasopism z zakresu BIN (1999–2000), uzyskali dzięki Internet Archive dostęp do 47% niedostępnych już, a cytowanych wcześniej zasobów (Casserly & Bird, 2003, 310). Wu udało się odnaleźć aż 74.4% brakujących dokumentów. Z tego 62.8% uzyskano wykorzystując wyszukiwarki, 24.4% serwis Internet Archive, a 12.8% przeszukując oryginalne strony WWW zawierające cytowane materiały (Wu, 2008, 494–495). D. Vinay Kumar, B. T. Sampath Kumar oraz D. R. Parameshwarappa uwzględnili w swoim badaniu zależność między różnymi analizowanymi czynnikami, np. długością łączy a możliwością odszukania zasobów w Internet Archive (Kumar et al., 2015). Zaprezentowane rezultaty wskazują, że większość cytowanych zasobów nadal istnieje w Internecie. Zmianie uległa tylko ich lokalizacja w sieci WWW. Pokazuje to zasadność utrzymywania stałych identyfikatorów odsyłających do materiałów sieciowych. Kwestia ta zostanie opisana dokładniej w dalszej części artykułu.

Porównanie wyników uzyskanych przez różnych badaczy umożliwia współczynnik wskazujący na czas połowicznej dezaktywacji łączy (ang. *half-life*). Pokazuje on czas po jakim połowa cytowanych linków nie będzie aktywna. Badacze wyliczali wartość tego współczynnika zarówno dla konkretnych czasopism, jak i dziedzin wiedzy. W tym drugim przypadku czas połowicznej dezaktywacji łączy był równy średniej uzyskanej dla kilku czasopism z danej dziedziny. Współczynnik ten jest odmienny dla różnych dziedzin

wiedzy i czasopism (Koehler, 2004), jednak nawet w obrębie jednej dziedziny jego wartość różni się niekiedy znacznie dla konkretnych publikacji. Goh i Ng (1997–2003) oraz K. R. Prithviraj i B. T. Sampath Kumar (2001–2010) wyliczyli jego wartość dla BIN na około 5 lat (Goh & Ng, 2007, 20; Prithviraj & Kumar, 2014, 44), natomiast A. H. Mardani i M. Sangari dla irańskich czasopism z tej dziedziny (2006–2010) uzyskali wynik wynoszący 3.63 roku (Mardani & Sangari, 2013, 34–35). Jednocześnie dla internetowego czasopisma „D-Lib Magazine” (1995–2004) uzyskano wynik 10 lat (McCown, Sheffan, Nelson, & Bollen, 2005, 5), a dla czasopisma „Information Research” (1995–2008) – 14.04 lata (Moghaddam, Saberi, & Esmaeel, 2010, 71–72). Daniela V. Dimitrova i Michael Bugeja dla czasopism z zakresu dziennikarstwa i komunikacji (2000–2003) uzyskali czas połowicznej dezaktywacji na poziomie 3.17 lat (Dimitrova & Bugeja, 2007, 820), natomiast Habibzadeh dla prestiżowych międzynarodowych czasopism medycznych uzyskał wynik 4.7 lata, a dla publikacji regionalnych – 3.5 roku (Habibzadeh, 2013, 463).

Zestawiano również kody odpowiedzi Hypertext Transfer Protocol (HTTP) uzyskane w przypadku nieaktywnych łączy. Najczęstszym z nich był kod 404 wskazujący na brak możliwości odnalezienia poszukiwanego pliku. Odsetek takich odpowiedzi wahał się w przypadku badań dotyczących literatury naukowej z zakresu BIN od 36.52% (Kumar et al., 2015, 61–62) do 63.84% (Sampath Kumar & Prithvi Raj, 2012, 25). W pierwszym z wymienionych badań, dotyczącym dwóch czasopism z okresu 2008–2012, najczęściej pojawiającą się odpowiedzią był kod 500, oznaczający wewnątrz błąd serwera (39.18%). Autorzy nie odnieśli się jednak do przyczyn tej odmienności (Kumar et al., 2015, 61–62). W pozostałych badaniach kod 500 pojawiał się najczęściej na drugim miejscu, po nim zaś kod 403 oznaczający dostęp zabroniony. Poza tym wśród zwracanych oznaczeń pojawiały się kody 502, 504, 401 oraz 400 (Goh & Ng, 2007, 19; Gul, Mahajan, & Ali, 2014, 32; Kumar et al., 2015, 61–62; McCown et al., 2005, 5; Moghaddam et al., 2010, 70; Prithviraj & Kumar, 2014, 41; Sadat-Moosavi, Isfandyari-Moghaddam, & Tajeddini, 2012, 186–187; Sampath Kumar & Prithvi Raj, 2012, 25).

Diomidis Spinellis przedstawił hipotezę mówiącą, że wraz ze wzrostem głębokości adresu URL, czyli liczby katalogów w łączy, wzrasta liczba niedostępnych zasobów (Spinellis, 2003, 75). Naukowcy uwzględniają w swoich badaniach analizę związków między głębokością łączy a czasem ich aktywności. W badaniu Goha oraz Nga liczba nieaktywnych adresów URL wzrastała od 25.6% w przypadku poziomu 0 do 39.4% w przypadku poziomu 5. Goh i Ng posługując się współczynnikiem korelacji Pearsona nie stwierdzili jednak zależności między badanymi zmiennymi (Goh & Ng, 2007, 21). Istotny związek między nimi stwierdzili natomiast Prithviraj i Kumar. Ich badanie pokazało na niemożność odnalezienia od 25.44% zasobów (poziom 0) aż do 64.16% (poziom 5) (Prithviraj & Kumar, 2014, 43). Kwestią rzadko poruszaną w badaniach dotyczących literatury z zakresu BIN jest problem zależności między liczbą znaków adresu URL a trwałością łączy. Zbadali ją Kumar, Kumar oraz Parameshwarappa wskazując na większy poziom niedostępności zasobów identyfikowanych za pomocą dłuższych URL. Badacze nie przeprowadzili analizy korelacji zmiennych (Kumar et al., 2015, 63–64).

Kolejnym czynnikiem, który badano był rodzaj domeny i jego związek ze stopniem nieaktywności łączy. W badaniach dotyczących literatury z zakresu BIN stosunkowo najczęściej cytowanymi domenami były domeny *com* (*co*), *org* i *edu* (*ac*) (Goh & Ng, 2007, 22; Gul et al., 2014, 32; Moghaddam et al., 2010, 67; Prithviraj & Kumar, 2014, 42; Sampath Kumar &

Prithvi Raj, 2012, 26). W pracy Mardaniego i Sangariego, której celem była analiza przypisów z sześciu irańskich czasopism BIN, wśród często cytowanych domen występuje też domena *gov* (Mardani & Sangari, 2013, 36). Popularność domeny *com* można wiązać ze stopniem jej rozpowszechnienia w Internecie, domena *edu* jest związana z treściami naukowymi i edukacyjnymi (Goh & Ng, 2007, 22), natomiast domena *org* wydaje się być popularna ze względu na wiarygodność treści (Sampath Kumar & Prithvi Raj, 2012, 25). Wśród nieaktywnych łączy duży odsetek stanowią zazwyczaj w różnych badaniach URL z domeną *edu*. Goh i Ng stwierdzają, że taki wynik może w pierwszej chwili wydawać się niezrozumiały, ponieważ uniwersytety dysponują zazwyczaj dobrą infrastrukturą oraz wiedzą techniczną. Przyczyn takiego wyniku upatrują w dynamicznym charakterze zawartości akademickich stron WWW, w przypadku których stara zawartość jest często zastępowana nową (Goh & Ng, 2007, 22). Między czasopismami z różnych dziedzin wiedzy zachodzą różnice w liczebności cytowanych adresów z poszczególnych domen (Yang et al., 2012, 783–784).

Druga grupa badań dotyczy sposobów radzenia sobie z problemem nietrwałości łączy i zasobów internetowych cytowanych w pracach naukowych. W jednym z najwcześniejszych artykułów odnoszących się do BIN, Casserly oraz Bird zwrócili uwagę na konieczność formułowania instrukcji dla autorów, tak, aby zawierały wskazówki jak ocenić jakość i trwałość zasobów. Badacze podkreślali, że niezbędne jest uwzględnienie sposobów cytowania zasobów internetowych, które powinny zawierać jak najpełniejsze informacje na temat dokumentu, w tym datę dostępu. Redaktorzy powinni przeglądać i korygować instrukcje dla autorów, dostosowując je do aktualnych trendów (Casserly & Bird, 2003, 315–316). Zalecenia te nie straciły swojej aktualności. W związku z rozpowszechnieniem i wzrostem liczby zasobów naukowych w postaci cyfrowej, coraz więcej redakcji tworzy własne lub wykorzystuje istniejące style cytowań odnośnie dokumentów internetowych. Niektórzy wskazywali także, że lepiej jest cytować zasoby z kolekcji cyfrowych takich jak biblioteki cyfrowe czy komercyjne bazy danych, niż ze stron WWW (Goh & Ng, 2007, 23). Nie zawsze jest to jednak możliwe i wydaje się, że niepotrzebnie ogranicza możliwości autorów prac naukowych. W niektórych przypadkach rozwiązaniem może być stosowanie stałych identyfikatorów takich jak Persistent Uniform Resource Locator (PURL) lub Digital Object Identifier (DOI) (zob. np. Gul et al., 2014, 36). Przegląd 1.8 miliona tekstów opublikowanych w latach 1997–2012 w bazach: arXiv.org, Elsevier i PubMed Central pokazał, że autorzy w dużym stopniu używają Uniform Resource Identifier (URI) zamiast DOI w bibliografiach załącznikowych i przypisach. Ostatecznie po wyłączeniu zasobu Elsevier uzyskano 397 412 DOI, 505 647 identyfikatorów, które powinny być zapisane jako DOI oraz 737 847 identyfikatorów niewyrażanych w opisywanym systemie. Autorzy badania, Herbert Van de Sompel, Martin Klein i Shawn M. Jones podkreślili, że nie znają przyczyny, dla której odsetek identyfikatorów należących do drugiej kategorii jest tak wysoki. Postawili hipotezę, że jest to związane ze sposobem działania programów do zarządzania bibliografią, które pobierają lokalizujące URI zamiast identyfikującego (Van de Sompel, Klein, & Jones, 2016). Jedną z przyczyn może być również nieznanie systemu DOI i korzyści wynikających z jego stosowania. Należy podkreślić, że rozwiązanie to jest trudne do wprowadzenia w przypadku stron internetowych, np. blogów, ponieważ wymaga sporego wysiłku od osób tworzących i utrzymujących dany serwis (Zittrain et al., 2014, 189).

Oprócz wspomnianych wyżej sposobów, innym rozwiązaniem jest archiwizacja cytowanych zasobów internetowych. Najprostszym technicznie sposobem przywoływanym

w literaturze przedmiotu, jest ich autoarchiwizacja przez autorów publikacji naukowych oraz redakcje (Prithviraj & Kumar, 2014, 45). Problemem wydaje się w tym przypadku stabilność takiej formy zachowania treści cyfrowych. Kolejnym sposobem jest skorzystanie z największego obecnie rozwijanego serwisu do archiwizacji zasobów Internetu – Internet Archive. Pozwala on zapisać na żądanie użytkownika konkretną stronę WWW dzięki opcji „Save Page Now” (Internet Archive, b.d.). Wątpliwości może jednak budzić stabilność tego serwisu związana z prywatnym charakterem instytucji (non profit) i jej dobrowolnym wsparciem (Zittrain et al., 2014, 190). Wspomniano, że w części prac serwis ten był wykorzystywany do odszukania treści niedostępnych pod zamieszczonym w bibliografii łączem URL. Badania Ke Zhoua, Claire Grover, Richarda Tobina oraz Martina Kleina przeprowadzone na kolekcji 648 388 artykułów z publikacji wydawnictwa Elsevier, pokazało, że z 70 270 nieaktywnych łączy, archiwizacji w wybranych aktualnie funkcjonujących archiwach internetowych poddano tylko 43 745 linków (62.3%). W celu zwiększenia odsetka zarchiwizowanych zasobów, badacze zaproponowali automatyczny sposób przewidywania dezaktywacji łączy w pracach naukowych oparty na uczeniu maszynowym (Zhou et al., 2015). W sieci WWW istnieją specjalne narzędzia przeznaczone do archiwizacji materiałów internetowych cytowanych w pracach naukowych. Najpopularniejsze z nich to serwisy WebCite (WebCite, b.d.), archive.is (archive.is, b.d.) oraz Perma (Perma.cc, b.d.). Jonathan Zittrain, Kendra Albert oraz Lawrence Lessig, autorzy artykułu dotyczącego serwisu Perma, opisali przyczyny powstania narzędzia wynikające ze słabości innych dostępnych rozwiązań. Zwrócili uwagę na problemy związane z finansowaniem serwisu WebCite oraz brakiem wsparcia instytucjonalnego dla archive.is. Ponadto, oba rozwiązania bazują na jednym źródle utrzymania serwisów, co nie zapewnia wystarczającej stabilności rozwiązania (Zittrain et al., 2014, 190–191). Inne rozwiązanie zostało wdrożone w przypadku serwisu Perma, który jest utrzymywany przez sieć współpracujących ze sobą instytucji, głównie amerykańskich bibliotek prawniczych (Zittrain et al., 2014, 191). Narzędzie jest przeznaczone głównie dla instytucji amerykańskich, jednak każdy może utworzyć darmowe konto umożliwiające archiwizację 10 stron miesięcznie za darmo (Perma.cc Docs, b.d.).

### 3. Metodologia

Przedmiotem badań przeprowadzonych przez autorów niniejszego artykułu był zbiór artykułów opublikowanych w latach 2009–2015 w czterech czasopismach z zakresu BIN. Były to trzy czasopisma naukowe – *Bibliotheca Nostra. Śląski Kwartalnik Naukowy, Przegląd Biblioteczny, Zagadnienia Informacji Naukowej – Studia Informacyjne* i jedno fachowe – *Biuletyn EBIB*. W badaniu nie uwzględniono danych z ostatniego (4) zeszytu czasopisma *Bibliotheca Nostra. Śląski Kwartalnik Naukowy* z 2015 r. w związku z jego niedostępnością w momencie prowadzenia prac badawczych. Zbiór publikacji tworzą wyłącznie artykuły o charakterze badawczym lub przeglądowym, w których autorzy realizują przyjęte cele badawcze posługując się aparatem naukowym w postaci wykazu cytowanych publikacji. Tym samym przedmiotem badania były teksty opublikowane we wskazanych czasopismach, które zamieszczono w działach lub sekcjach wskazujących na ich badawczy charakter. Nie uwzględniono m.in. recenzji oraz sprawozdań. Łącznie przeanalizowano 670 artykułów, których pełne bibliografie załącznikowe i przypisy manualnie przeniesiono do arkusza

kalkulacyjnego. Nie uwzględniono adresów URL zamieszczonych w tekstach publikacji. Ten sposób odwołań do zasobów internetowych był stosowany w niektórych numerach *Biuletynu EBIB*. Następnie oznaczono pozycje, w których wystąpiły odwołania do zasobów sieciowych i w sposób automatyczny wyodrębniono występujące tam adresy URL. Uzyskano w ten sposób zbiór 4867 adresów URL, co po usunięciu duplikatów dało 4593 unikalnych odesłań.

Podczas przygotowywania danych do analizy natrafiono na kilka problemów. Pierwszy z nich dotyczył formatowania niektórych adresów URL. W przypadku długich adresów stosunkowo często stosowany był znak przeniesienia części adresu do następnej linii, co skutkowało jego formalną modyfikacją i uzyskaniem komunikatu o niedostępności zasobu. Niekiedy adres łączy był również zakończony kropką. W adresach występowały również zbędne spacje. Zdecydowano się usunąć z adresów URL wszystkie dodatkowe znaki nie wchodzące w skład oryginalnego linku. Wśród błędów w zapisie łączy pojawiały się tzw. literówki, np. zamiast „http:” występował zapis „http;”. Tego rodzaju błędy również zostały skorygowane. Autorzy wykorzystywali również w przypisach bibliograficznych identyfikatory DOI, które w większości przypadków nie miały postaci URL, np. doi:10.1007/978–3–642–04346–8. W celu zapewnienia bezpośredniego dostępu do zasobu zdecydowano się w sposób automatyczny dokonać konwersję tych zapisów do postaci URL poprzez dodanie do identyfikatorów DOI formuły „http://dx.doi.org/”, np. http://dx.doi.org/doi:10.1007/978–3–642–04346–8. Do analizy odesłań wykorzystano aplikację HEADMasterSEO (HEADMasterSEO, 2016). Jest to oprogramowanie, które na podstawie wskazanej listy adresów URL pobiera odpowiedzi serwera w postaci kodów odpowiedzi http. Aplikacja ta przeprowadza pierwszy etap interakcji między klientem a serwerem realizującym dostęp do danego dokumentu, czyli wysłała zapytanie o dostępność zasobu otrzymując ustandaryzowany kod odpowiedzi wskazujący na status jego dostępności. Według dokumentacji protokołu HTTP kody odpowiedzi przyjmują wartość trzycyfrową (od 100 do 511) i można je zorganizować w pięciu grupach (HTTP, 2004):

- (1) Kody informacyjne, np. 110: Przekroczono czas połączenia. Serwer zbyt długo nie odpowiada.
- (2) Kody powodzenia, np. 200: Zawartość żądanego dokumentu.
- (3) Kody przekierowania, np. 301: Trwale przeniesiony – żądany zasób zmienił swój URI i w przyszłości zasób powinien być szukany pod wskazanym nowym adresem.
- (4) Kody błędu aplikacji klienta, np. 404: Nie znaleziono – serwer nie odnalazł zasobu według podanego URL ani niczego co by wskazywało na istnienie takiego zasobu w przeszłości.
- (5) Kody błędu serwera http, np. 500: Wewnętrzny błąd serwera – serwer napotkał niespodziewane trudności, które uniemożliwiły zrealizowanie żądania.

Aplikacja HEADMasterSEO zwracała dla każdego adresu URL kod odpowiedzi serwera HTTP, a w przypadku kodów wskazujących na przekierowania dodatkowo ich liczbę oraz kod odpowiedzi dla ostatecznie ustanowionego adresu URL. W celu określenia wiarygodności uzyskanych danych na losowo dobranej próbie 10% badanych adresów URL przeprowadzono ponowne testy. Wyniki te potwierdziły wcześniej uzyskane kody odpowiedzi.

Badanie przeprowadzono w dniach 9–13 maja 2016 r. korzystając z łącza internetowego o przepustowości 8 Mb/s. z zalecaną przez producenta konfiguracją aplikacji HEADMasterSEO.

## 4. Wyniki

W tej części artykułu zaprezentowano wyniki badania czterech polskich czasopism z zakresu BIN. Zostały one podzielone na zagadnienia badawcze zgodnie z podziałem przedstawionym w części poświęconej stanowi badań. Wyniki są podstawą do porównania liczby cytowań zasobów internetowych i aktywności cytowanych materiałów w polskich i zagranicznych czasopismach z zakresu BIN oraz przedstawienia rekomendacji związanych z problemem braku stabilności odwołań do zasobów sieciowych.

### 4.1. Analiza bibliografii załącznikowych i przypisów

W analizowanym zbiorze ok. 37% wszystkich pozycji w bibliografiach załącznikowych i przypisach zawierało odwołania do ich lokalizacji w Internecie. Wskaźnik ten utrzymuje się na poziomie 20–30% dla trzech tytułów czasopism z wyjątkiem *Biuletynu EBIB*, gdzie średnio ponad połowa (58.6%) cytowań, zawiera odwołania do zasobów sieciowych (Tab. 1).

Tabela 1. Liczba i odsetek cytowań z URL w badanych czasopismach

| Tytuł czasopisma                                      | Liczba cytowań z URL | Odsetek cytowań z URL | Średnia liczba cytowań z URL na artykuł |
|-------------------------------------------------------|----------------------|-----------------------|-----------------------------------------|
| Przegląd Biblioteczny                                 | 923                  | 31.4                  | 10.3                                    |
| Bibliotheca Nostra. Śląski Kwartalnik Naukowy         | 920                  | 23.4                  | 6.4                                     |
| Biuletyn EBIB                                         | 2317                 | 58.6                  | 6.7                                     |
| Zagadnienia Informacji Naukowej – Studia Informacyjne | 707                  | 29.0                  | 7.5                                     |

Z punktu widzenia pojedynczej publikacji, średnia liczba cytowań dokumentów sieciowych wynosi 7.3 pozycji. Jedynie w odniesieniu do *Przeglądu Bibliotecznego* zaobserwowano wyższą wartość dla tej zmiennej (średnio 10.3). Rozkład wartości dla tego wskaźnika z podziałem na rok publikacji artykułu cytującego prezentuje tabela 2. Największą wartość zmienna ta przyjmuje dla artykułów opublikowanych w 2010 i 2014 r. (odpowiednio 9 oraz 8.1).

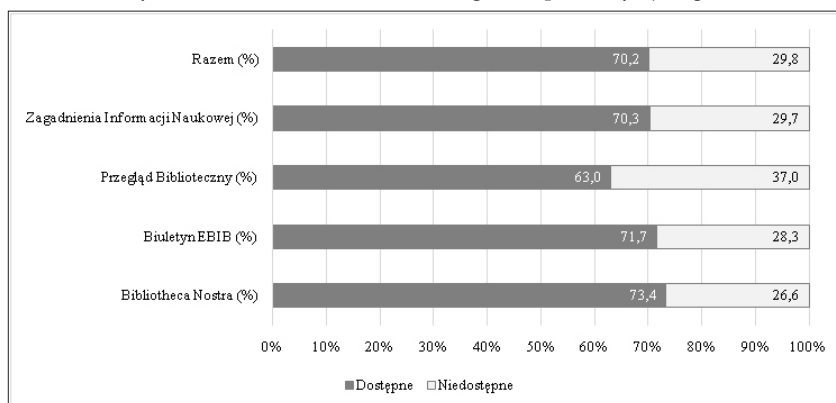
Tabela 2. Średnia liczba cytowań z URL na artykuł wg roku wydania publikacji cytującej

| Rok publikacji | Liczba artykułów | Liczba cytowań z URL | Średnia liczba cytowań z URL na artykuł |
|----------------|------------------|----------------------|-----------------------------------------|
| 2009           | 97               | 585                  | 6.0                                     |
| 2010           | 78               | 705                  | 9.0                                     |
| 2011           | 90               | 541                  | 6.0                                     |
| 2012           | 113              | 854                  | 7.6                                     |
| 2013           | 96               | 724                  | 7.5                                     |
| 2014           | 104              | 846                  | 8.1                                     |
| 2015           | 92               | 612                  | 6.7                                     |
| Razem          | 670              | 4867                 | 7.3                                     |

#### 4.2. Aktualność adresów URL w bibliografiach załącznikowych i przypisach

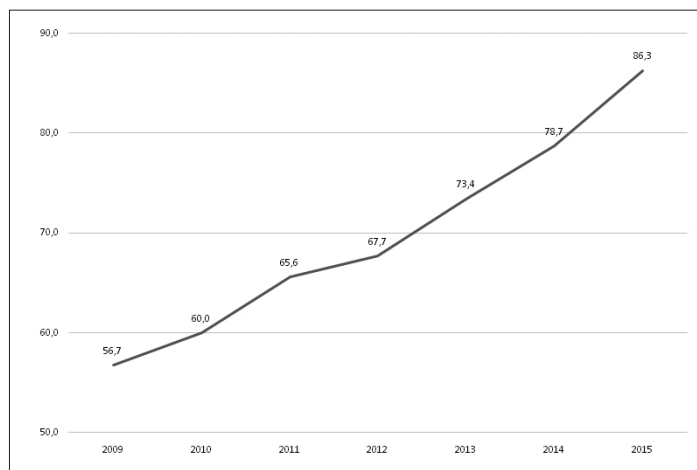
W wyniku przeprowadzonych testów na zbiorze 4593 adresów URL określono, że 70.2% odesłań do zasobów internetowych było aktywnych w maju 2016 r. Największy odsetek nieaktywnych odwołań odnotowano dla *Przeglądu Bibliotecznego* i wynosi on 37%. W pozostałych przypadkach aktualność odesłań waha się w przedziale 70–73% (Rys. 1).

Rys. 1. Aktualność adresów URL wg czasopisma cytującego



Korelując wartość tego wskaźnika z rokiem publikacji artykułu cytującego można odnotować jego sukcesywny wzrost na przestrzeni siedmiu lat. W artykułach z 2009 r. tylko ponad połowa (56.7%) adresów URL w bibliografiach załącznikowych jest obecnie aktualna. W ciągu kolejnych lat odsetek nadal aktywnych odwołań rośnie średnio o 5% i dla artykułów z 2015 r. wynosi on 86.3% (Rys. 2).

Rys. 2. Odsetek aktualnych adresów URL wg roku wydania publikacji cytującej



Szczegółowa analiza zmian wartości wskaźnika dla aktualności adresów URL na przestrzeni siedmiu lat z odniesieniem do źródła publikacji cytującej nie potwierdza jednak

stabilnego wzrostu. O ile w odniesieniu do artykułów opublikowanych w 2015 r. odsetek aktywnych URL jest wyraźnie większy niż dla artykułów z 2009 r., to w okresie pomiędzy tymi datami nie można zauważyć jednolitej tendencji wzrostowej. Co więcej, w każdym z analizowanych tytułów odnotowano przynajmniej jedną taką sytuację, w której odsetek aktywnych URL zmalał w stosunku do jego wartości z roku poprzedniego. Największą stabilność w tym ujęciu reprezentuje *Biuletyn EBIB*. Uwagę zwraca stosunkowo niski odsetek aktywnych adresów URL zamieszczanych w *Przeglądzie Bibliotecznym* w latach 2009–2010 (Tab. 3). Nie udało się jednak ustalić przyczyn tego zjawiska.

Tabela 3. Odsetek aktualnych adresów URL dla poszczególnych tytułów czasopism w latach 2009–2015

|                                                       | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 |
|-------------------------------------------------------|------|------|------|------|------|------|------|
| Bibliotheca Nostra. Śląski Kwartalnik Naukowy         | 70.0 | 65.4 | 68.9 | 66.4 | 79.9 | 75.3 | 86.7 |
| Biuletyn EBIB                                         | 60.2 | 64.7 | 63.5 | 69.4 | 76.0 | 79.0 | 85.4 |
| Przegląd Biblioteczny                                 | 39.7 | 35.7 | 67.6 | 69.3 | 62.6 | 78.5 | 82.4 |
| Zagadnienia Informacji Naukowej – Studia Informacyjne | 52.5 | 71.5 | 60.6 | 63.0 | 70.5 | 82.7 | 93.7 |

### 4.3. Odpowiedzi serwera

Analizując bezpośrednie odpowiedzi serwerów na żądanie pobrania nagłówka HTTP przez aplikację HEADMasterSEO (Tab. 4) można zaobserwować, że tylko w ok. 44% wskazań zasoby internetowe były bezpośrednio dostępne pod wskazanym adresem. Jednocześnie odnotowano duży odsetek przekierowań (32.3%). W celu uzyskania pełnego obrazu poziomu dostępności do cytowanych zasobów sieciowych szczegółowej analizie poddano grupę adresów URL, dla których narzędzie badawcze zwróciło kody nagłówków z grupy 3XX. Wyniki tych badań wskazują, że w 79% przypadków przekierowania doprowadziły do uzyskania prawidłowego rezultatu, czyli zgłoszenia kodu 200 – zasób sieciowy jest dostępny. Tym samym średni rzeczywisty poziom dostępności dla analizowanego zbioru odwołań wynosi, tak jak wspomniano wcześniej, 70.2%.

Tabela 4. Grupy kodów http uzyskane dla badanych adresów URL

| Kod odpowiedzi  | Odsetek wskazań |
|-----------------|-----------------|
| 2XX             | 44.6            |
| 3XX             | 32.3            |
| 4XX             | 18.0            |
| 5XX             | 1.0             |
| Pozostałe błędy | 4.1             |

W zbiorze adresów URL, dla których odnotowano kod odpowiedzi z grupy 3XX występują łańcuchy przekierowań o wartościach od 1 do 10. Jednokrotne przekierowania

dotyczą 72% adresów z tej puli, dwukrotne – 19%, trzykrotne – 6%, czterokrotne – 1.6%, a pozostałe uzyskały wartości mniejsze niż 1%.

Spośród odpowiedzi odnoszących się do nieaktywnych adresów URL największy odsetek stanowił kod 404 (75.2%) oznaczający, że serwer nie odnalazł zasobu według podanego URL. Na drugim miejscu (16.4%) pod względem liczby wystąpień znalazła się grupa odpowiedzi nie w postaci kodu HTTP, lecz odpowiedzi aplikacji związanej z brakiem możliwości nawiązania połączenia (np. przekroczony czas oczekiwania, błąd połączenia). Pozostałe odpowiedzi pojawiały się bardzo rzadko, np. kody 403 (4.2%), 500 (3.6%) oraz 410 (2.7%).

#### 4.4. Wskaźnik połowicznej dezaktywacji łączy (*half-life*)

Kolejnym etapem analizy uzyskanych danych badawczych było obliczenie wskaźnika połowicznej dezaktywacji łączy. Określa on okres liczony w latach, po którym połowa identyfikatorów sieciowych przestaje być aktywna. W celu obliczenia wskaźnika posłużono się następującym wzorem (Dimitrova & Bugeja, 2007, 820):

$$t_h = \frac{t l_n(0,5)}{l_n W_{(t)} - l_n W_{(0)}}$$

Zmienna  $t_h$  oznacza czas połowicznej dezaktywacji łączy,  $W_{(0)}$  oznacza liczbę aktywnych łączy w chwili publikacji artykułu, a  $W_{(t)}$  – liczbę aktywnych łączy w późniejszym punkcie czasu  $t$ . Dla całego zbioru danych badawczych, a więc w odniesieniu do czterech analizowanych tytułów czasopism, wskaźnik połowicznej dezaktywacji łączy przyjął wartość 7.39. Tabela 5 przedstawia rozkład jego wartości dla poszczególnych tytułów czasopism w odniesieniu do roku publikacji artykułów cytujących.

Tabela 5. Rozkład wartości wskaźnika połowicznej dezaktywacji łączy (*half-life*)

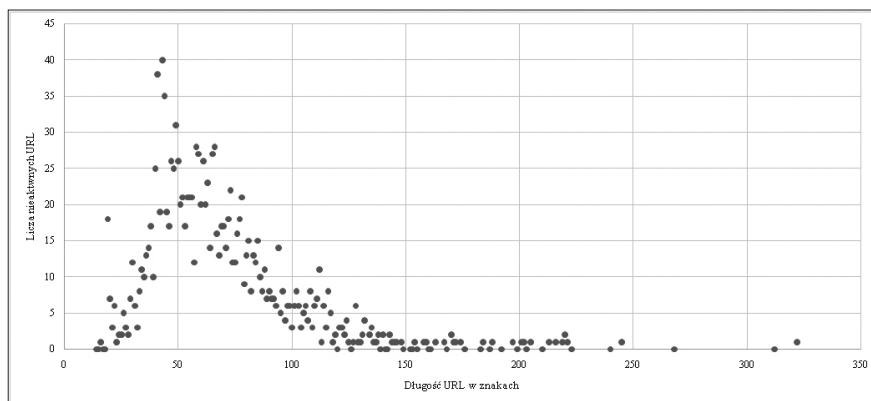
|                                                       | 2009  | 2010  | 2011 | 2012 | 2013 | 2014 | 2015  | Średnia |
|-------------------------------------------------------|-------|-------|------|------|------|------|-------|---------|
| Bibliotheca Nostra. Śląski Kwartalnik Naukowy         | 13.60 | 9.79  | 9.32 | 6.76 | 9.25 | 4.90 | 4.87  | 8.36    |
| Biuletyn EBIB                                         | 9.55  | 9.55  | 7.64 | 7.58 | 7.59 | 5.89 | 4.39  | 7.46    |
| Przegląd Biblioteczny                                 | 5.26  | 4.04  | 8.86 | 7.56 | 4.44 | 5.74 | 3.58  | 5.64    |
| Zagadnienia Informacji Naukowej – Studia Informacyjne | 7.53  | 12.42 | 6.92 | 5.99 | 5.96 | 7.28 | 10.60 | 8.10    |

Wartość wskaźnika jest różna dla poszczególnych czasopism na przestrzeni badanego okresu. Stały trend jest widoczny jedynie w przypadku łączy zamieszczonych w *Biuletynie EBIB*. Najniższą wartość współczynnik przyjął w przypadku *Przeglądu Bibliotecznego* (5.64). Jest to jedyne czasopismo spośród analizowanych w artykule, w którym połowa cytowanych łączy dla wybranych lat była nieaktywna w momencie przeprowadzania badania. Dotyczyło to lat 2009 (5.26) i 2010 (4.04).

#### 4.5. Budowa adresów URL a poziom dostępności

Omawiając stan badań wspomniano, że długość adresu URL liczona w znakach może być rozważana jako jeden z czynników wpływających na aktywność odesłania. Średnia długość adresu URL w analizowanym zbiorze wyniosła 60 znaków. Dla aktywnych odwołań wskaźnik ten przyjął wartość 57 znaków, a dla nieaktywnych 67. W wyniku analizy danych nie stwierdzono liniowego wzrostu liczby nieaktywnych łączy, czyli zależności między długością adresu URL a dostępnością zasobu sieciowego (Rys. 3).

Rys. 3. Zależność między długością adresu URL a jego dostępnością



Źródło: badania własne.

Analiza tego zjawiska z zastosowaniem organizacji danych w przedziałach znakowych odpowiadających długości adresu URL oraz odsetka ich dostępności również nie potwierdza tej hipotezy (Tab. 6). Widać jednak, że największy odsetek odwołań to adresy URL o długości od 41 do 60 znaków i to w tej grupie odnotowano największą liczbę nieaktywnych łączy.

Tabela 6. Zależność między długością adresu URL a jego dostępnością

| Długość URL | Liczba URL | Odsetek URL | Odsetek nieaktywnych URL |
|-------------|------------|-------------|--------------------------|
| 1–20        | 170        | 3.70        | 1.90                     |
| 21–40       | 952        | 20.73       | 11.70                    |
| 41–60       | 1645       | 35.82       | 35.41                    |
| 61–80       | 934        | 20.34       | 26.04                    |
| 81–100      | 484        | 10.54       | 12.66                    |
| 101–120     | 245        | 5.33        | 7.24                     |
| 121–140     | 96         | 2.09        | 2.85                     |
| >141        | 67         | 1.46        | 2.19                     |

Nawiązując do hipotezy wspomnianej w pierwszej części tekstu, wskazującej na zależność poziomu dostępności zasobu od głębokości adresu URL, czyli liczby katalogów w łańcuchu, nie stwierdzono takiej zależności w odniesieniu do zgromadzonych danych (Tab. 7).

Tabela 7. Zależność między głębokością adresu URL a jego dostępnością

| Głębokość URL | Liczba URL | Liczba nieaktywnych URL | Odsetek nieaktywnych URL |
|---------------|------------|-------------------------|--------------------------|
| 0             | 1          | 0                       | 0                        |
| 1             | 1136       | 225                     | 19.8                     |
| 2             | 1208       | 352                     | 29.1                     |
| 3             | 977        | 342                     | 35.0                     |
| 4             | 624        | 217                     | 34.8                     |
| 5             | 345        | 140                     | 40.6                     |
| 6             | 179        | 54                      | 30.2                     |
| 7             | 81         | 22                      | 27.2                     |
| 8             | 16         | 8                       | 50.0                     |
| 9             | 19         | 6                       | 31.6                     |
| 10            | 4          | 0                       | 0                        |
| 11            | 2          | 1                       | 50.0                     |
| 13            | 1          | 0                       | 0                        |

Co prawda, w przypadku pięciopoziomowych adresów URL połowa z odwołań była nieaktywna, lecz tendencja ta nie utrzymywała się na wyższych poziomach, a liczba takich adresów nie była znacząca. Ponad 70% adresów URL w zbiorze wskazywało na dosyć płytką strukturę odwołań (do trzech podkatalogów) i w tej grupie można zaobserwować wzrost poziomu niedostępności zasobu wraz ze zwiększającą się głębokością odwołania.

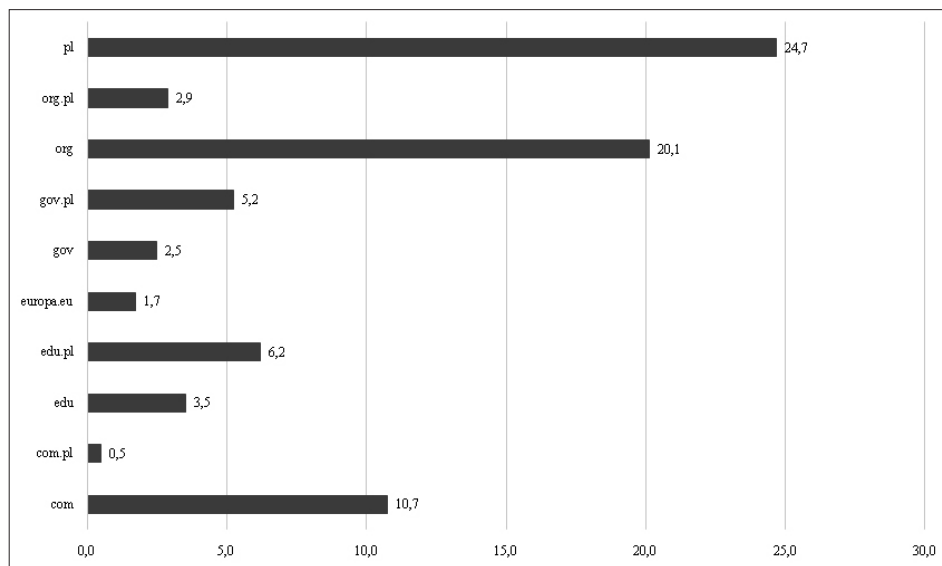
#### 4.6. Pochodzenie cytowanych zasobów sieciowych oraz jego wpływ na poziom ich dostępności

Celem prac badawczych, oprócz określenia poziomu dostępności cytowanych zasobów internetowych, było również naszkicowanie ogólnego obrazu tego, co cytują autorzy publikacji. W tym celu z puli adresów URL automatycznie wyodrębniono typy domen internetowych oraz nazwy domenowe. W pierwszym przypadku zdecydowano zastosować podstawową typologię domen internetowych z dodatkowym uwzględnieniem rozszerzenia *pl* jako wskazania na rejestrację domeny w Polsce oraz *europa.eu* jako wskazania na rejestrację domeny w ramach usług informacyjnych Unii Europejskiej (Rys. 4).

Z punktu widzenia typu domeny największy odsetek adresów URL wskazuje na zasoby cyfrowe udostępniane przez organizacje pozarządowe i non profit (*org* – 20.1%). Jedynie niecałe 3% adresów URL wskazywało na tego typu źródła zarejestrowane w Polsce (*org.pl*). Ponad 11% adresów URL było związanych z podmiotami komercyjnymi, z czego tylko 0.5% miało zarejestrowano w domenie *com.pl*. Warto podkreślić, że domena ta może być rejestrowana nie tylko przez firmy, ale również przez osoby prywatne. Bardzo podobny odsetek wystąpień, co domena *com*, miały również strony należące do instytucji edukacyjnych, z tym

że domena ogólna *edu* (3.5%) miała mniej wystąpień niż jej polska odmiana *edu.pl* (6.2%). Około 8% adresów przekierowywało do serwisów informacyjnych instytucji rządowych (*gov* i *gov.pl*). Uwzględniając narodowy charakter domeny, prawie jedna czwarta adresów zawierała wyłącznie rozszerzenie *pl*. Nieuwzględnione na rysunku 4 pozostałe typy domen były reprezentowane przez ok. 22% adresów URL ze zbioru badawczego.

Rys. 4. Typy domen (%)



Analiza danych pod kątem nazw domenowych, występujących w adresach URL, nie doprowadziła do wskazania dominujących sieciowych źródeł informacji występujących w przypisach i bibliografiach załącznikowych publikacji cytujących. W tabeli 8 przedstawiono jednak dziesięć internetowych źródeł informacji o największym odsetku cytowań.

Tabela 8. Nazwy domenowe w adresach URL (%)

| Domena                                                            | Odsetek wskazań |
|-------------------------------------------------------------------|-----------------|
| <a href="http://www.ebib.info/">http://www.ebib.info/</a>         | 2.3             |
| <a href="http://goo.gl/">http://goo.gl/</a>                       | 2.2             |
| <a href="http://pl.wikipedia.org/">http://pl.wikipedia.org/</a>   | 1.9             |
| <a href="http://ur.oss.wroc.pl/">http://ur.oss.wroc.pl/</a>       | 1.6             |
| <a href="http://www.ifla.org/">http://www.ifla.org/</a>           | 1.1             |
| <a href="http://dx.doi.org/">http://dx.doi.org/</a>               | 1.0             |
| <a href="http://www.ala.org/">http://www.ala.org/</a>             | 0.8             |
| <a href="http://eprints.rclis.org/">http://eprints.rclis.org/</a> | 0.7             |
| <a href="http://www.bn.org.pl/">http://www.bn.org.pl/</a>         | 0.7             |
| <a href="http://www.nauka.gov.pl/">http://www.nauka.gov.pl/</a>   | 0.7             |

W analizowanym zbiorze największą liczbę wskazań odnotowano do zasobów pochodzących z serwisu internetowego EBIB. Serwis ten jest notabene niedostępny pod tym adresem, a usługodawca nie zapewnił przekierowania do aktualnej lokalizacji zasobów, co spowodowało, że adresy URL zostały oznaczone przez narzędzie badawcze jako niedostępne. Odnotowano również 99 adresów URL, które wskazywały na wykorzystanie serwisu Google umożliwiającego utworzenie tymczasowego aliasu dla długiego adresu URL. Niestety, 21% z nich było aktualnie nieaktywnych. Na ten problem zwracają uwagę m.in. Klein et al. (2014). W bibliografiach załącznikowych i w przypisach widoczne jest wykorzystanie zasobów polskojęzycznej Wikipedii. W zbiorze badawczym odnotowano 88 odwołań do tego źródła, co stanowi niecałe 2% wszystkich wskazań. Autorzy publikacji w niewielkim stopniu wykorzystują stałe identyfikatory DOI. W zbiorze badawczym odnotowano tylko 44 adresy URL, które wykorzystywały ten system. Trudno jednak, wyłącznie na podstawie tej informacji, stawiać wnioski bez znajomości danych wskazujących na liczbę zarejestrowanych DOI dla wszystkich cytowań w zbiorze danych badawczych.

Interesujące wyniki uzyskano zestawiając typy domen, które reprezentowały zgromadzone adresy URL z poziomem ich aktualności (Tab. 9). Największą grupę odwołań (24.7%) stanowiły adresy URL zakwalifikowane do typu domeny *pl*. Jedna czwarta z nich okazała się być nieaktywna w okresie, w którym prowadzono badanie. Podobną zależność odnotowano w odniesieniu do typu *com*. W przypadku zasobów sieciowych opublikowanych w serwisach instytucji edukacyjnych, jedna trzecia adresów URL była nieaktywna. W przypadku polskich zasobów edukacyjnych odsetek ten wynosi 21.8%. Wysoką liczbę nieaktywnych adresów URL odnotowano dla instytucji rządowych. W przypadku Polski, to prawie połowa odesłań. Również ponad połowa adresów URL do zasobów instytucji Unii Europejskiej okazała się być nieaktywna.

Tabela 9. Pochodzenie cytowanych zasobów a stopień ich dostępności

| Typ domeny | Dostępne (%) | Niedostępne (%) | Udział (%) |
|------------|--------------|-----------------|------------|
| com        | 77.3         | 22.7            | 10.7       |
| com.pl     | 69.6         | 30.4            | 0.5        |
| edu        | 69.8         | 30.2            | 3.5        |
| edu.pl     | 78.2         | 21.8            | 6.2        |
| europa.eu  | 46.8         | 53.2            | 1.7        |
| gov        | 69.3         | 30.7            | 2.5        |
| gov.pl     | 50.6         | 49.4            | 5.2        |
| org        | 75.5         | 24.5            | 20.1       |
| org.pl     | 81.8         | 18.2            | 2.9        |
| pl         | 74.3         | 25.7            | 24.7       |
| pozostałe  | 60.4         | 39.6            | 21.9       |

Dopełnieniem badań nad dostępnością odwołań do zasobów sieciowych w bibliografiach załącznikowych i przypisach była próba weryfikacji pośredniej dostępności do zasobów zakwalifikowanych jako niedostępne. W tym celu wykorzystano cyfrowe archiwum zasobów

Internetu Wayback Machine (Wayback Machine, b.d.), które jest tworzone w ramach projektu Internet Archive. Na potrzeby badań wykorzystano interfejs programistyczny tej kolekcji Wayback Machine API (Application Programming Interface) (Wayback Machine API, b.d.), który posłużył do weryfikacji dostępności adresów URL. Uzyskane dane wskazują, że 53% zasobów sieciowych zidentyfikowanych w badaniu jako niedostępne posiada kopię ich zawartości w bazie Wayback Machine. Oznacza to, że faktyczny poziom dostępności tych odwiedzin w całym zbiorze badawczym jest znacznie wyższy. Niestety, dostęp do tych źródeł internetowych wymaga od osób zainteresowanych manualnego ich odnalezienia w zasobach Internetu.

#### 4.7. Zmiana treści cytowanych zasobów internetowych

Istotnym problemem, obok nieaktywności łączy do cytowanych dokumentów, jest kwestia zmiany treści tych zasobów. W badaniu podjęto próbę określenia odsetka cytowanych zasobów internetowych, których treść uległa modyfikacji od czasu odwołania się do nich w analizowanym zbiorze. Analizę przeprowadzono dla każdego czasopisma osobno. Pod uwagę wzięto wyłącznie te łączy, w przypadku których ostateczny kod odpowiedzi HTTP przyjął wartość 200. W celu oszacowania próby wykorzystano wzór (Jabłońska & Sobieraj, 2013, 35):

$$n_b = \frac{N}{1 + \frac{4d^2(N-1)}{Z^2}}$$

Zmienna  $N$  oznacza wielkość populacji, wielkość  $d$  oznacza zakładany błąd (5%), a wielkość  $Z$  wartość standardową 1.96 dla poziomu istotności  $p = 0.05$ . W analizie zmiany treści przywoływanych zasobów wykorzystano informacje bibliograficzne zamieszczone w przypisach bądź w bibliografiach załącznikowych. Jako zasoby o zmienionej treści oznaczano wyłącznie dokumenty, w przypadku których nie było żadnych wątpliwości, że taka zmiana nastąpiła (Tab. 10). Oznacza to, że wyliczony odsetek zmienionych treściowo zasobów to wartość minimalna, która jest zapewne nieco mniejsza od faktycznie występującej.

Tabela 10. Zmiana treści cytowanych zasobów w poszczególnych czasopismach

| Czasopismo                                            | Odsetek zasobów ze zmienioną treścią |
|-------------------------------------------------------|--------------------------------------|
| Bibliotheca Nostra. Śląski Kwartalnik Naukowy         | 11.16%                               |
| Biuletyn EBIB                                         | 9.71%                                |
| Przegląd Biblioteczny                                 | 15.32%                               |
| Zagadnienia Informacji Naukowej – Studia Informacyjne | 14.15%                               |

Średni odsetek zmienionych treściowo zasobów wyniósł 12.28%. Wyniki powyżej tej średniej uzyskano dla *Przeglądu Bibliotecznego* oraz *Zagadnień Informacji Naukowej*. Dla porównania w badaniu Mardaniego i Sangariego uzyskano wynik na poziomie 11% (Mardani & Sangari, 2013, 35–36).

## 5. Dyskusja

W artykule zaprezentowano analizę pozycji z bibliografii załącznikowych oraz przypisów zawierających odesłania do zasobów internetowych. Jako źródło danych badawczych wybrano cztery polskie czasopisma z zakresu BIN. Trzy z nich znajdują się na liście czasopism punktowanych opublikowanej przez Ministerstwo Nauki i Szkolnictwa Wyższego, jedno zaś jest internetowym czasopismem fachowym. Duża część poczynionych ustaleń jest zbieżna z wynikami badań przytoczonymi w części poświęconej stanowi badań, co potwierdza występowanie pewnych ogólnych trendów w zakresie aktywności łączy do cytowanych zasobów internetowych.

W czasopismach uwzględnionych w badaniu ok. 37% wszystkich pozycji w bibliografiach załącznikowych i przypisach zawierało łączy do zasobów internetowych. Średnia ta jest wyższa od uzyskanej w innych badaniach. Te niestety obejmują wcześniejsze lata publikacji, co w zestawieniu z prezentowanymi wynikami badań nie pozwala na jednoznaczne wyciąganie wniosków, ale może stanowić pewien punkt odniesienia. Ali Sadat-Moosavi, Alireza Isfandiyari-Moghaddam oraz Oranus Tajeddini obliczyli, że średni odsetek cytowanych zasobów sieciowych w czterech czasopismach dotyczących BIN w latach 2005–2008 wynosił 24% (Sadat-Moosavi et al., 2012, 184), a Kumar, Kumar oraz Parameshwarappa określili jego poziom na 23.81% (Kumar et al., 2015, 61) dla publikacji z lat 2008–2012. W ostatnim przytoczonym badaniu zaobserwowano spadek odsetka cytowanych zasobów internetowych. Procent cytowań z URL w czasopiśmie *Bibliotheca Nostra. Śląski Kwartalnik Naukowy* osiągnął najbliższą wartość w stosunku do średniej wyliczonej w innych badaniach, a internetowe czasopismo fachowe *Biuletyn EBIB* znacznie ją przekroczyło. Trudno jest jednak wskazać pewne przyczyny odmiennego stosunku pomiędzy powoływaniem się na materiały drukowane, a odwoływaniem się do zasobów sieciowych w polskich czasopismach z zakresu BIN. Może się to wiązać m.in. z czasem pochodzenia materiału badawczego. Należy podkreślić, że zwiększa to zagrożenie niedostępności dokumentów przytaczanych przez autorów artykułów.

Średnia liczba cytowań zasobów internetowych na artykuł dla polskich czasopism jest zazwyczaj wyższa (7.3) niż średnia dla ich zagranicznych odpowiedników. W jednym z badań czasopism z zakresu BIN (*The Electronic Library, Program*) z lat 2008–2012 wyliczono średnią liczbę cytowanych zasobów sieciowych na artykuł w wysokości 6.10 (Kumar et al., 2015, 61). Niekiedy jednak wynik jest znacznie wyższy, jak dla internetowego czasopisma *Ariadne* (2010–2012), w przypadku którego średnia wyniosła aż 16.21 (Gul et al., 2014, 31). Wynik ten jest rozbieżny z uzyskanym dla *Biuletynu EBIB*. Wysoka średnia w przypadku *Przeglądu Bibliotecznego* nie wiąże się z ogólnie wysoką liczbą pozycji bibliograficznych przywoływanych w tym czasopiśmie. Zaprezentowane wyniki są często rozbieżne, co wskazuje na ścisłą zależność omawianego wskaźnika od charakteru czasopisma.

W wyniku analiz określono, że 70.2% odesłań do zasobów internetowych było aktywnych w maju 2016 r. Jest to stosunkowo wysoki wskaźnik w porównaniu z uzyskanymi w innych badaniach czasopism z dziedziny BIN. Najbardziej zbliżony wynik uzyskali Goh i Ng – 69% (Goh & Ng, 2007, 19). Wskazuje to jednak, że zgodnie z oczekiwaniami niedostępność cytowanych zasobów internetowych jest istotnym problemem w przypadku polskich czasopism z dziedziny BIN. Problem ten nie jest równie widoczny we wszystkich badanych czasopismach. Najmniejszy odsetek aktywnych odesłań odnotowano dla *Przeglądu*

*Bibliotecznego* – 63%, co pokazuje wagę omawianego zagadnienia w przypadku tego periodyku. Problem występuje jednak również w innych czasopismach. Należy podkreślić, że redakcje czasopism mają tylko ograniczone możliwości wpływania na zwiększenie dostępności cytowanych zasobów. We wskazówkach dla autorów w trzech analizowanych czasopismach naukowych znajdują się dokładne zalecenia dotyczące sposobu cytowania dokumentów sieciowych. Poziom dostępności dokumentów internetowych można poprawić sprawdzając wszystkie linki przed ich użyciem i opublikowaniem w czasopiśmie, co było już podkreślane przez niektórych autorów (zob. np. Dimitrova & Bugeja, 2007, 822; Prithviraj & Kumar, 2014, 45). W przypadku *Biuletynu EBIB* problemem jest zaobserwowana w czasie badań duża zmienność sposobu cytowania zasobów internetowych. Niekiedy nawet w ramach jednego numeru stosowane były różne rozwiązania: przypisy, bibliografia załącznikowa bądź zamieszczanie linków w tekście artykułu. Z poruszonymi zagadnieniami ściśle wiąże się wartość współczynnika połowicznej dezaktywacji łączy. Średnia dla analizowanych czasopism jest stosunkowo wysoka i wynosi 7.39 lat. Wartość tego wskaźnika jest wyraźnie niższa dla *Przeglądu Bibliotecznego* (5.64). Jest to jednak wartość wyższa, niż uzyskiwana często w innych badaniach. Podkreślono już, że wysokie wyniki uzyskano dla czasopism internetowych *D-Lib Magazine* – 10 lat (McCown et al., 2005, 5) oraz *Information Research* – 14.04 lat (Moghaddam et al., 2010, 71–72). W przypadku polskiego czasopisma *Biuletyn EBIB* wynik jest dużo niższy.

Najczęściej występującą odpowiedzią HTTP w przypadku nieaktywnych łączy był kod 404 stanowiący 75.2% odpowiedzi. Jest to wynik podobny do uzyskanych w innych badaniach. Odmienność polega na dużo większym odsetku tego typu odpowiedzi w prezentowanym badaniu polskich czasopism. Inne kody stanowiły niewielki odsetek wszystkich błędnych odpowiedzi. Istotnym problemem pozostaje zatem zmiana układu serwisów i przenoszenie zasobów.

W badaniach uwzględniono również hipotezę mówiącą, że budowa adresu URL może mieć wpływ na jego aktualność. Założenie to przyjęto na podstawie wyników stanu badań, gdzie badacze (np. Spinellis, 2003; Goh & Ng, 2007; Prithviraj & Kumar, 2014; Kumar et al., 2015) zaobserwowali pewne zależności w tym kontekście w odniesieniu do długości adresu URL oraz jego głębokości. W trakcie badania nie stwierdzono zależności między długością adresu URL a dostępnością zasobu. Liczba znaków dla zasobów niedostępnych jest co prawda nieco większa, jednak nie jest to czynnik, który z dużym prawdopodobieństwem pozwala określić stabilność łączy. Podobnie nie stwierdzono zależności pomiędzy poziomem dostępności zasobu a głębokością adresu URL.

W przeprowadzonym badaniu popularne rodzaje domen nie różnią się od wyników wcześniej przeprowadzonych analiz. Najczęściej cytowaną domeną była domena *org*, a następnie *com*, *edu* oraz *gov*. Hipotezy dotyczące popularności poszczególnych domen zostały przedstawione w części poświęconej stanowi badań. Łączy z najpopularniejszej domeny *org* należały do jednych z najbardziej trwałych. Z kolei do najmniej stabilnych linków należały, oprócz łączy zawierających nazwę domeny wskazującej na serwisy Unii Europejskiej *europa.eu*, te należące do polskich instytucji rządowych. Wskazuje to na dużą zmienność tych stron. Być może wśród cytowanych dokumentów dużą liczbę stanowiły różnego rodzaju ogłoszenia, które z racji swego tymczasowego charakteru są usuwane ze stron jako nieaktualne. W celu potwierdzenia tej hipotezy należałoby dokładnie przeanalizować rodzaje cytowanych zasobów ze stron rządowych.

W badaniu podjęto również próbę analizy możliwości odszukania zasobów zakwalifikowanych jako niedostępne wykorzystując internetowe archiwum Wayback Machine. Za jego pomocą udało się uzyskać dostęp do 53% wcześniej niedostępnych dokumentów. Wskazuje to na możliwość manualnego odnalezienia niedostępnych pierwotnie zasobów, co jest jednak dość kłopotliwe.

Oprócz niedostępności zasobów pod linkami zamieszczonymi w bibliografiach załącznikowych, problemem wydaje się być również zmiana treści cytowanych dokumentów sieciowych. W badaniu określono poziom tego zjawiska na 12.28%. Z jednej strony odsetek ten nie jest wysoki, jednak w połączeniu z liczbą niedostępnych zasobów, powoduje poważne ograniczenia w dostępie do cytowanych materiałów zmniejszając tym samym możliwość późniejszej analizy prac naukowych.

## 6. Zakończenie

Na koniec warto zastanowić się nad możliwymi do wdrożenia rozwiązaniami, które zapewniłyby większą dostępność cytowanych zasobów internetowych. Ciągłe aktualne zostaje zalecenie dla redakcji czasopism, które powinny prowadzić spójną politykę odnośnie cytowania materiałów sieciowych. Wspomniano już o tym, że redakcje czasopism powinny sprawdzać aktywność łączy przed publikacją artykułów. Tymczasowe rozwiązanie może opierać się na istniejących, amerykańskich serwisach archiwizacji Internetu, takich jak Internet Archive, WebCite lub Perma. Na początku artykułu wspomniano o ograniczeniach związanych z korzystaniem z tych narzędzi. Wydaje się, że mimo ich występowania redakcje czasopism mogłyby umożliwić zapisywanie obok oryginalnego adresu URL cytowanego zasobu, łączy do serwisu archiwizacyjnego. Na przykład styl cytowania Bluebook umożliwia zamieszczanie linków do serwisu perma.cc (Kribble, 2015). Bardziej trwałe rozwiązanie powinno opierać się na stworzeniu polskiego serwisu archiwizującego cytowane zasoby, czy szerzej dokumenty ważne dla polskiej nauki i kultury. W celu jego stworzenia można by się posłużyć istniejącymi zasobami, np. kod Perma jest dostępny w serwisie GitHub (GitHub, b.d.). Obsługą takiego serwisu powinna zająć się instytucja z zapewnionym stałym finansowaniem. Wśród bibliotek najlepiej przygotowana do tego zadania wydaje się być Biblioteka Narodowa, dysponująca dobrym zapleczem technicznym. Inne rozwiązanie może polegać na oparciu archiwizacji na sieci instytucji, podobnie jak zrobiono to w Stanach Zjednoczonych w przypadku serwisu Perma. W Polsce największa liczba bibliotek naukowych i akademickich jest skupiona wokół Narodowego Uniwersalnego Katalogu Centralnego (NUKAT). Wydaje się, że Centrum NUKAT zarządzające projektem dysponuje sporym doświadczeniem w zakresie organizacji współpracy między różnymi bibliotekami. Inną instytucją, która mogłaby podjąć się tego zadania może być Ośrodek Przetwarzania Informacji – Państwowy Instytut Badawczy, operator Polskiej Bibliografii Naukowej. Niezależnie od przyjętego rozwiązania organizacyjnego wydaje się, że stworzenie polskiego serwisu archiwizacyjnego jest sprawą istotną i wymagającą rozwiązania w najbliższej przyszłości. Dalsze badania cytowanych zasobów internetowych mogłyby objąć polskie czasopisma z innych dziedzin wiedzy. Wydaje się jednak, że ważniejszą sprawą jest opracowanie metodyki archiwizacji zasobów i wdrożenie rozwiązania.

## Bibliografia

- archive.is (b.d.). [online], [09.06.2016], <http://archive.is/>
- Carnevale, R. J.; Aronsky, D. (2007). The life and death of URLs in five biomedical informatics journals [online]. *International Journal of Medical Informatics*, 76(4), 269–273, [09.06.2016], <http://doi.org/10.1016/j.ijmedinf.2005.12.001>
- Casserly, M. F.; Bird, J. E. (2003). Web citation availability: Analysis and implications for scholarship [online]. *College and Research Libraries*, 64(4), 300–317, [09.06.2016], <http://doi.org/10.5860/crl.64.4.300>
- Dimitrova, D. V.; Bugeja, M. (2007). The half-life of internet references cited in communication journals [online]. *New Media & Society*, 9(5), 811–826, [09.06.2016], <http://doi.org/10.1177/1461444807081226>
- GitHub (b.d.). GitHub – harvard-lil/perma: Indelible links [online]. [09.06.2016], <http://github.com/harvard-lil/perma>
- Goh, D. H.-L.; Ng, P. K. (2007). Link decay in leading information science journals [online]. *Journal of the American Society for Information Science and Technology*, 58(1), 15–24, [09.06.2016], <http://doi.org/10.1002/asi.20513>
- Gul, S.; Mahajan, I.; Ali, A. (2014). The growth and decay of URLs citation: A case of an online library & information science journal [online]. *Malaysian Journal of Library and Information Science*, 19(3), 27–39, [09.06.2016], <http://ejum.fsktm.um.edu.my/ArticleInformation.aspx?ArticleID=1502>
- Habibzadeh, P. (2013). Decay of references to Web sites in articles published in general medical journals: mainstream vs small journals [online]. *Applied Clinical Informatics*, 4(4), 455–464, [09.06.2016], <http://doi.org/10.4338/ACI-2013-07-RA-0055>
- HEADMasterSEO. (2016). Wersja 1.07. [online]. [09.06.2016], <http://headmasterseo.com/>
- HTTP (2004). Hypertext Transfer Protocol (HTTP/1.1): Semantics and Content [online], [09.06.2016], <http://tools.ietf.org/html/rfc7231>
- Internet Archive (b.d.). Internet Archive Frequently Asked Questions [online], [09.06.2016], <http://archive.org/about/faqs.php>
- Jabłońska, K.; Sobieraj, A. (2013). Metodyka dobierania próby badawczej w naukach społecznych [online]. *Bezpieczeństwo i Technika Pożarnicza*. 32(4), 31–36, [09.06.2016], <http://doi.org/10.12845/bitp.32.4.2013.3>
- Klein, M.; Van de Sompel, H.; Sanderson, R.; Shankar, H.; Balakireva, L.; Zhou, K.; Tobin, R. (2014). Scholarly context not found: one in five articles suffers from reference rot [online]. *PloS One*, 9(12), [09.06.2016], <http://doi.org/10.1371/journal.pone.0115253>
- Koehler, W. (2004). A longitudinal study of Web pages continued: a consideration of document persistence [online]. *Information Research: An International Electronic Journal*, 9(2), [09.06.2016], <http://www.informationr.net/ir/9-2/paper174.html>
- Kribble, M. (2015). 20th edition of the Bluebook: including Perma.cc! [online], [09.06.2016], <http://etseq.law.harvard.edu/2015/06/20th-edition-of-the-bluebook-including-perma-cc/>
- Kumar, D. V.; Kumar, B. T. S.; Parameshwarappa, D. R. (2015). URLs Link Rot: Implications for Electronic Publishing [online]. *World Digital Libraries*, 8(1), 59–65, [09.06.2016], <http://doi.org/10.18329/09757597/2015/8105>
- Mardani, A. H.; Sangari, M. (2013). An analysis of the availability and persistence of web citations in Iranian LIS journals [online]. *International Journal of Information Science and Management*, 11(1), 29–42, [09.06.2016], <http://ijism.ricest.ac.ir/index.php/ijism/article/view/202>
- McCown, F.; Sheffan, C.; Nelson, M. L.; Bollen, J. (2005). The Availability and Persistence of Web References in D-Lib Magazine [online]. In: *5th International Web Archiving Workshop (IWA05)*, [09.06.2016], <http://arxiv.org/abs/cs/0511077v1>
- Moghaddam, A. I.; Saberi, M. K.; Esmaeel, S.M. (2010). Availability and Half-life of Web References Cited in Information Research Journal: A Citation Study [online]. *International Journal of*

- Information Science and Management (IJISM)*, 8(2), 57–75, [09.06.2016], <http://ijism.ricest.ac.ir/index.php/ijism/article/view/49>
- Oguz, F.; Koehler, W. (2016). URL decay at year 20: A research note [online]. *Journal of the Association for Information Science and Technology*, 67(2), 477–479, [09.06.2016], <http://doi.org/10.1002/asi.23561>
- Perma.cc (b.d.). [online], [09.06.2016], <http://perma.cc/>
- Perma.cc Docs (b.d.). [online], [09.06.2016], <http://perma.cc/docs>
- Prithviraj, K. R.; Kumar, B. T. S. (2014). Corrosion of URLs: Implications for electronic publishing [online]. *IFLA Journal*, 40(1), 35–47, [09.06.2016], <http://doi.org/10.1177/0340035214526529>
- Sadat-Moosavi, A.; Isfandyari-Moghaddam, A.; Tajeddini, O. (2012). Accessibility of online resources cited in scholarly LIS journals: A study of Emerald ISI-ranked journals [online]. *Aslib Proceedings*, 64(2), 178–192, [09.06.2016], <http://doi.org/10.1108/00012531211215196>
- Sampath Kumar, B. T.; Prithvi Raj, K. R. (2012). Availability and persistence of web citations in Indian LIS literature [online]. *The Electronic Library*, 30(1), 19–32, [09.06.2016], <http://doi.org/10.1108/02640471211204042>
- Spinellis, D. (2003). The decay and failures of web references [online]. *Communications of the ACM*, 46(1), 71–77, [09.06.2016], <http://doi.org/10.1145/602421.602422>
- Van de Sompel, H.; Klein, M.; Jones, S.M. (2016). Persistent URIs Must Be Used To Be Persistent [online]. In: *Proceedings of the 25th International Conference Companion on World Wide Web*, 119–120, [09.06.2016], <http://arxiv.org/abs/1602.09102v1>
- Wayback Machine (b.d.). [online], [09.06.2016], <http://archive.org/web>
- Wayback Machine APIs (b.d.). [online], [09.06.2016], [http://archive.org/help/wayback\\_api.php](http://archive.org/help/wayback_api.php)
- WebCite (b.d.). [online]. [09.06.2016], <http://www.webcitation.org/>
- Wu, Z. (2008). An empirical study of the accessibility of web references in two Chinese academic journals [online]. *Scientometrics*, 78(3), 481–503, [09.06.2016], <http://doi.org/10.1007/s11192-007-1951-1>
- Yang, S.; Han, R.; Ding, J.; Song, Y. (2012). The distribution of Web citations [online]. *Information Processing and Management*, 48, 779–790, [09.06.2016], <http://doi.org/10.1016/j.ipm.2011.10.002>
- Zhou, K.; Grover, C.; Tobin, R.; Klein, M. (2015). No More 404s: Predicting Referenced Link Rot in Scholarly Articles for Pro-Active Archiving [online]. In: *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries* (2015th-June ed., Vol. 2015-June, 233–236), [09.06.2016], <http://doi.org/10.1145/2756406.2756940>
- Zittrain, J.; Albert, K.; Lessig, L. (2014). Perma: Scoping and addressing the problem of link and reference rot in legal citations [online]. *Legal Information Management*, 14(02), 88–99, [09.06.2016], <http://doi.org/10.1017/S1472669614000255>
- 

## Web Citations in Polish Library and Information Science Journals: the Analysis of URL Validity

### Abstract

**Purpose/Thesis:** This paper is intended to provide a comprehensive literature review on the issue of link rot, that is, the inaccessibility of Web citations in academic journals and describe this problem based on articles published in four Polish journals in the field of library and information science.

**Approach/Research methods:** Using critical literature review, the authors identified main areas of interest related to the problem of link rot in scientific journals, as well as key methodological approaches.

The subject of empirical studies were web citations in research papers published in four Polish library and information science journals from 2009 to 2015: *Bibliotheca Nostra*, *Śląski Kwartalnik Naukowy*, *Biuletyn EBIB*, *Przegląd Biblioteczny*, *Zagadnienia Informacji Naukowej - Studia Informacyjne*. The level of URL accessibility was examined automatically using HTTP server responses.

**Results and conclusions:** Over 37% of examined papers contained Web citations. At the time of research over 70% of Web citations were still accessible, but this level was different for each journal. The half-life value of Web citations is approximately 7 years.

**Originality/Value:** The paper presents a comprehensive literature review on the problem of link rot which can be used as a background for further investigations related to this issue and as a reference for methodological considerations. According to the authors' knowledge, the issue of link rot had not been so far the subject of interest for Polish library and information science researchers. Therefore, the results of the research presented in the paper are the first attempt to describe the scale of this problem.

### Keywords

Citation analysis. Polish scientific journals. Library and Information Science journals. The half-life value of Web citations. Web citations.

---

*Dr MARCIN ROSZKOWSKI – adiunkt w Katedrze Informatologii na Wydziale Dziennikarstwa, Informacji i Bibliologii Uniwersytetu Warszawskiego. Jego zainteresowania badawcze obejmują problematykę organizacji wiedzy w środowisku sieciowym, metadanych oraz technologii Sieci Semantycznej. Najważniejsze publikacje: M. Roszkowski: Kartoteka haseł wzorcowych jako usługa sieciowa – automatyczna identyfikacja nazw osobowych z wykorzystaniem kartoteki VIAF. W: Bibliografia – teoria, praktyka, dydaktyka, Warszawa 2016 (w druku); M. Roszkowski: Do czego potrzebne są nam ontologie? Charakterystyka funkcjonalna ontologii jako narzędzi reprezentacji wiedzy. W: Nauka o informacji w okresie zmian, Warszawa 2014, 151–167; M. Roszkowski: Od MARC 21 do Semantic Web: reprezentacja metadanych bibliograficznych w środowisku sieciowym. W: Bibliografi@: źródła – standardy – zasoby. Warszawa 2013, 5–25.*

Kontakt z autorem:

[m.roszkowski@uw.edu.pl](mailto:m.roszkowski@uw.edu.pl)

Katedra Informatologii

Wydział Dziennikarstwa, Informacji i Bibliologii

Uniwersytet Warszawski

ul. Nowy Świat 69

00–927 Warszawa

*Dr BARTŁOMIEJ WŁODARCZYK – adiunkt w Katedrze Bibliografii i Dokumentacji na Wydziale Dziennikarstwa, Informacji i Bibliologii Uniwersytetu Warszawskiego. Jego zainteresowania badawcze dotyczą problematyki organizacji informacji i wiedzy. Najważniejsze publikacje: B. Włodarczyk: Mapy tematów jako podstawa semantyzacji języka haseł przedmiotowych. Warszawa 2016; B. Włodarczyk: Deskryptory formalne w Bibliotece Kongresu a tematy formalne w języku haseł przedmiotowych Biblioteki Narodowej i w języku KABA. Przegląd Biblioteczny 2014, 4, 518–536; B. Włodarczyk: Topic Map as a Method for the Development of Subject Headings Vocabulary: An Introduction to the Project of the National Library of Poland. Cataloging & Classification Quarterly 2013, 7, 816–829.*

Kontakt z autorem:

[bm.wlodarczyk@uw.edu.pl](mailto:bm.wlodarczyk@uw.edu.pl)

Katedra Bibliografii i Dokumentacji

Wydział Dziennikarstwa, Informacji i Bibliologii

Uniwersytet Warszawski

ul. Nowy Świat 69

00–927 Warszawa

# Polskie repozytoria instytucjonalne jako miejsce dla otwartych zasobów naukowych i edukacyjnych

Ewa Głowacka

*Instytut Informacji Naukowej i Bibliologii, Uniwersytet Mikołaja Kopernika w Toruniu*

## Abstrakt

**Cel/teza:** Otwarte repozytoria cyfrowe to obecnie bardzo dynamicznie rozwijający się rodzaj źródeł informacji naukowej. Celem artykułu jest omówienie wyników badań nad stanem funkcjonujących w Polsce naukowych repozytoriów tworzonych przez wyższe uczelnie i inne jednostki naukowe.

**Koncepcja/metody badań:** Przedstawiono polskie repozytoria naukowe od strony ich rozwoju, liczebności, struktury zasobów oraz funkcji. Zdecydowano się na analizę jedynie takich inicjatyw, które gromadzą zasoby nowe, a nie – jak w przypadku licznie funkcjonujących w Polsce bibliotek cyfrowych – zdigitalizowane kolekcje dziedzictwa narodowego. Pomocą w wyborze serwisów były listy repozytoriów naukowych zawartych w portalu Uwolnij Naukę, EBIB oraz w bazie Agregator CEON. Każdy z serwisów przebadano oddzielnie. Dla przedstawienia liczby inicjatyw zagranicznych oraz funkcji repozytoriów polskich i obcych wykorzystano serwis OpenDOAR oraz literaturę przedmiotu.

**Wyniki i wnioski:** Autorce udało się odnaleźć 33 aktualnie funkcjonujące polskie otwarte repozytoria naukowe. Ogółem zgromadziły one ok. 270 tys. różnorodnych zasobów, przede wszystkim prac dyplomowych i doktorskich, artykułów z czasopism oraz materiałów dydaktycznych. Projekty te rozwijają się bardzo aktywnie, zważywszy na to, że mają jeszcze bardzo krótką historię. Tworzenie repozytoriów przynosi wiele korzyści dla uczelni. Zaliczymy do nich wzrost jej prestiżu jako skutek wzrostu widoczności badań naukowych prowadzonych na wydziałach, promowanie kadry naukowej, wspieranie edukacji studentów poprzez ułatwienie im dostępu do materiałów dydaktycznych.

**Zastosowania praktyczne:** Wyniki analizy polskich repozytoriów pokazują obraz ich obecnego stanu rozwoju oraz mogą stanowić zachętę do deponowania prac oraz korzystania z ich zasobów.

## Słowa kluczowe

Open Access. Polskie repozytoria instytucjonalne. Polskie repozytoria dziedziczne. Zasoby cyfrowe.

*Otrzymano: 21 lipca 2016. Zrecenzowano: 20 sierpnia 2016. Zaakceptowano: 1 września 2016.*

## 1. Wstęp

Współczesne przemiany w komunikacji naukowej i edukacji często rozpatrywane są w aspekcie zjawiska „Nauka 2.0”, czyli procesów oraz tendencji związanych z wykorzystaniem w nauce współczesnych technologii informacyjno-komunikacyjnych. Głównymi przejawami rozwoju zastosowań nowych technologii i nowego, bardziej otwartego spojrzenia na komunikację naukową, jest rozwój idei Open Access (OA) oraz przepływ informacji naukowej nie tylko przez tradycyjne, dotychczas rozwijane kanały (publikacje tradycyjne i elektroniczne, głównie książki i czasopisma), lecz także wykorzystanie w komunikacji naukowej

technologii Web 2.0 i związanych z nią mediów społecznościowych (np. blogów, serwisów społecznościowych i forów internetowych). OA to ruch społeczny działający na rzecz swobodnego upowszechnienia wiedzy, który w praktyce oznacza wolny i nieograniczony dostęp do cyfrowych publikacji naukowych oraz otwarty model komunikacji naukowej.

Jedną ze strategii ruchu OA – „zielona droga” to tworzenie otwartych archiwów i repozytoriów, które gromadzą i udostępniają treści w formie cyfrowej:

Repozytorium otwarte (ang. *open repository*, *open archive*) jest to narzędzie służące zarządzaniu i długoterminowemu przechowywaniu dokumentów cyfrowych. Repozytoria wspierają rozwój badań naukowych, procesy uczenia się oraz prace administracyjne. Stosują one otwarte standardy tak, aby ich zasób był w pełni dostępny. Standardy te pozwalają na stosowanie mechanizmów, które umożliwiają importowanie, eksportowanie, przechowywanie i wyszukiwanie cyfrowych dokumentów w repozytorium. Zawartość otwartego archiwum może być bardzo różna i może służyć różnym celom oraz użytkownikom (Rychlik & Karwasińska, b.d.).

W cyfrowych repozytoriach i archiwach mogą się znaleźć zarówno teksty niepublikowane i nierecenzowane (preprinty), jak i publikacje, które przeszły proces recenzji i zostały zaakceptowane do druku (postprinty). Repozytorium może gromadzić wszystkie formy dokumentów czy postaci danych, które powstają na uczelni: artykuły z czasopism (postprinty), rękopisy (preprinty), dane bibliograficzne, książki lub ich części, np. rozdziały, materiały konferencyjne, postery, prezentacje, prace magisterskie i doktoraty, prace habilitacyjne, materiały edukacyjne, multimedia i materiały audio-wideo, inne (Bednarek-Michalska, 2011). Obecnie rozwijane są różne typy repozytoriów, do których zaliczamy:

- repozytoria instytucjonalne (ang. *institutional repository*) – cyfrowe archiwa gromadzące i rozpowszechniające dorobek intelektualny konkretnej społeczności naukowej (uczelni, instytutu badawczego),
- repozytoria dziedzinowe (ang. *subject repository*), które gromadzą dokumenty spójne tematycznie,
- repozytoria danych badawczych (tzw. danych surowych) (ang. *data repository*) – archiwa gromadzące surowe dane eksperymentalne, np. genomy bakterii, zwierząt czy człowieka (Rychlik & Karwasińska, b.d.).

Repozytoria, w przeciwieństwie do czasopism, nie mają własnych recenzentów i nie finansują procesu recenzowania. Archiwa takie zazwyczaj funkcjonują przy ośrodkach akademickich lub instytucjach badawczych. Materiały gromadzi się w nich na zasadzie autoarchiwizacji (ang. *self-archiving*), tzn. autorzy sami deponują kopie swoich prac. Repozytoria zawierają publikowany i niepublikowany dorobek naukowy jednego lub kilku ośrodków w postaci dokumentów utworzonych jako elektroniczne (ang. *born digital*).

W polskiej literaturze przedmiotu polskie repozytoria otwarte kilka lat temu analizowała Emilia Karwasińska (Karwasińska, 2012) przedstawiając dziewięć takich serwisów. W niniejszym artykule omówiono działające obecnie polskie otwarte repozytoria naukowe, analizując ich rozwój, wielkość zasobów i podstawowe funkcje. Badania ukończono 23 sierpnia 2016 r. Zdecydowano się na analizę jedynie takich inicjatyw, które gromadzą zasoby nowe, a nie – jak w przypadku licznie funkcjonujących w Polsce bibliotek cyfrowych – zdigitalizowane kolekcje dziedzictwa narodowego. W odnalezieniu polskich otwartych repozytoriów wykorzystano wykazy repozytoriów naukowych z portalu Uwolnij Naukę<sup>1</sup>,

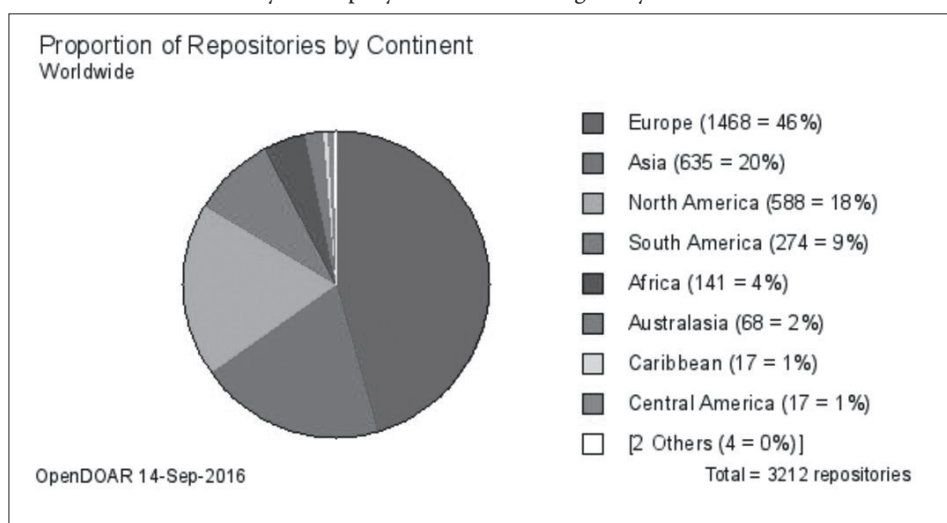
<sup>1</sup> <http://uwolnijnauke.pl/baza-wiedzy/>

EBIB<sup>2</sup> oraz Agregatora CEON<sup>3</sup>. Każdy z serwisów WWW repozytoriów cyfrowych przebadano oddzielnie. Dla przedstawienia liczby inicjatyw zagranicznych oraz funkcji repozytoriów polskich i obcych wykorzystano serwis Directory of Open Access Repositories (OpenDOAR)<sup>4</sup> oraz literaturę przedmiotu. Liczbę publikacji, strukturę oraz lata powstania repozytoriów ustalano analizując ich strony WWW.

## 2. Rozwój oraz funkcje repozytoriów

Rozwój repozytoriów związany jest z procesem ewolucji systemu komunikacji naukowej w kierunku nadania znaczenia również nieformalnym, wirtualnym kanałom komunikacji. Od 2006 r. na całym świecie rośnie liczba repozytoriów OA – obecnie jest ich już ponad 3000, z czego najwięcej jest w Europie (Rys. 1)<sup>5</sup>.

Rys. 1. Repozytoria na świecie wg kontynentów



Źródło: OpenDOAR, dostęp: 14.09.2016.

Repozytorium, oprócz swoich funkcji edukacyjnych, upowszechniających, archiwizujących i zabezpieczających dorobek naukowy, może być traktowane jako platforma wydawnicza, daje ono bowiem możliwość archiwizowania kolekcji czasopism uczelnianych. Ciekawym zjawiskiem jest też wydawanie czasopism-nakładek (ang. *overlay journals*) bazujących na zasobie repozytorium. Jest to wysokiej jakości czasopismo Open Access, którego zasób znajduje się w jednym lub większej liczbie repozytoriów. Selekcjonuje ono

<sup>2</sup> [http://www.ebib.pl/?page\\_id=673](http://www.ebib.pl/?page_id=673)

<sup>3</sup> <http://agregator.ceon.pl/>

<sup>4</sup> <http://www.opendoar.org/>

<sup>5</sup> Katalog OpenDOAR rejestruje repozytoria otwarte, ale również biblioteki cyfrowe gromadzące dziedzictwo kulturowe.

teksty dostępne w otwartym trybie. Bywa też tak, że część czasopism posiada zasób wydany głównie przez komercyjnych wydawców z odsyłaczami do pełnych tekstów zarchiwizowanych w repozytoriach w formie pre- lub postprintu. Przykładem czasopisma nakładki jest *Lund Medical Faculty Monthly*, które gromadzi artykuły z czasopism komercyjnych publikowane przez autorów Wydziału Medycznego na Lund University. Część uczelnianych repozytoriów tworzy dodatkowe usługi służące promocji naukowców. Jedną z nich jest tzw. e-portfolio, które daje możliwość umieszczenia swojego profilu w repozytorium. W jednym miejscu umieszczane są informacje o autorze (np. życiorys naukowy), prezentowane są publikacje naukowe, oraz popularnonaukowe, zdobyte granty, nagrody oraz zainteresowania naukowe. Takie usługi świadczą np. University of Rochester, Bond University oraz Digital Academic Repository Universiteit van Amsterdam (Rychlik, 2011, 12).

W wielu krajach istnieją instytucje koordynujące budowanie systemów informacyjnych nie tylko dla nauki, ale także dla edukacji. Przykładem bardzo dobrej praktyki w tym zakresie jest Wielka Brytania i działający w tym kraju Joint Information Systems Committee (JISC). Komitet JISC przeprowadził wiele badań związanych z Open Access, w tym także na temat modeli narodowych platform naukowych. W wyniku tych badań wyróżniono trzy rozwiązania (Bednarek-Michalska, 2011):

- model scentralizowany – metadane i zasoby są gromadzone w jednym miejscu,
- model rozproszony – istnieje wiele instytucjonalnych repozytoriów oraz wspólna wyszukiwarka,
- model agregacyjny lub hybrydowy – zasoby są gromadzone w różnych miejscach, ale metadane są agregowane w centralnej bazie danych i tam następuje proces wyszukiwania i ujednolicania zasobu.

Współcześnie tworzone są też narodowe portale zapewniające dostęp do zasobów naukowych, np. holenderski projekt NARCIS – The gateway to scholarly information in The Netherlands<sup>6</sup> lub HAL<sup>7</sup> – wielodyscyplinarne archiwum francuskie (Bednarek-Michalska, 2011, 51).

Przedstawiciele nauk ścisłych coraz częściej postulują rozszerzenie Open Access o udostępnianie w repozytoriach nie tylko artykułów, lecz również danych badawczych, co może znacząco wpłynąć na postęp naukowy. Powstają więc repozytoria danych badawczych (tzw. danych surowych) Tworzone przez wielkie konsorcja instytucji naukowych, centra badawcze a czasem biblioteki uniwersyteckie. Obecny system publikacji nie sprzyja upowszechnianiu danych, również sami naukowcy niechętnie je udostępniają. Dostępność danych jest także ograniczana przez uwarunkowania prawne oraz czynniki ekonomiczne i techniczne. Brak odpowiednich standardów publikacji surowych danych powoduje również, że ich gromadzenie, przetwarzanie i agregowanie jest utrudnione. W 2009 r. w Londynie podczas spotkania przedstawicieli British Library, Technical Information Center of Denmark, TU Delft Library, National Research Council's Canada Institute for Scientific and Technical Information (NRC-CISTI), California Digital Library, Purdue University oraz German National Library of Science and Technology postanowiono utworzyć stowarzyszenie zajmujące się upowszechnianiem informacji o danych badawczych. Pierwszym wykazem repozytoriów danych badawczych był serwis internetowy DataCite<sup>8</sup> (Bednarek-Michalska, 2012).

<sup>6</sup> <http://www.narcis.nl/>

<sup>7</sup> <http://hal.archives-ouvertes.fr/>

<sup>8</sup> <http://www.datacite.org/>

### 3. Polskie otwarte repozytoria naukowe

Autorce udało się ustalić, że w Polsce obecnie funkcjonują 33 otwarte repozytoria cyfrowe (Tab. 1). Mają one charakter archiwów gromadzących dorobek jednej uczelni, bądź wybranych jednostek naukowych uczelni wyższych i innych instytucji naukowych. Udostępniają materiały dydaktyczne oraz dorobek naukowy pracowników i doktorantów. Ich struktura odwzorowuje strukturę uczelni. Wyjątkiem jest Repozytorium UW, które gromadzi tylko rozprawy doktorskie. Materiały edukacyjne są udostępniane w ramach kolekcji wydziałowych w osobnych podkolekcjach. Ten rodzaj kolekcji cyfrowych rozwija się w Polsce bardzo aktywnie, zważywszy na to, że mają jeszcze krótką historię – pierwsze (repozytorium PCSS) powstało w 2006 r., ale najwięcej tego typu projektów zaczęto rozwijać w latach 2012–2013.

Tabela 1. Polskie otwarte repozytoria naukowe (stan na 23.08.2016)

| Nazwa repozytorium                                                      | Liczba publikacji | Rok powstania |
|-------------------------------------------------------------------------|-------------------|---------------|
| 1                                                                       | 2                 | 3             |
| Repozytorium Cyfrowe Instytutów Naukowych Polskiej Akademii Nauk (RCIN) | 58 566            | 2008          |
| Repozytorium Politechniki Warszawskiej                                  | 57 170*           | 2013          |
| Repozytorium Uniwersytetu Jagiellońskiego                               | 28 404            | 2015          |
| Repozytorium Uniwersytetu Łódzkiego (RUŁ)                               | 16 819            | 2011          |
| Repozytorium Uniwersytetu A. Mickiewicza (AMUR)                         | 12 939            | 2009          |
| Repozytorium Politechniki Krakowskiej (SUW)                             | 8151              | 2009          |
| Repozytorium Centrum Otwartej Nauki (CEON)                              | 8727              | 2012          |
| Repozytorium Krakowskiej Akademii im. A.F. Modrzewskiego (eRIKA)        | 7283              | 2013          |
| Repozytorium Uniwersytetu Ekonomicznego we Wrocławiu (WIR)              | 7264              | 2014          |
| Repozytorium Akademii Wychowania Fizycznego w Warszawie                 | 4368              | b.d.          |
| Repozytorium Uniwersytetu w Białymstoku (RUB)                           | 3928              | 2013          |
| Repozytorium Uniwersytetu M. Kopernika w Toruniu (RUMAK)                | 3490              | 2012          |
| Repozytorium Uniwersytetu im. Kazimierza Wielkiego                      | 3292              | 2013          |
| Repozytorium Politechniki Poznańskiej                                   | 1293              | 2011          |
| Repozytorium Uniwersytetu Rzeszowskiego                                 | 1280              | 2013          |
| Repozytorium Uniwersytetu Wrocławskiego                                 | 1237              | 2013          |
| Repozytorium Politechniki Łódzkiej (Cyrena)                             | 1121              | 2015          |
| Repozytorium Uniwersytetu Warszawskiego                                 | 1051              | 2012          |
| Repozytorium Uniwersytetu Techniczno-Przyrodniczego w Bydgoszczy        | 918               | 2011          |
| Repozytorium Wiedzy Politechniki Wrocławskiej (SUW)                     | 913               | 2014          |
| Biblioteka Cyfrowa Uniwersytetu Papieskiego Jana Pawła II               | 910               | 2013          |

| 1                                                                                                                                           | 2   | 3    |
|---------------------------------------------------------------------------------------------------------------------------------------------|-----|------|
| Repozytorium UP w Krakowie                                                                                                                  | 848 | b.d. |
| Repozytorium WSB-NLU                                                                                                                        | 812 | 2013 |
| Repozytorium IBB PAN                                                                                                                        | 640 | 2010 |
| Repozytorium Instytucjonalne PCSS                                                                                                           | 640 | 2006 |
| Repozytorium Instytutu Medycyny Pracy w Łodzi (ECNIS)                                                                                       | 601 | b.d. |
| Repozytorium Uniwersytetu Przyrodniczo-Humanistycznego w Siedlcach                                                                          | 556 | 2015 |
| Repozytorium Wydziału Elektrycznego Politechniki Wrocławskiej (ENY)                                                                         | 430 | 2008 |
| Biblioteka Cyfrowa Katedry Lingwistyki Formalnej UW                                                                                         | 387 | 2008 |
| Repozytorium Publikacji Naukowych Politechniki Śląskiej (Repolis)                                                                           | 353 | 2012 |
| Otwarte Repozytorium Nauk Historycznych (Lectorium)                                                                                         | 257 | 2014 |
| Repozytorium Dolnośląskiej Szkoły Wyższej (oPUB)                                                                                            | 257 | 2014 |
| Repozytorium Polsko-Japońskiej Wyższej Szkoły Technik Komputerowych                                                                         | 177 | b.d. |
| * Baza wiedzy PW liczy 104823 obiektów, takich jak publikacje, opisy projektów, wydarzeń, profile naukowców. Samych publikacji jest 57 170. |     |      |

Źródło: Opracowanie własne na podstawie analizy danych z portalu Uwolnij Naukę, Agregatora CEON oraz witryn repozytoriów.

### 3.1. Repozytoria gromadzące dorobek pracowników i studentów pojedynczych uczelni

Obecnie największą tego typu inicjatywą jest Repozytorium Politechniki Warszawskiej, które stanowi integralną część systemu Bazy Wiedzy PW. W repozytorium archiwizuje się w wersji cyfrowej pełne teksty materiałów dokumentujących prowadzone prace, w tym: monografie, artykuły z czasopism, rozdziały z książek i raporty, patenty, a także teksty utworów stanowiących podstawę do nadawania stopni i tytułów naukowych, prace licencjackie i magisterskie, projekty badawcze. Repozytorium działa od 2013 r. Koncepcja i projekt techniczny oraz oprogramowanie Bazy Wiedzy PW – OMEGA-PSIR zrealizowane zostały na Wydziale Elektroniki i Technik Informacyjnych PW w ramach projektu SYNAT<sup>9</sup>. Do dużych inicjatyw tego rodzaju należy również Repozytorium Uniwersytetu Łódzkiego (RUŁ), które zajęło czwarte miejsce wg międzynarodowego Ranking Web of Repositories<sup>10</sup> w kategorii repozytoriów polskich oraz pierwsze miejsce na liście Agregatora Centrum Otwartej Nauki (CEON)<sup>11</sup> w kategorii repozytoriów o największej liczbie dokumentów. Jednym z pierwszych, wzorcowych i bogatych w zasoby repozytoriów instytucjonalnych jest

<sup>9</sup> Projekt był realizowany w Polsce w okresie 08.2010–08.2013. Jego celem było utworzenie platformy dla sieciowych zasobów wiedzy dla nauki, edukacji.

<sup>10</sup> <http://repositories.webometrics.info/>, stan na lipiec 2016

<sup>11</sup> Informacja ze strony repozytorium RUŁ, <http://repozytorium.uni.lodz.pl>

też utworzone w 2009 r. Repozytorium Uniwersytetu im. Adama Mickiewicza — AMUR. Według wspomnianego rankingu Ranking Web of World Repositories, repozytorium AMUR zajęło 287 miejsce wśród najlepszych repozytoriów na świecie. Te cztery projekty są najzasobniejsze w dokumenty spośród polskich repozytoriów pojedynczych instytucji, jednak liczba materiałów cyfrowych zgromadzonych w Repozytorium PW jest kilkakrotnie większa niż wszystko, co udostępniono w ramach trzech pozostałych inicjatyw.

Kolejną grupę projektów stanowią archiwa liczące kilka tysięcy zbiorów. Największe z nich to Repozytorium Politechniki Krakowskiej (RPK) udostępniające zasoby naukowe oraz dydaktyczne autorstwa pracowników i studentów PK, a także publikacje wydawane przez wydawnictwo uczelni oraz cyfrowe wersje zbiorów Biblioteki Politechniki Krakowskiej (BPK). Następne pod względem liczebności zasobów są Repozytorium Krakowskiej Akademii im. Andrzeja Frycza Modrzewskiego (eRIKA) oraz zapoczątkowane w 2014 r. Repozytorium Uniwersytetu Ekonomicznego we Wrocławiu (WIR), które udostępniają już ok. 8 tys. publikacji. Mniej więcej o połowę mniej zasobów udostępnia Repozytorium Akademii Wychowania Fizycznego w Warszawie, Repozytorium Uniwersytetu w Białymstoku, Repozytorium UKW, a także Repozytorium Uniwersytetu Mikołaja Kopernika w Toruniu — RUMAK. Prawie wszystkie powstały w latach 2012–2013 i rozwijają się w podobnym tempie — kolekcja każdego z nich liczy ok. 3 tys. dokumentów, oprócz inicjatywy warszawskiej AWF udostępniającej ponad 4300 publikacji. Z kolei projektem koordynowanym przez UW oraz Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego (ICM), a więc również służącym całej uczelni, jest Repozytorium Uniwersytetu Warszawskiego (RUW). Jednak inicjatywa ta, jak już wspomniano, ogranicza swoją działalność tylko do udostępniania kolekcji rozpraw doktorskich. W przyszłości twórcy RUW zamierzają rozszerzyć jego zasób o inne materiały (artykuły, książki, materiały dydaktyczne i pokonferencyjne)<sup>12</sup>.

Do grupy repozytoriów instytucjonalnych o mniejszych kolekcjach (ok. 1000) zaliczają się: Uniwersytetu Jagiellońskiego, Uniwersytetu Techniczno-Przyrodniczego w Bydgoszczy, Uniwersytetu Wrocławskiego, Wyższa Szkoła Biznesu — National-Louis University w Nowym Sączu (WSB-NLU) — uwzględniono w niej również uczelnie partnerskie (Sidor, 2014), Repozytorium Publikacji Naukowych Politechniki Śląskiej, Uniwersytetu Przyrodniczo-Humanistycznego w Siedlcach, Dolnośląskiej Szkoły Wyższej oPUB, Polsko-Japońskiej Wyższej Szkoły Technik Komputerowych oraz Biblioteka Cyfrowa Uniwersytetu Papieskiego Jana Pawła II. Bogatym w zasoby jest Repozytorium Uniwersytetu Jagiellońskiego. Liczy obecnie ponad 28 tys. pozycji, jednak przeważającą część tej kolekcji stanowi bibliografia prac pracowników uczelni, która z czasem będzie uzupełniana o pełne teksty publikacji. Publikacje pełnotekstowe udostępniane są w odrębnym dziale *Zasoby repozytorium*, który w sierpniu 2016 r. liczył nieco ponad 1000 pozycji.

Obok repozytoriów instytucjonalnych, które powstają na wyższych uczelniach dla wszystkich ich jednostek, w Polsce tworzone są odrębne kolekcje cyfrowe dla poszczególnych instytutów naukowych i badawczych. W 2006 r. powstało Repozytorium Poznańskiego Centrum Superkomputerowo-Sieciowego (PCSS). Znajduje się w nim 640 publikacji dotyczących m.in. systemu dLibra, usług sieciowych, tworzenia środowiska informacyjnego bibliotek cyfrowych. Do repozytoriów obejmujących prace jednego wydziału

<sup>12</sup> Repozytorium Uniwersytetu Warszawskiego, <http://depotuw.ceon.pl/browse?type=title>

należy uruchomiona w 2008 r. Biblioteka Cyfrowa Katedry Lingwistyki Formalnej UW. Do omawianej grupy zalicza się również Repozytorium Wydziału Elektrycznego Politechniki Wrocławskiej (ENY) oraz Repozytorium Instytutu Biochemii i Biofizyki PAN (IBB PAS)<sup>13</sup>. W Polsce funkcjonuje też Repozytorium Instytutu Medycyny Pracy w Łodzi (ECNIS). Zbiory wymienionych repozytoriów nie przekraczają kilkuset pozycji.

### 3.2. *Repozytoria współtworzone przez wiele instytucji*

Obok repozytoriów instytucjonalnych, które są prowadzone przez pojedyncze instytucje w Polsce obecne są również kolekcje cyfrowe, które obejmują zespoły zasobów kilku lub kilkunastu różnych instytucji naukowych. Obecnie funkcjonuje kilka takich repozytoriów, z czego jedno wpisuje się w szerszą inicjatywę międzynarodową.

Do tej grupy repozytoriów należy Repozytorium Cyfrowe Instytutów Naukowych Polskiej Akademii Nauk (RCIN). Założeniem projektu jest utworzenie ponadregionalnego i multidyscyplinarnego repozytorium cyfrowego złożonego ze zdigitalizowanych materiałów archiwalnych, publikacji naukowych, dokumentacji badań oraz piśmienniczego dziedzictwa kulturowego wyselekcjonowanych ze zbiorów 16 polskich instytutów naukowych oraz ich bibliotek tworzących Konsorcjum (RCIN, 2014). RCIN funkcjonuje od 2008 r. i obecnie daje dostęp do ok. 60 tys. dokumentów.

Sieć Doskonałości ECNIS<sup>14</sup> (Environmental Cancer Risk, Nutrition and Individual Susceptibility – Rak Środowiskowy, Dieta i Indywidualna Wrażliwość) rozpoczęła działalność w 2005 r. w ramach Szóstego Programu Ramowego Badań i Rozwoju Unii Europejskiej. Zrzesza ona przedstawicieli europejskich ośrodków naukowych zajmujących się badaniami nad rakiem środowiskowym z 13 krajów (Belgia, Dania, Finlandia, Grecja, Holandia, Hiszpania, Francja, Niemcy, Polska, Szwecja, Węgry, Wielka Brytania, Włochy). Celem jej założycieli było stworzenie trwałej struktury instytucji partnerskich, umożliwiającej prowadzenie wspólnych badań nad rakiem. Wśród uczestników programu w 2007 r. powstała idea budowania otwartego repozytorium dziedzinowego zawierającego dokumenty będące wynikiem wykonanych prac. Koordynatorem działań sieci jest Instytut Medycyny Pracy (IMP) w Łodzi, a Bibliotece Naukowej IMP powierzono zadania związane z administrowaniem repozytorium (Przyłuska et al., 2007). Repozytorium ECNIS jest otwartym archiwum dokumentów, daje możliwość zamieszczania pre- i post-printów, raportów i prezentacji oraz wyników badań i materiałów źródłowych. Celem utworzenia repozytorium ECNIS było jak najszybsze archiwizowanie prac naukowych i udostępnianie ich szerokiej opinii publicznej. System wykorzystuje najnowsze zasady udostępniania pełnych tekstów dokumentów, posiada funkcję autoarchiwizacji, system alertów o nowych publikacjach, pozwala na tworzenie bezpośrednich połączeń do serwisów społecznościowych typu Connotea, Delicious, Citeulike, Facebook, zakładanie indywidualnych stron dla autorów, które informują o zainteresowaniach badawczych i dorobku publikacyjnym. Wpisuje się tym samym w nurt przemian otwartości w komunikacji naukowej. Jest pierwszym repozytorium tego typu budowanym przez polskie biblioteki (Przyłuska & Radomska, 2010).

<sup>13</sup> W sierpniu 2016 r. strona była nieaktywna.

<sup>14</sup> ECNIS, <http://ecnis.openrepository.com/ecnis/>

Instytucje naukowe, które nie prowadzą własnego repozytorium, mogą skorzystać z Repozytorium CEON (Repozytorium Centrum Otwartej Nauki) prowadzonego przez ICM UW. W tym repozytorium prace mogą zamieszczać, przechowywać i publicznie udostępniać wyłącznie zarejestrowani autorzy będący osobami fizycznymi. We współpracy z Biblioteką Uniwersytecką w Warszawie CEON prowadzi również repozytorium UW. W Archiwum CEON prezentowana jest też kolekcja wybranych książek z zakresu humanistyki i nauk społecznych, zdigitalizowaną w 2007 r. oraz serwis „Otwórz książkę”. Jest to cyfrowa kolekcja współczesnych książek naukowych, w której skład wchodzi przede wszystkim książki niedostępne w wersji drukowanej, poszukiwane ze względu na wartość naukową i edukacyjną. Obecnie serwis udostępnia ponad 300 tytułów, przede wszystkim z nauk społecznych (socjologia, psychologia, ekonomia i in.) oraz humanistycznych (historia, kulturoznawstwo, literaturoznawstwo)<sup>15</sup>.

Dla nauk humanistycznych powstaje od 2014 r. repozytorium dziedzinowe – Lectorium (Otwarte Repozytorium Nauk Historycznych)<sup>16</sup>. Materiały prezentowane są w sześciu głównych działach tematycznych: historia, archeologia, historia sztuki, bibliologia i informacja naukowa, etnologia i antropologia kulturowa oraz muzykologia. Lectorium jest wspólnym projektem Wydziału Historycznego Uniwersytetu Warszawskiego oraz ICM UW.

W Polsce działa też Agregator Centrum Otwartej Nauki<sup>17</sup>. Jest to wspólny punkt dostępu do zasobów polskich otwartych repozytoriów. Zapewnia ich większą widoczność, szerszą dostępność i łatwiejsze wyszukiwanie. Pozwala na przeszukiwanie polskich repozytoriów wg tytułu, daty publikacji, języka i typu. Niestety, nie wszystkie archiwa są w nim zarejestrowane. Aktualnie (sierpień 2016 r.) agregator rejestruje 19 repozytoriów i ok. 65 tys. pozycji. Agregator nie rejestruje zbiorów kilkunastu polskich zasobów tego rodzaju. Dobrym źródłem informacji o polskich repozytoriach otwartych jest też baza wiedzy serwisu Uwolnij Naukę zawierająca obszerną listę (31 pozycji) takich projektów.

#### 4. Podsumowanie

Obecnie otwarte repozytoria w Polsce zgromadziły ok. 27 tys. różnorodnych zasobów, przede wszystkim prac dyplomowych i doktorskich, artykułów z czasopism oraz materiałów dydaktycznych. Omawiane projekty rozwijają się bardzo aktywnie, zważywszy na ich krótką historię. Tworzone są przez uczelnie publiczne i niepubliczne, stanowią efekt wspólnych inicjatyw różnych jednostek naukowych bądź są projektami pojedynczych instytutów naukowo-badawczych. Ich liczba stale rośnie. Większość z nich wykorzystuje oprogramowanie Open Source, szczególnie Dspace i EPrints. Część z nich funkcjonuje na specjalnie utworzonych własnych platformach informatycznych. Często wykorzystywane jest oprogramowanie dLibra.

Do zasobów repozytoriów można przekazywać materiały, do których prawa należą do naukowców lub ich rodzimych instytutów (np.: sprawozdania, raporty, analizy, wyniki badań, publikacje wydawnictw uczelnianych). W przypadku opublikowania pracy u komercyjnego wydawcy należy upewnić się czy posiada się prawo do zamieszczenia jej w repozytorium.

<sup>15</sup> Otwórz książkę, <http://otworzksiazke.pl>

<sup>16</sup> Lectorium, <http://lectorium.edu.pl/pl/>

<sup>17</sup> Agregator CEON, <http://agregator.ceon.pl/>

Pomocą może służyć portal SHERPA, w którym znajdują się informacje o warunkach umieszczania w repozytoriach materiałów, opublikowanych wcześniej u danego wydawcy. Zasady przestrzegania prawa autorskiego powinny być opisane i udostępnione na platformie repozytorium (Rychlik & Karwasińska, b.d.).

Tworzenie repozytoriów przynosi wiele korzyści dla uczelni. Należą do nich m.in. wzrost jej prestiżu jako skutek zwiększenia widoczności w sieci wyników badań naukowych prowadzonych na wydziałach, promowanie kadry naukowej, wspieranie edukacji studentów poprzez ułatwienie im dostępu do materiałów dydaktycznych. Autorzy deponujący prace w otwartych repozytoriach mogą również zaobserwować różne korzyści z tego faktu, np. wzrost cytowalności prac umieszczonych w otwartym dostępie, zwiększenie „widoczności” w sieci dorobku naukowego – prace umieszczane w repozytoriach są indeksowane przez światowe wyszukiwarki np. Google Scholar, powszechność i szybkość dostępu do badań oraz możliwość utworzenia cyfrowego CV – prezentacja własnego dorobku z dostępem do pełnych tekstów. Coraz powszechniej obiektem zainteresowania są tzw. dane surowe (ang. *raw data*), czyli dane badawcze przetworzone w niewielkim stopniu, np. dokumenty finansowe, filmy, fotografie, nagrania audio, notatki, pamiętniki, rysunki, statystyki demograficzne i ekonomiczne, itp. Tak rozumiane dane przeciwstawione są opracowaniom, czyli tradycyjnym publikacjom naukowym, fachowym, publicystycznym, jak artykuły, eseje, książki, podręczniki, raporty, referaty. Powstaje coraz więcej archiwów surowych danych, co pozwala na wielokrotne wykorzystanie tych samych danych w wielu badaniach. Jest to jednak temat na odrębne opracowanie.

## Bibliografia

- Bednarek-Michalska, B. (2011). *Rola bibliotek naukowych we wdrażaniu rozwiązań otwartych. Repozytorium Open Access – model dla uczelni* [online]. [19.05.2016], [http://suw.biblos.pk.edu.pl/resources/i5/i7/i8/i0/r5780/BednarekMichalskaB\\_RolaBibliotek.pdf](http://suw.biblos.pk.edu.pl/resources/i5/i7/i8/i0/r5780/BednarekMichalskaB_RolaBibliotek.pdf)
- Bednarek-Michalska, B. (2012) Repozytoria surowych danych – dlaczego biblioteki powinny je znać? [online]. *Biuletyn EBIB*, 135, [20.11.2015], [http://www.nowyebib.info/images/stories/numery/135/135\\_michalska\\_.pdf](http://www.nowyebib.info/images/stories/numery/135/135_michalska_.pdf)
- Karwasińska, E. (2012). Polskie repozytoria otwarte. [online] *Biuletyn EBIB*, 20.08.2016], <http://open.ebib.pl/ojs/index.php/ebib/article/view/157/296>
- Przyłuska J.; Radomska, A. (2010). Repozytorium tematyczne publikacji powstałych w ramach sieci doskonałości ECNIS [online]. *Forum Bibl. Med.*, 1(5), [29.08.2016], <http://cybra.lodz.pl/Content/4644>
- Przyłuska J.; Radomska, A.; Kłosiński, R. (2007). Budowanie repozytorium dziedzinowego – doświadczenia Biblioteki Naukowej Instytutu Medycyny Pracy w Łodzi i Biblioteki Politechniki Łódzkiej [online]. W: *EBIB Materiały konferencyjne nr 18*. [29.08.2016], [http://www.ebib.pl/publikacje/matkonf/mat18/przyluska\\_radomska\\_klosinski.php](http://www.ebib.pl/publikacje/matkonf/mat18/przyluska_radomska_klosinski.php)
- RCIN (2014). Informacje o Projekcie RCIN [online], [29.08.2016], <http://www.rcin.org.pl/dlibra/text?id=aboutRCIN>
- Rychlik, M.; Karwasińska, E. (b.d.). *Repozytoria otwarte – zielony kanał Open Access* [online]. [24.08.2016], <http://elearning.pwste.edu.pl/moodle/mod/book/view.php?id=93&chapterid=28>
- Rychlik M. (2011). *Co repozytorium instytucjonalne oferuje swojej społeczności akademickiej? – na przykładzie repozytorium AMUR oraz repozytoriów na świecie* [online]. [29.08.2016], <http://fbc.pionier.net.pl/details/nn2Zvh5>
- Sidor, M. (2014). *Repozytorium instytucjonalne – jego funkcje i narzędzia w komunikacji naukowej w oparciu o doświadczenia WSB – [Preprint]* [on-line]. [29.08.2016], <http://repozytorium.wsb-nlu.edu.pl/handle/11199/4384>

## Polish Institutional Repositories as a Place for Open Scientific and Educational Resources

### Abstract

**Purpose/Thesis:** Open digital repositories are rapidly developing type of information sources for science and education. The aim of this article is to discuss the results of research on the current state of digital repositories maintained by universities and other research institutions in Poland.

**Approach/Research methods:** Polish digital repositories were described in terms of their current state of development, abundance, structure, collections and functions. It was decided to analyze only the digital collections recording new resources, and not - as in the case of many digital libraries in Poland – digitized collections of national heritage. The repositories in question were selected based on the list of repositories from the following web portals: Uwolnij Naukę, EBIB and CEON database aggregator. Each website was analyzed individually. Data from OpenDOAR was used to present some statistical background for this type of projects in Poland.

**Results and conclusions:** The author found 33 currently available open digital repositories in Poland. They provide access to approximately 270,000 resources of various types, mostly theses and dissertations, journal articles and teaching materials. Open digital repositories are developing very actively, given that they have a very short history. Creating and maintaining digital repositories by the universities brings many benefits such as leveraging reputation as a result of increased visibility of the scientific research conducted at the departments, promoting the academic staff, supporting the students by facilitating their access to teaching materials.

**Originality/Value:** The results of the analysis of Polish digital repositories shows their current state of development and may encourage researchers to deposit their works there and use repository resources in their research process.

### Keywords

Open Access. Polish institutional repositories. Polish domain repositories. Digital resources.

---

*Dr hab. EWA GŁOWACKA pracuje w Instytucie Informacji Naukowej i Bibliologii UMK. Obecnie jest jego dyrektorem. Jej zainteresowania badawcze koncentrują się wokół elektronicznych zasobów informacji, zarządzania jakością i oceną jakości zasobów informacyjnych, zarządzania informacją, zarządzania wiedzą, architektury informacji.*

*Kontakt z autorką:*

*egt@umk.pl*

*Instytut Informacji Naukowej i Bibliologii, Uniwersytet Mikołaja Kopernika w Toruniu*

*ul. Władysława Bojarskiego 1*

*87–100 Toruń*

# Media społecznościowe jako źródło informacji bibliograficznej w naukach humanistycznych.

## Komunikat z badań

Piotr Tańkowski

*Katedra Informatologii, Wydział Dziennikarstwa, Informacji i Bibliologii, Uniwersytet Warszawski*

### Abstrakt

**Cel/teza:** Celem artykułu jest przedstawienie wyników pilotażowego badania dotyczącego pewnych elementów zachowań informacyjnych humanistów, pracowników naukowych polskich uczelni, w sieciach społecznościowych. Chodzi o wykorzystywanie ich jako źródła informacji bibliograficznej w zakresie nowych publikacji we własnej dyscyplinie.

**Koncepcja/metody badań:** Pilotażowe badanie sondażowe tego typu zachowań informacyjnych przeprowadzone zostało w maju 2016 r. za pomocą ankiety udostępnionej na Facebooku i skierowanej do polskich badaczy humanistyki lub nauk społecznych. Przeprowadzono także wyszukiwanie artykułów zagranicznych poświęconych tej problematyce, by przedstawić światowy stan badań tej problematyki i na tym tle zarysować sytuację w Polsce.

**Wyniki i wnioski:** Badacze zagraniczni zgodni są co do tego, że korzystanie przez naukowców z mediów społecznościowych zwiększa liczbę cytowań oraz podnosi ogólny wpływ społeczny akademików aktywnie korzystających z tej formy popularyzacji. Wyniki badania sondażowego przeprowadzonego wśród badaczy polskich pokazały, że coraz chętniej korzystają oni z mediów społecznościowych poszukując informacji oraz tworząc sieci kontaktów. Nadal jednak możliwości oferowane przez nowe media nie są w pełni wykorzystywane.

**Oryginalność/wartość poznawcza:** W Polsce nie prowadzono dotychczas badań nad tą formą nieformalnych zachowań komunikacyjnych pracowników wyższych uczelni.

### Słowa kluczowe

Nieformalne zachowania komunikacyjne. Komunikacja naukowa. Niewidzialne kolegia. Sieci kontaktów. Dzielenie się informacją i wiedzą. Media społecznościowe. Nauka 2.0.

*Otrzymano: 10 lipca 2016. Zrecenzowano: 15 lipca 2016. Zaakceptowano: 18 lipca 2016.*

## 1. Wprowadzenie

Problematyka współczesnej komunikacji naukowej, a w jej ramach dzielenia się informacją i wiedzą, była już niejednokrotnie badana i omawiana w literaturze informatologicznej (z nowszej literatury polskiej zob. np. Świgoń, 2015; Kotuła, 2013; Nahotko, 2010; Sapa, 2009; Krakowska, 2008; Cisek & Sapa, 2007. Tam też dalsze odsyłacze bibliograficzne). Współcześnie komunikacja naukowa coraz częściej wykorzystuje nowe technologie, przenosząc się do świata wirtualnego oraz nabiera cech zachowań nieformalnych (zob. Genoni et al., 2014; Nahotko, 2008, 171–175; Barjak, 2006; Genoni et al., 2005. Warto również zwrócić uwagę na książkę Baykoucheva, 2015, poświęconą różnorodnym zagadnieniom zbierania informacji

i zarządzania nimi w środowisku cyfrowym). Tematyce obiegu informacji w epoce nowych mediów poświęciło swoje dysertacje doktorskie kilkoro informatologów młodego pokolenia. Przedmiot bibliografii w zmieniającym się środowisku informacyjnym badał Jarosław Pacek (Pacek, 2010), przy czym, jak sądzę, po upływie kilku już lat od publikacji jego książki refleksji należy poddać również rolę bibliografii w środowisku sieci społecznościowych. Przedmiotem analizy stać się powinna także funkcja bibliotek i placówek informacyjnych jako szeroko pojętych brokerów informacji bibliograficznej 2.0 (Gmiterek, 2012; Wójcik, 2013). Na zagadnienie konwergencji komunikacji naukowej zwrócił uwagę m.in. Emanuel Kulczycki (2012). Sieci społeczne naukowców korzystających z jednego z najpopularniejszych serwisów społecznościowych poddał analizie Veslava Osińska i Tomasz Komendziński (Osińska & Komendziński, 2014, 269–280). Przedmiotem ich zainteresowania była jednak wizualizacja sieci kontaktów, a nie sposób dzielenia się wiedzą w tejże sieci.

W sieciach społecznościowych zachodzi dwukierunkowe (czy może raczej: wielokierunkowe) dzielenie się informacją. Każdy użytkownik może być nie tylko odbiorcą, ale też nadawcą oraz bardzo często pośrednikiem (brokerem) informacji przesyłanej różnymi kanałami, a także przenoszonej z jednego kanału informacyjnego do innego. Dzielenie się wiedzą przez pracowników naukowych w sieciach społecznościowych może przyjmować różne formy: prowadzenia blogów naukowych, dyskusji na forach, redagowania haseł w Wikipedii (czy ogólniej: aktywności w serwisach typu Wiki), zamieszczania wpisów na Twitterze czy Facebooku, udostępniania preprintów artykułów w społecznościowych serwisach naukowych itd. Rozwój nowych technologii umożliwia także włączanie się do dyskursu naukowego szeroko pojętym „amatorom” – osobom zainteresowanym daną problematyką, nieposiadającym jednak (jeszcze) stopni naukowych, stanowiska akademickiego czy ugruntowanej pozycji w środowisku naukowym (m.in. studentom i doktorantom).

Z powodu szybkiego przyrostu liczby publikacji naukowych w każdej dyscyplinie (por. np. Nahotko, 2010, 37)<sup>1</sup>, coraz większym problemem staje się nie tylko kwestia ich przeczytania, ale nawet uzyskiwania na bieżąco informacji o ukazujących się nowościach, monografiach i artykułach w czasopismach czy tomach zbiorowych. Tradycyjne sposoby pozyskiwania informacji bibliograficznej na temat nowości (bibliografie, cytowania, recenzje) stają się z różnych względów nieefektywne. Opóźnienia w publikowaniu recenzji monografii w czasopismach naukowych sięgać mogą nawet dwóch lat, zaś rozrastające się w nieskończoność tradycyjne spisy bibliograficzne mogą być niewygodne w użyciu. Nie do przecenienia pozostają bezpośrednie kontakty osobiste pomiędzy uczonymi, ustna wymiana informacji co do potrzeb oraz stanu badań w danej dyscyplinie, w tym także informacji i opinii na temat nowo ukazujących się publikacji. Z oczywistych względów jednak skala i zasięg tego typu komunikacji są znacznie ograniczone. Potrzebujemy zatem nowych narzędzi czy też sposobów docierania do tego typu informacji, narzędzi i metod nie tylko szybkich, ale też precyzyjnych. Konieczna wydaje się realizacja postulatów zarówno konstruowania nowych instrumentów, opartych na nowych technologiach, jak i personalizacji

<sup>1</sup> Jeden z respondentów przeprowadzonej przeze mnie ankiety, której wyniki omówione są w dalszej części artykułu, napisał: „Podstawowym problemem jest zachwianie proporcji między nowymi publikacjami a czasem możliwym nie tylko na ich przeczytanie, ale twórcze wykorzystanie – przemysł, zrobienie notatek, inspirację do własnych ścieżek badawczych”. O nadprodukcji publikacji naukowych, których ilość zabija jakość oraz o konieczności ograniczenia liczby publikowanych artykułów, piszą Daniel Sarewitz (2016) i Lindsay Waters (2009).

usług wyszukiwania informacji bibliograficznej (por. Machlarz 2014, 231–240). Takimi systemami filtrującymi mogą stać się w pewnym stopniu technologie zaliczane ogólnie do zjawisk określanych mianem Web 2.0 (definicje Web 2.0 zob. Gmiterek, 2012, 37–45; por. także Nahotko, 2010, 202–203. Na temat społeczności tworzących się wokół nowych serwisów zob. Jung, red., 2010). Mowa oczywiście nie o automatycznych filtrach personalizujących wyświetlane informacje, opartych na specjalnych algorytmach, takich jak te używane przez informatycznych gigantów (Facebook, Google, Amazon), choć takie również byłyby przydatne, lecz o świadomym i celowym konstruowaniu i kształtowaniu przez każdego użytkownika własnego środowiska informacyjnego. Proces ten polega na wyborze pewnych technologii, narzędzi czy kanałów, uznanych arbitralnie za najodpowiedniejsze, najwygodniejsze i w najlepszy sposób zaspokajające indywidualne potrzeby informacyjne, odrzuceniu zaś innych. W realiach społeczeństwa informacyjnego w coraz większym stopniu zmuszani jesteśmy, by przeistaczać się ze zwykłych konsumentów, użytkowników informacji, w graczy informacyjnych (Nicholas & Dobrowolski, 2001, 4–7).

Wymienione powyżej zjawiska, wraz z zyskującym na popularności ruchem Open Access, należą w znacznej mierze do nieformalnych zachowań komunikacyjnych (Pikas, 2006). Przestrzeń wirtualna sprzyja tworzeniu się niewidzialnych kolegiów (ang. *invisible colleges*), czyli swego rodzaju nieformalnych grup uczonych z różnych instytucji, zazwyczaj odległych od siebie w przestrzeni geograficznej. John Gresham pisał o transformacji niewidzialnego kolegium, którego koncepcja powstała jeszcze w XVII stuleciu, do postaci kolegium cyfrowego, czy też kolegium w cyberprzestrzeni (ang. *cyberspace college*) (Gresham, 1994, 39). Znalazło to także oddźwięk w odpowiedzi jednej z respondentek omawianej poniżej ankiety, która stwierdziła, że nowe media dają jej możliwość korzystania w pracy zawodowej z „konsultacji ze znanymi postaciami ze świata nauki”. Przy tym jednak wyłącznie od nastawienia samych akademików, członków owych kolegiów zależeć będzie, na ile, w jaki sposób i z kim zechcą oni dzielić się swoją wiedzą (będzie o tym mowa w dalszej części tekstu). Innymi słowy, dzięki technologiom sieciowym realizowany może być społecznościowy paradygmat uprawiania nauki, nazywany również niekiedy Nauką 2.0 (Cisek, 2008a; na temat koncepcji „nauki opartej na współpracy” zob. Tapscott & Williams, 2008, 219–260).

## 2. Problematyka badań

Celem artykułu jest przedstawienie wyników pilotażowego badania dotyczącego pewnych elementów zachowań informacyjnych humanistów, pracowników naukowych polskich uczelni, w sieciach społecznościowych, a mianowicie wykorzystywania ich jako źródła informacji bibliograficznej w zakresie nowych publikacji we własnej dyscyplinie. Prowadzono już badania dotyczące narzędzi i źródeł informacji bibliograficznej wykorzystywanych przez polskich pracowników naukowych (ich wyniki nie zostały jeszcze dotychczas opublikowane, choć mówiono o nich w czasie konferencji „Nauka o informacji (informacja naukowa) w okresie zmian”, zorganizowanej przez IINiSB Uniwersytetu Warszawskiego 11–12 maja 2015 r.). Problemowi temu poświęcona została także książka Anny Mierzeckiej-Szczepańskiej, która podjęła się studiów nad zachowaniami informacyjnymi pracowników Wydziału Historycznego Uniwersytetu Warszawskiego. Zgodnie z jej ustaleniami:

naukowcy zbierający informacje w pierwszej kolejności korzystają z komunikacji nieformalnej, wymiany informacji wśród kolegów specjalistów o prowadzonych pracach badawczych z danego zakresu, a dopiero później sięgają do źródeł formalnych (Mierzecka-Szczepeńska, 2013, 123–124).

Jednakże społeczniowy aspekt tego zagadnienia nie był dotychczas analizowany przez polskich specjalistów, aczkolwiek nie ulega wątpliwości, że włączenie go do badań nad nieformalnymi zachowaniami informacyjnymi pracowników nauki, jako integralnej części tego zjawiska, staje się niezbędne. Fakt ten już przed kilku laty uświadomili sobie badacze z krajów anglosaskich. Przykładowo, problem dzielenia się wiedzą ukrytą za pomocą narzędzi społeczniowych stanowił przedmiot analizy akademików australijskich: Sirousa Panahiego, Jasona Watsona oraz Helen Partridge (Panahi et al., 2013; Panahi et al., 2012). Z kolei międzynarodowa grupa specjalistów zaproponowała analizę śladów zachowań informacyjnych naukowców w sieciach społeczniowych, ich elektronicznych aktów komunikacyjnych oraz rozpatrywanie ich jako tzw. wskaźników altmetrycznych. Jak piszą:

*Termin ten można zdefiniować jako tworzenie i badanie nowych wskaźników wpływu opartych o media społeczniowe, służących analizowaniu aktywności w szkolnictwie wyższym. Wskaźniki altmetryczne ilustrują aspekt zmieniających się procesów naukowo-badawczych, dialog i wykorzystanie wyników badań. Wskaźniki altmetryczne pokazują także zmieniające się metody bibliometryczne i naukometryczne (Puckett Rodgers & Barbrow, 2014, 2)<sup>2</sup>.*

Na temat aktywności naukowców w serwisach społeczniowych pisał w *Nature* Richard Van Noorden (Van Noorden, 2014). Cytaty można by mnożyć. Oczywiście staje się zatem, że studia na tym polu podjąć muszą także polscy informatolodzy.

### 3. Wykorzystanie Web 2.0 przez humanistów – wyniki badań pilotażowych

Pilotażowe badanie tego typu zachowań informacyjnych przeprowadzono w maju 2016 r., rozsyłając ankietę składającą się z sześciu pytań (zob. załącznik). Choć prośba o jej wypełnienie w samym tylko serwisie Facebook dotarła – zgodnie z dostępnymi statystykami – co najmniej do 2500 użytkowników, ostatecznie zdecydowały się ją wypełnić jedynie 42 osoby, 20 mężczyzn i 22 kobiety. Wśród nich były dwie osoby z dyplomem studiów licencjackich, 25 – magisterskich, 10 osób z doktoratem oraz pięcioro doktorów habilitowanych. Największy odsetek respondentów stanowili reprezentanci bibliologii i informatologii (18 osób) oraz historii (12 osób). Ankietę wypełniło także troje historyków sztuki oraz pojedynczy reprezentanci pedagogiki, filologii klasycznej, archeologii, sztuk pięknych, etnologii, nauk o polityce, socjologii, kulturoznawstwa, językoznawstwa oraz literaturoznawstwa. Największą popularnością wśród nich cieszą się serwisy: Facebook (39 użytkowników), Academia.edu (23 użytkowników), Lubimyczytac.pl (14), Researchgate.net i Mendeley (po 7 wskazań), a wreszcie Biblionetka.pl oraz Twitter (po 5 wskazań). Jaki obraz nieformalnej komunikacji naukowej w mediach społeczniowych wyłania się z tego badania?

Stosunkowo najwięcej zainteresowania badaczy zajmujących się współczesną komunikacją naukową wzbudzają prowadzone przez naukowców blogi (zob. np. Kulczycki, 2012; Cisek, 2008b, 170–177; Kjellberg, 2009) – prawdopodobnie dlatego, że najłatwiej poddają

<sup>2</sup> Zob. też Priem et al., 2012; Priem et al., 2011.

się obserwacji i opisowi. Powyżej wspomniałem już o recenzjach, będących w tradycyjnym systemie komunikacji naukowej jednym ze źródeł informacji o nowych (choć może już nie najnowszych) publikacjach. W danym kontekście stwierdzić trzeba, że recenzja książki naukowej opublikowana na blogu ma tę przewagę nad recenzją w czasopiśmie, że nie tylko ma szansę ukazać się szybciej, lecz także potencjalnie posiada większy zasięg, być może przeczyta ją (a przynajmniej do niej dotrze) więcej czytelników niż do tekstu drukowanego. Ponadto będą oni mogli także wyrazić swoje opinie, podać uzupełnienia (także bibliograficzne), sprostowania czy w komentarzach zadać dodatkowe pytania autorowi recenzji.

W drugiej połowie XX w. recenzje i abstrakty były jednymi z filarów rozwoju nauki w Polsce (szczególnie w przypadku niedostępnej literatury zagranicznej). Jednakże wraz ze zmianą rzeczywistości społeczno-politycznej, stanowiącej otoczenie nauki, a także samej nauki i funkcjonującej w jej ramach komunikacji, również i rola recenzji naukowych musiała ulec zmianie. Obecnie w poszukiwaniu informacji, także tej o krajowych i zagranicznych nowościach wydawniczych, sięga się po źródła i narzędzia innego rodzaju niż jeszcze dwie dekady temu. Zresztą proces publikacji czasopism trwa na tyle długo, że recenzowana książka często zdąży zniknąć z półek księgarskich, zanim jej recenzja ukaże się drukiem.

Wszystko to sprawia, że wielu pracowników naukowo-dydaktycznych, lecz także amatorów, wierzy w sens wykorzystania narzędzi społecznościowych, w tym platform blogowych (zarówno ogólnych, jak i specjalistycznych platform dedykowanych dla blogów naukowych, na przykład ScienceBlogs czy ResearchBlogging), w celu pisania o książkach i artykułach naukowych, recenzowania i dyskutowania o nich. Teksty na blogach naukowych różnią się od tych publikowanych na blogach literackich m.in. tym, że nie powstają w celach marketingowych. Już samo czytanie dobrze napisanej, rzetelnej recenzji dać może sporo wiadomości, rzucać nowe światło na znane problemy, pozwala spojrzeć na nie z innej perspektywy, podsunąć nowe interpretacje, zachęcać do własnych przemyśleń. Nauka rozwija się w dialogu, nie w monologu, jakim jest opublikowanie książki czy artykułu, a blogujący recenzenci włączają się w debatę i zachęcają do tego innych. Teraz, gdy na publikację odpowiedzi autorowi recenzji nie trzeba czekać do następnego numeru czasopisma (ukazującego się co kwartał lub rzadziej) dialog ten toczyć się może o wiele łatwiej i skuteczniej. Także wielu z respondentów uważa blogi za istotne źródło informacji o nowych publikacjach, jak na przykład „blogi specjalistyczne prowadzone przez informatologów i bibliotekarzy”.

Drugim ważnym medium społecznościowym, często omawianym w kontekście zastosowań w komunikacji naukowej, są mikroblogi. Analizie poddaje się możliwości wykorzystania ich (w szczególności dotyczy to popularnego Twittera) w celu promowania publikacji naukowych. W Polsce Twitter wciąż nie cieszy się zbyt wielkim uznaniem w środowiskach akademickich (jak wspomniano, zaledwie pięć wskazań w ankiecie), inaczej natomiast jest w Ameryce Północnej, gdzie pracownicy wyższych uczelni korzystają z niego dość intensywnie. Szczególnie wiele interesujących wiadomości na ten temat zawiera artykuł Gunthera Eysenbacha, który zaproponował nowatorskie ujęcie naukometryczne z wykorzystaniem nowego źródła danych, czyli właśnie Twittera. Serwisu tego użyć można do pomiaru oddziaływania społecznego publikacji, uzupełniając tradycyjną analizę cytowań. Badania kanadyjskiego uczonego dowodzą, że tweetowanie na temat artykułów naukowych może wielokrotnie zwiększyć ich zasięg społeczny oraz liczbę cytowań. Podkreśla on także, że ocena naukometryczna wpisów zamieszczanych w mediach społecznościowych (ang. *social media impact metrics*) jest problemem ważnym z kilku powodów. Po pierwsze, może

być dość łatwo stosowana przez samych akademików, instytucje i redakcje czasopism celem kontrolowania ogólnego wpływu społecznego konkretnych badań w wybranych przedziałach czasowych. Po drugie, dzięki tej metodzie można poddawać ewaluacji różne metody upowszechniania wiedzy. Po trzecie wreszcie, może ona znaleźć zastosowanie jako narzędzie, dzięki któremu będzie można kierować czytelników do tych artykułów, które w największym stopniu przyciągają uwagę publiczności czy też społeczności uczonych (Eysenbach, 2011). Przydatność praktycznego zastosowania tego serwisu mikroblogowego dla udostępniania informacji o nowych publikacjach, powiększania grona odbiorców tekstów oraz jako narzędzia naukometrycznego, pozwalającego śledzić *impact factor* własnych publikacji, podkreślają także autorzy poradnika *Using Twitter in University Research, Teaching and Impact Activities* (Mollett et al., 2011). Co ciekawe, badacze, którzy zadali sobie pytanie, czy wskaźniki altmetryczne rzeczywiście działają, doszli do wniosków negatywnych. Nie dostrzegli oni znaczących korelacji pomiędzy wpisami w serwisach społecznościowych a wzrostem liczby cytowań. Mimo to wskazali na wyjątkową, pozytywną pozycję Twittera w tym kontekście (Thelwall et al., 2013)<sup>3</sup>.

Bez wątpienia ze wszystkich społecznościowych kanałów komunikacyjnych to właśnie Twitter doczekał się najgłębszej refleksji naukometrycznej, komunikologicznej i informatologicznej. Jak już wspomniano, u nas nie cieszy się on takim powodzeniem, jak za Oceanem i dotychczas niewielu polskich naukowców z niego korzysta. W warunkach polskich serwis ten wykorzystywany jest głównie przez polityków i dziennikarzy, natomiast przeciętnych użytkowników posiada stosunkowo niewielu.

Doskonałym miejscem udostępniania informacji o nowych publikacjach wydaje się być serwis Facebook. Podkreślało to wielu z respondentów ankiety. Komunikacja taka odbywać się może na różne sposoby: w grupach, zarówno o tematyce ogólnej (np. „Publikacje humanisty”), jak i specjalistycznych, fachowych czy zawodowych (np. „Biblio”, „Mediewiści polscy”, „Imago Turci”; czasem użytkownicy publikują w grupach nie tylko informacje o nowych książkach, polskich i zagranicznych, lecz także spisy treści wydawnictw zbiorowych czy najnowszych numerów czasopism), poprzez profile wydawców (bardzo wiele wydawnictw naukowych posiada już swoje profile w serwisach społecznościowych i mniej lub bardziej systematycznie udostępnia na nich informacje na temat nowości wydawniczych) czy też „brokerów informacji bibliograficznej” (np. „Classica, mediaevalia et cetera”, „New Books in History”, „Libros arte medieval”). Informacje o nowych książkach publikują czasem także na swoich profilach sami autorzy, jak również jednostki naukowe, w których są zatrudnieni (katedry, instytuty czy wydziały wyższych uczelni), a wreszcie różnego rodzaju instytucje (np. archiwa, muzea, biblioteki naukowe). Dzięki polubieniom i udostępnieniom informacja taka potencjalnie docierać może do tysięcy użytkowników.

Zauważyć trzeba, że zamknięte grupy specjalistyczne, takie jak wspomniane powyżej, niekiedy wykazują tendencję do wykluczania amatorów, osób spoza ścisłego środowiska zawodowego (a także spoza własnych kręgów towarzyskich), jednakże zainteresowanych problematyką. Dążą do elitarności, a tym samym nie wykorzystują szansy na szersze upowszechnienie wyników badań swojej dyscypliny czy popularyzację nauki w ogóle. Zjawisko to ocenić należy zdecydowanie negatywnie, ponieważ popularyzacja jest jednym

<sup>3</sup> Korpus literatury anglojęzycznej poświęconej naukowemu tweetowaniu jest już pokaźny. W tym miejscu warto przywołać jeszcze jeden tylko artykuł: Holmberg & Thelwall, 2013.

z obowiązków pracowników nauki, szczególnie tych korzystających z mediów społecznościowych. O konieczności wykorzystywania tychże przez uczonych w celu popularyzacji nauki pisała w 2011 r. na swoim blogu w witrynie „Scientific American” Christine Wilcox (Wilcox, 2011). Na problem niechęci naukowców do popularyzacji nauki, wynikającej z ich obaw przed utratą naukowej powagi („saganizacji”<sup>4</sup>), zwróciła uwagę m.in. grupa amerykańskich medioznawców i komunikologów (Liang et al., 2014, 772–775). Równocześnie autorzy ci wskazali, że rozwój technologii informatycznych daje uczonym nowe możliwości docierania do różnych kręgów opinii publicznej i aktywnego kształtowania form komunikacji oraz sformułowali hipotezę, zgodnie z którą istnieje pozytywna korelacja pomiędzy komunikowaniem się badaczy z publicznością spoza świata nauki a wpływem naukowym (ang. *scientific impact*) akademików posiadających tę umiejętność. Uczeni, którzy angażują się w działania popularyzujące naukę, publikują więcej artykułów w czasopiśmie recenzowanych i bywają częściej cytowani<sup>5</sup>.

Facebook cieszy się największą popularnością wśród respondentów ankiety (korzysta z niego niemal 93% badanych), choć w pewnej mierze wynik ten może być spowodowany faktem, że prośbę o udział w badaniu rozsyłałem za pomocą tego serwisu, licząc na efekt wirtualnej kuli śnieżnej (co już samo w sobie jest dość symptomatyczne; wspomniałem powyżej, iż media społecznościowe coraz częściej wykorzystywane są do prowadzenia badań naukowych). Jak napisał jeden z respondentów, profesor nadzwyczajny Uniwersytetu Adama Mickiewicza w Poznaniu, „Korzystam z fb. Mam polubione strony kilku wydawnictw i na bieżąco pojawiają mi się informacji o nowościach, szczególnie z interesujących mnie dziedzin. Obserwuję również profile znajomych naukowców, którzy często podają informacje o nowościach wydawniczych”. Inny badany chwali, że Facebook umożliwia mu „zbieranie wszelkich możliwych nowinek historycznych”. Kolejna respondentka wyznaje: „gdyby nie różne grupy, nie korzystałabym z tego [serwisu] – nie mam czasu na ‘zabawy’ w czytanie nieinteresujących rzeczy. Jedynie sprawy związane z nauką, nowymi publikacjami, badaniami, eventami kulturalnymi etc. To trzeba przyznać – jest [to] dobre źródło informacji”.

Dobrym sposobem społecznościowego udostępniania informacji bibliograficznej są fora tematyczne. Dają one użytkownikom możliwość wymieniać się informacjami na temat nowości, ale także uzupełniania własnych spisów o pozycje dawniejsze, a dotychczas nieznane, polecane przez innych forumowiczów. Interesujący wydaje się przykład publikującego książki historyczne oświęcimskiego wydawnictwa Napoleon V, którego właściciel jest aktywnym użytkownikiem największego forum historycznego w Polsce – historycy.org. Za jego pośrednictwem komunikuje się ze swoimi czytelnikami: zbiera informacje na temat ich oczekiwań, pozycji, których brakuje na polskim rynku, rekomendacji zagranicznej literatury etc. Ponadto na subforum wydawniczym otwiera w nowych wątkach dyskusje nad ostatnio opublikowanymi książkami. Użytkownicy zatem nie tylko bardzo szybko dowiadują się o nowościach wydawniczych, lecz także mają możliwość podyskutowania o nich, wymienienia się opiniami, jak też kształtowania polityki wydawniczej. Dla wydawcy

<sup>4</sup> Termin ukuto od nazwiska amerykańskiego astronoma i popularyzatora nauki, Carla Edwarda Sagana, który z powodu swojego zaangażowania w działalność popularyzatorską stracił autorytet naukowy w oczach swoich kolegów.

<sup>5</sup> Konieczność wykorzystywania mediów 2.0 podkreślana jest także przez wielu innych autorów. Przegląd odnośnej literatury dostępny jest online pod adresem <https://socialnetworkingforscientists.wikispaces.com/General> (dostęp: 27.05.2016).

zaś jest to nie tylko forma promocji, lecz także platforma, na której może nawiązać kontakt z autorami czy tłumaczami, budując swoje sieci kontaktów.

Zarządzaniu własnymi zestawieniami bibliograficznymi, ale także dzieleniu się tego typu informacją z innymi użytkownikami, służą narzędzia zwane menedżerami bibliografii, takie jak Mendeley, Zotero, EndNote, Refworks czy CiteULike, te jednak nie są w Polsce zbyt popularne (Wiorogórska, 2015). Spośród badanych humanistów jedynie siedmiu wskazało, iż korzysta z Mendeley, nikt natomiast nie używa pozostałych. Trzeba na koniec wspomnieć także o społecznościowych repozytoriach tekstów naukowych, takich jak Academia.edu („jest genialna”) czy Researchgate.net (odpowiednio 23 i 7 użytkowników w badaniu ankietowym). Dają one możliwość nie tylko publikowania własnych tekstów naukowych, lecz także prowadzenia na ich temat dyskusji. Ich użytkownicy mogą tworzyć własne sieci kontaktów, obserwować wybrane osoby, a dzięki temu mogą na bieżąco dowiadywać się o ich nowych publikacjach.

Istotnym źródłem informacji stają się newslettery księgarń naukowych i specjalistycznych, których bazy danych

zwłaszcza za granicą, stają się ważnym darmowym źródłem informacji o produkcji wydawniczej (Jachimczyk, 2008, 245).

Z powiadomień różnego rodzaju korzysta znaczna część ankietowanych.

Uzupełniającym sposobem zdobywania informacji bibliograficznej mogą być popularne społecznościowe serwisy czytelnicze, np. Lubimyczytac.pl. Dokonywane tam przypadkowe odkrycia oraz rekomendacje innych użytkowników, obserwowanych i znajomych, mogą stanowić cenne uzupełnienie własnych zestawień bibliograficznych. Problematyka niezamierzonego pozyskiwania informacji o literaturze doczekała się już także refleksji w piśmiennictwie informatologicznym, a co więcej, postuluje się konstruowanie naukowych cyfrowych systemów informacyjnych nie tylko umożliwiających, ale wręcz

inspirujących przypadkowe odkrywanie informacji i relacji między nimi (Sapa, 2008, 81. Tam też literatura przedmiotu).

Obecnie uzasadniona jest aktualizacja tej obserwacji poprzez włączenie w jej zakres także bibliofilskich serwisów społecznościowych.

## 4. Zakończenie

Podsumowując powyższe rozważania, bazując nie tylko na wynikach pilotażowej ankiety, lecz także na obserwacjach zachowań informacyjnych moich koleżanek i kolegów w serwisach społecznościowych, których sam jestem użytkownikiem, stwierdzić trzeba, iż polscy humaniści coraz aktywniej wykorzystują owe media do poszukiwania informacji, w tym także informacji bibliograficznej (choć pięcioro respondentów udzieliło odpowiedzi, że nie korzysta z nich w tym celu w ogóle<sup>6</sup>). Warto podkreślić raz jeszcze, że udostępnianie za ich pomocą informacji o nowych publikacjach powoduje wzrost liczby cytowań oraz

<sup>6</sup> Jedna z badanych osób napisała: „Szczepnie mówię, nie przyszło mi to jeszcze do głowy. Ale bardzo dziękuję za sugestię!”.

ogólnie podnosi poziom wpływu społecznego i naukowego (*vulgo*: popularność) obecnego w sieci uczonego. Nowe nowe media służą polskim humanistom zasadniczo do wymiany informacji o monografiach, rzadziej zaś powiadamia się za ich pośrednictwem o artykułach, choć niekiedy udostępniania się skany spisów treści nowych numerów czasopism lub też materiałów konferencyjnych, a także preprinty artykułów za pośrednictwem takich serwisów, jak Academia.edu.

Za swego rodzaju *post scriptum* do moich rozważań uznać można wyniki ankiety „W sieci o książkach”, dotyczącej sposobów lektury Polaków, opublikowane 24 maja 2016 r. przez wspomniany serwis *Lubimy Czytać*<sup>7</sup>. W ankiecie wzięło udział bez mała 24 tys. respondentów – użytkowników serwisu i członków sieciowej społeczności czytelniczej. Sondaż ten wskazuje, jak ważnym medium stał się Internet dla miłośników literatury. Poszukując oni za jego pośrednictwem informacji o książkach, uczestniczą w dyskusjach, korzystają z mediów „dwuzerowych”, także tych poświęconych szczególnie tematyce literackiej i to one stanowią dla nich ważne źródło informacji na temat produkcji wydawniczej. Jako takie wymieniano tutaj m.in. *Lubimy Czytać*, strony księgarni internetowych, blogi i vlogi książkowe, Facebook, Twitter, Instagram itd.

Nie ulega zatem wątpliwości, że mamy do czynienia z istotnym, nabierającym – w miarę upowszechniania się technologii „dwuzerowych” w różnych dziedzinach życia – coraz większej wagi zjawiskiem społecznym, wartym pogłębionych badań i szerszego opisu. Jest to bardzo skomplikowany i pojemny treściowo fenomen, z wieloma mniej lub bardziej bezpośrednimi i oczywistymi implikacjami, w którego omówieniu trzeba odwoływać się do różnych zagadnień, umieszczając je w szerokim kontekście (przypomnijmy dla uporządkowania: nieformalne zachowania komunikacyjne w komunikacji naukowej, tworzenie niewidzialnych kolegiów oraz sieci kontaktów, dzielenie się informacją i wiedzą, rozwój nowych nowych mediów, społecznościowy paradygmat uprawiania nauki). Problem do tychczas pozostaje w Polsce raczej niezauważony, natomiast wiele uwagi poświęcają mu naukowcy z zachodniego kręgu kulturowego.

W gruncie rzeczy nie jest istotne, jakich technologii i mediów będziemy używali, by zaspokajać własne potrzeby informacyjne. Ważne jest, by wybrać sobie odpowiednie dla indywidualnych upodobań, potrzeb oraz realizacji stających przed nami zadań narzędzia i korzystać z nich rozbudowując sieci kontaktów, a przede wszystkim by świadomie zarządzać swoim środowiskiem informacyjnym, stając się ekspertem w „grze o informację”.

## Bibliografia

- Barjak, F. (2006). The Role of the Internet in Informal Scholarly Communication. *Journal of the American Society for Information Science and Technology*, 57 (10), 1350–1367.
- Baykoucheva, S. (2015). *Managing Scientific Information and Research Data*. Amsterdam: Elsevier.
- Cisek, S. (2008a). Nauka 2.0. Nowe narzędzia komunikacji naukowej [online]. e-LiS, [11.05.2016], <http://eprints.rclis.org/11212/>
- Cisek, S. (2008b). Weblogi (blogi) – nowe narzędzie komunikacji w nauce. W: D. Pietruch-Reizes (red.) *Zarządzanie informacją w nauce*. Katowice: Wydawnictwo Uniwersytetu Śląskiego, 170–177.
- Cisek, S.; Sapa, R. (2007). Komunikacja naukowa w Internecie – mity i rzeczywistość. W: W. Lubaszewski (red) *Komputer – Człowiek – Prawo*. Kraków: Wydaw. UJ, 39–49.

<sup>7</sup> <http://lubimyczytac.pl/aktualnosci/7122/w-sieci-o-ksiazkach---wyniki-ankiety>.

- Eysenbach, G. (2011). Can Tweets Predict Citations? Metrics of Social Impact Based on Twitter and Correlation with Traditional Metrics of Scientific Impact. *Journal of Medical Internet Research*, 13 (4), [27.04.2016], <http://www.jmir.org/2011/4/e123/>
- Genoni, P.; Merrick, H.; Willson, M. (2014). Virtual Symposia: An Investigation into Scholarly Communities Online [online]. Academia.edu, [18.05.2016], [https://www.researchgate.net/publication/240636706\\_Virtual\\_symposia\\_an\\_investigation\\_into\\_scholarly\\_communities\\_online](https://www.researchgate.net/publication/240636706_Virtual_symposia_an_investigation_into_scholarly_communities_online)
- Genoni, P.; Merrick, H.; Willson, M. (2005). Community, Communication, Collaboration: Scholarly Practice in Transformation. In: *Proceedings of the Next Wave of Collaboration: Educause Australasia 2005, April 5–8, 2005* [online]. Auckland, NZ: Curtin University, [18.05.2016], [http://espace.library.curtin.edu.au/R?func=dbin-jump-full&local\\_base=gen01-era02&object\\_id=19204](http://espace.library.curtin.edu.au/R?func=dbin-jump-full&local_base=gen01-era02&object_id=19204)
- Gmiterek, G. (2012). *Biblioteka w środowisku społecznościowego Internetu: biblioteka 2.0*. Warszawa: Wydaw. SBP.
- Gresham, J. (1994). From Invisible College to Cyberspace College: Computer Conferencing and the Transformation of Informal Scholarly Communication Networks. *Interpersonal Computing and Technology Journal*, 2 (4), 37–52.
- Holmberg, K.; Thelwall, M. (2013). Disciplinary Differences in Twitter Scholarly Communication. In: J. Gorraiz; E. Schiebel; C. Gumpenberger; M. Hörlesberger; H. Moed (eds.) *Proceedings of the 14th International Society of Scientometrics and Informetrics Conference*, Vienna, Austria. Wien: Austrian Institute of Technology, 567–582.
- Jachimczyk, A. (2008). Bazy internetowe księgarń jako źródło informacji. W: D. Pietruch-Reizes, W. Babik (red.) *Wymiana informacji i rozwój profesjonalnych usług informacyjnych w edukacji, nauce i kulturze na rzecz społeczeństwa opartego na wiedzy*. Katowice: PTIN, 245–253.
- Jung, B. red. (2010). *Wokół mediów ery Web 2.0*. Warszawa: Wydaw. Akademickie i Profesjonalne.
- Kjellberg, S. (2009). Blogs as Interface Between Several Worlds. A Case Study of the Swedish Academic Blogosphere. *Human IT*, 10 (3), 1–45.
- Kotuła, S.D. (2013). *Komunikacja bibliologiczna wobec World Wide Web*. Lublin: Wydaw. UMCS.
- Krakowska, M. (2008). *Nowe formy komunikacji społecznej w europejskiej przestrzeni edukacyjnej*. Kraków: Wydaw. UJ.
- Kulczycki, E. (2012). Blogi i serwisy naukowe. Komunikacja naukowa w kulturze konwergencji. W: I. Sójkowska (red.) *Materiały konferencyjne EBIB*, 22. Toruń: Stowarzyszenie EBIB, [16.05.2016], <https://repozytorium.amu.edu.pl/handle/10593/2520>
- Liang, X.; Su, L.Y.-F.; Yeo, S.K.; Scheufele, A.A.; Brossard, D.; Xenos, M.; Nealey, P.; Corley, E.A. (2014). Building Buzz: (Scientists) Communicating Science in New Media Environments. *Journalism & Mass Communication Quarterly*, 91 (4), 772–791.
- Machlarz, A. (2014). Nadinformatywność systemów informacyjnych. W: E. Głowacka, M. Kowalska, P. Krysiński (red.) *Współczesne oblicza komunikacji i informacji. Problemy, badania, hipotezy*. Toruń: Wydaw. Nauk. UMK, 231–242.
- Mierzecka-Szczepeńska, A. (2013). *Badania zachowań informacyjnych*. Warszawa: Wydaw. SBP.
- Mollett, A.; Moran, D.; Dunleavy P. (2011). Using Twitter in University Research, Teaching and Impact Activities [online]. LSE Research Online, [11.05.2016], [http://eprints.lse.ac.uk/38489/1/Using\\_Twitter\\_in\\_university\\_research%2C\\_teaching\\_and\\_impact\\_activities\\_\(LSE\\_RO\).pdf](http://eprints.lse.ac.uk/38489/1/Using_Twitter_in_university_research%2C_teaching_and_impact_activities_(LSE_RO).pdf)
- Nahotko, M. (2008) Czy nastąpi koniec druku? Komunikacja naukowa XXI wieku oparta na publikacjach elektronicznych. W: D. Pietruch-Reizes, W. Babik (red.) *Wymiana informacji i rozwój profesjonalnych usług informacyjnych w edukacji, nauce i kulturze na rzecz społeczeństwa opartego na wiedzy*. Katowice: PTIN, 78–91.
- Nahotko, M. (2010). *Komunikacja naukowa w środowisku cyfrowym: globalna biblioteka cyfrowa w informatycznej infrastrukturze nauki*. Warszawa: Wydaw. SBP.
- Nicholas, D.; Dobrowolski, Z. (2001). Informacyjny gracz: nowa koncepcja użytkownika informacji. *Praktyka i Teoria Informacji Naukowej* 1–2, 4–9.

- Osińska, V.; Komendziński, T. (2014). Naukowcy na Facebooku. Wizualizacja sieci społecznych w nauce. W: E. Głowacka, M. Kowalska, P. Krysiński (red.) *Współczesne oblicza komunikacji i informacji. Problemy, badania, hipotezy*. Toruń: Wydaw. Nauk. UMK, 269–282.
- Pacek, J. (2010). *Bibliografia w zmieniającym się środowisku informacyjnym*. Warszawa: Wydaw. SBP.
- Panahi, S.; Watson, J.; Partridge, H. (2013). Towards Tacit Knowledge Sharing over Social Web Tools. *Journal of Knowledge Management* 17 (3), 379 – 397.
- Panahi, S.; Watson, J.; Partridge, H. (2012). Social Media and Tacit Knowledge Sharing: Developing a Conceptual Model, *World Academy of Science, Engineering and Technology* 64, 1095–1102.
- Pikas, Ch.K. (2006). The Impact of Information and Communication Technologies on Informal Scholarly Scientific Communication: A Literature Review [online]. University of Maryland, [11.05.2016], [http://terpconnect.umd.edu/~cpikas/878/Pikas\\_The\\_Impact\\_of ICTs\\_on\\_ISSC\\_0506.pdf](http://terpconnect.umd.edu/~cpikas/878/Pikas_The_Impact_of ICTs_on_ISSC_0506.pdf)
- Priem, J.; Taraborelli, D.; Groth, P.; Neylon, C. (2012). Altmetrics: A Manifesto [online]. Altmetrics, [18.05.2016], <http://altmetrics.org/manifesto>
- Priem, J.; Piwowar, H.A.; Hemminger, B.M. (2011). Altmetrics in the Wild: Using Social Media to Explore Scholarly Impact [online]. *Altmetrics*, [18.05.2016], <http://altmetrics.org/altmetrics12/priem/>
- Puckett Rodgers, E.; Barbrow, S. (2014). Wskaźniki altmetryczne i ich rosnące znaczenie w bibliotekach naukowych [online]. *Biuletyn EBIB*, 6 (151), [18.05.2016], <http://open.ebib.pl/ojs/index.php/ebib/article/view/249>
- Sapa, R. (2009). *Metodologia badań obszaru pośredniczenia w komunikacji naukowej z perspektywy nauki o informacji*. Kraków: Wydaw. UJ.
- Sapa, R. (2008). Przypadkowe pozyskiwanie informacji o literaturze naukowej. Inne spojrzenie na jakość systemów informacyjnych dla naukowców. W: D. Pietruch-Reizes, W. Babik (red.) *Wymiana informacji i rozwój profesjonalnych usług informacyjnych w edukacji, nauce i kulturze na rzecz społeczeństwa opartego na wiedzy*. Katowice: PTIN, 77–85.
- Sarewitz, D. (2016). The Pressure to Publish Pushes Down Quality [online]. *Nature*, 533 (12 May 2016), [18.05.2016], <http://www.nature.com/news/the-pressure-to-publish-pushes-down-quality-1.19887?>
- Świogoń, M. (2015). *Dzielenie się wiedzą i informacją : specyfika nieformalnej komunikacji w polskim środowisku akademickim*. Olsztyn: Wydaw. UWM.
- Tapscott, D.; Williams, A.D. (2008). *Wikinomia: o globalnej współpracy, która zmienia wszystko*. Warszawa: Wydaw. Akademickie i Profesjonalne.
- Thelwall, M.; Haustein, S.; Larivière, V.; Sugimoto, C.R. (2013). Do Altmetrics Work? Twitter and Ten Other Social Web Services [online]. *PLOS*, [18.05.2016], <http://dx.doi.org/10.1371/journal.pone.0064841>
- Van Norden, R. (2014). Online Collaboration: Scientists and the Social Network [online]. *Nature*, 512 (14 August 2014), [18.05.2016], <http://www.nature.com/news/online-collaboration-scientists-and-the-social-network-1.15711>
- Waters, L. (2009). *Zmierzch wiedzy. Przemiany uniwersytetu a rynek publikacji naukowych*. Kraków: Homini.
- Wilcox, C. (2011). Social Media for Scientists. Part 1: It's Our Job [online]. *Scientific American* [11.05.2016], <http://blogs.scientificamerican.com/science-sushi/social-media-for-scientists-part-1-it-8217-s-our-job/>
- Wiorogórska, Z. (2015). Field-Specific Information Needs of Doctoral Students in Psychology. A Comparative Study. In: S. Kurbanoglu, J. Boustany, S. Spiranec, E. S. Grassian, D. Mizrachi, & L. Roy (eds.) *Information Literacy: Moving Toward Sustainability*. Third European Conference, ECIL 2015, Tallinn, Estonia, October 19–22, 2015, Revised Selected Papers. Springer International Publishing Switzerland, 375–382.
- Wójcik, M. (2013). *Web 2.0 w działalności usługowej instytucji książki*. Kraków: Wydaw. UJ.

## Social Media as a Source of Bibliographic Information in the Humanities. Research Report

### Abstract

**Purpose/Thesis:** The purpose of this paper is to present the results of a pilot study on selected elements of information behavior of Polish humanists (academic staff of Polish universities) observed in social networks. The central idea is to analyze the use of the social networks as a source of bibliographic information on new publications in a given field of science. .

**Approach/Research methods:** The pilot study of this type of information behavior was conducted in May 2016 with a questionnaire made available on Facebook and addressed to Polish scholars in the humanities and social sciences. Moreover, the research was made to find foreign articles devoted to this issue in order to reflect a global state of research on this subject, using it as a background to outline the state of respective research in Poland.

**Results and conclusions:** Foreign researchers are unanimous in recognizing the use of the social media as the factor increasing the number of citations and improving the overall social impact of scholars actively applying this form of knowledge popularization. Polish scientists are increasingly eager to use social media while seeking information and creating a network of contacts, although the possibilities offered by the new media still are not used to their full extent.

**Originality/Value:** There has not been any research on this form of informal communication behavior of university employees conducted in Poland previous to the study described in this paper.

### Keywords

Informal information behavior. Scholarly communication. Invisible colleges. Networks of contacts. Knowledge and information sharing. Social media. Science 2.0.

---

*Dr hab. PIOTR TAŃKOWSKI, adiunkt w Katedrze Informatologii na Wydziale Dziennikarstwa, Informacji i Bibliologii Uniwersytetu Warszawskiego. W pracy naukowej i dydaktycznej zajmuje się bibliologią historyczną, dziejami komunikacji społecznej oraz zagadnieniami informatologii. Najważniejsze nowe publikacje: Rejestracja rękopiśmiennych i drukowanych poloników w zbiorach zagranicznych – problem wciąż otwarty, Roczniki Biblioteczne 58 (2014) [druk 2015], 87–100; Anti-Turkish Literature in 15<sup>th</sup>-16<sup>th</sup> Century Europe, Tarih İncelemeleri Dergisi, XXX/1 (2015), 231–280; Antyturska korespondencja Macieja Korwina i Sykstusa IV : przyczynek do dziejów propagandy w stosunkach międzynarodowych w późnym średniowieczu, Zeszyty Naukowe Uniwersytetu Jagiellońskiego. Prace Historyczne, 143 (2016), 1, 37–55.*

*Kontakt z autorem:*

*p.tanowski@gmail.com*

*Katedra Informatologii*

*Wydział Dziennikarstwa, Informacji i Bibliologii*

*Uniwersytet Warszawski*

*ul. Nowy Świat 69*

*00-927 Warszawa*

# Przekształcenie relacji występujących w hasłach rozwiniętych i opisach przedmiotowych KABA do postaci powiązań w mapie tematów

Bartłomiej Włodarczyk

*Katedra Bibliografii i Dokumentacji,*

*Wydział Dziennikarstwa, Informacji i Bibliologii, Uniwersytet Warszawski*

## Abstrakt

**Cel/teza:** Celem artykułu jest wskazanie możliwości i problemów transformacji wybranych haseł rozwiniętych KABA do postaci powiązań w mapie tematów opartej na JHP KABA.

**Koncepcja/metody badań:** Na podstawie analizy dokumentacji opublikowanej na stronie Centrum NUKAT przeprowadzono ocenę możliwości wprowadzenia bardziej szczegółowej kategoryzacji jednostek leksykalnych języka KABA. Następnie analizując noty stosowania określników rzeczowych, wskazano na rodzaje konstrukcji występujących we wspomnianym języku oraz możliwości ich przekształcenia w powiązania w mapie tematów.

**Wyniki i wnioski:** W języku KABA, oprócz podziału tematów i określników zgodnego z etykietami pól formatu MARC 21 dokonano kategoryzacji określników w celu kontroli połączeń między jednostkami leksykalnymi w hasłach rozwiniętych. Dane na ten temat są zapisane w rekordach wzorcowych określników (pole 073). Brakuje ich jednak w rekordach wzorcowych tematów (pole 072), co uniemożliwia automatyczną kontrolę poprawności haseł. Pokazano, że ich uzupełnienie tylko w ograniczonym zakresie może być pomocne przy przekształcaniu języka do postaci mapy tematów. W przypadku niektórych jednostek identyfikujących argumenty relacji konieczne było zaproponowanie dodatkowych kategorii, które pozwoliły na zapis ograniczeń dotyczących powiązań. W artykule zaprezentowano przykłady przekształcenia wybranych haseł i opisów przedmiotowych. Kryterium doboru było występowanie w ich składzie określników rzeczowych reprezentujących relacje między przedmiotami. Podział całości zasobu leksykalnego KABA byłby dużo bardziej złożony i wymagałby utworzenia licznych ograniczeń. Zalecane byłoby przynajmniej częściowe zautomatyzowanie tej pracy.

**Oryginalność/wartość poznawcza:** W artykule zaprezentowano propozycję przekształcenia relacji występujących w hasłach rozwiniętych i w opisach przedmiotowych KABA w mapę tematów. Przedstawiona propozycja z jednej strony wypełnia w pewnym stopniu lukę, jaka istnieje w tym zakresie w literaturze z zakresu informatologii. Z drugiej strony jest ona głosem w dyskusji na temat przyszłości języków haseł przedmiotowych, w szczególności JHP KABA, pokazując rozwiązanie wpisujące się w nurt współczesnych projektów semantycznych. W związku z kompatybilnością języka KABA z językami RAMEAU i LCSH, rezultaty badań mogą być częściowo zastosowane odnośnie do tych narzędzi.

## Słowa kluczowe

Hasła rozwinięte. Język haseł przedmiotowych KABA. Mapy tematów. Relacje.

*Otrzymany: 17 czerwca 2016. Zrecenzowany: 6 lipca 2016. Zatwierdzony: 1 sierpnia 2016.*

## 1. Wprowadzenie

W katalogach polskich bibliotek obecnie są wykorzystywane głównie dwa języki haseł przedmiotowych (JHP): Język Haseł Przedmiotowych KABA (Katalogów Automatycznych Bibliotek Akademickich, JHP KABA) oraz Język Haseł Przedmiotowych Biblioteki Narodowej (JHP BN). Pierwszy z nich jest stosowany w sieci bibliotek naukowych i akademickich współpracujących w ramach Narodowego Uniwersalnego Katalogu Centralnego (NUKAT), drugi zaś w Bibliotece Narodowej (BN) oraz w większości bibliotek publicznych.

W BN trwają prace nad przekształceniem JHP BN w Deskryptory Biblioteki Narodowej (DBN). Zgodnie z założeniami projektu, JHP o prekoordynowanych jednostkach leksykalnych, ma zostać przekształcony w język deskryptorowy, w którym „gramatykę stanowią najczęściej reguły indeksowania współrzednego” (*Język deskryptorowy*, 2002). Wśród przyczyn podjęcia decyzji o przekształceniu rozwijanego od 1957 r. języka wymieniono m.in. jego niedostosowanie do wymagań współczesnych systemów informacyjno-wyszukiwawczych, w tym wyszukiwania fasetowego (Cichoń et al., 2014, 174–177). W związku z przemianami zachowań informacyjnych użytkowników oraz rozwojem systemów informacyjno-wyszukiwawczych, również twórcy drugiego polskiego JHP stoją obecnie przed koniecznością podjęcia decyzji dotyczącej jego przyszłości. Jednym z rozwiązań jest całkowita fasetyzacja struktury języka, a więc przyjęcie modelu transformacji podobnego do rozwiązania opracowywanego obecnie w BN. Ostateczna decyzja powinna uwzględniać potrzeby informacyjne użytkowników, które w projekcie przekształcenia powinny być traktowane priorytetowo, a także możliwości bibliotek współpracujących w ramach NUKAT. W propozycji przekształcenia należy wziąć pod uwagę specyficzne cechy języka KABA.

Celem artykułu jest wskazanie odmiennej możliwości transformacji JHP KABA, tj. transformacji haseł rozwiniętych KABA w sieć powiązań w mapie tematów (ang. *topic map*). Propozycję przekształcenia JHP BN zgodnie z tym wzorcem przedstawiono w osobnej publikacji (Włodarczyk, 2016). Mapy tematów są standardem reprezentacji wiedzy, którego model zakłada istnienie „tematów” (ang. *topics*), reprezentujących różne pojęcia i obiekty (tzw. przedmioty), „powiązań” (ang. *associations*), czyli relacji łączących tematy, oraz „wystąpień” (ang. *occurrences*), będących relacjami między tematami a relewantnymi zasobami informacyjnymi. Element „zakres” (ang. *scope*) reprezentuje kontekst, w którym ważne są m.in. powiązania. W artykule standard map tematów (ISO, 2006) traktowany jest jako podstawa przekształcenia haseł rozwiniętych KABA. Hasło rozwinięte można potraktować jako zbiór jednostek połączonych różnymi rodzajami relacji. W JHP jednymi z najpowszechniej wykorzystywanych związków na poziomie syntagmatycznym są relacje łączące przedmiot dokumentu z jego lokalizacją przestrzenną i czasową. Te dwa podstawowe rodzaje relacji nie wyczerpują jednak katalogu związków występujących w hasłach rozwiniętych. Na pozostałe wskazują niektóre określniki rzeczowe oraz formy. Ponadto, relacje między różnymi przedmiotami są zawarte w ramach tematów ogólnych. Dotyczy to niektórych tematów składających się z elementów połączonych spójnikiem „i”, np. „Religia i państwo”. W artykule poddano analizie jedynie określniki rzeczowe KABA wskazujące na związki między przedmiotami. Obiektem badań były zarówno rekordy wzorcowe tematów, określników i haseł rozwiniętych KABA, jak i rekordy bibliograficzne dostępne w tym katalogu. Przedstawione w artykule wyniki mogą posłużyć jako jeden z elementów przekształcenia języka KABA w system organizacji wiedzy (SOW) uwzględniający formalną specyfikację relacji między pojęciami.

Zawartość artykułu została podzielona na trzy części. W pierwszej omówiono kategoryzację jednostek leksykalnych KABA. Zwrócono uwagę na możliwość formalizacji połączeń poszczególnych kategorii tematów z określnikami rzeczowymi. W drugiej części przedstawiono możliwości specyfikacji relacji zawartych w wybranych hasłach rozwiniętych KABA. Nie uwzględniono tematów, które identyfikują związki między przedmiotami. Zaprezentowano przykłady transformacji haseł zawierających wybrane określniki, wskazujące na relacje między przedmiotami, do postaci powiązań w mapie tematów. W trzeciej części przedstawiono sposób kontroli typów ról i tematów w powiązaniach. W zakończeniu zaprezentowano podsumowanie problematyki przedstawionej w artykule.

## 2. Kategoryzacja jednostek leksykalnych KABA

Rekordy wzorcowe języka KABA są zapisywane w formacie MACHine-Readable Cataloging 21 (MARC 21). Format ten zawiera odrębne pola służące do zapisu różnych rodzajów tematów – osobowego, korporatywnego dla tzw. ciał zbiorowych, korporatywnego dla tzw. imprez, tytułowego, ogólnego, geograficznego oraz formalnego. W formacie MARC 21 uwzględniono również podpola służące do zapisu określników, drugiej obok tematów kategorii leksykalnej JHP. Wyodrębniono podpola dla określników: rzeczowego, geograficznego, chronologicznego oraz formy. Podkategorie wyróżnione w ramach dwóch podstawowych kategorii leksykalnych JHP służą organizacji oraz kontroli słownictwa JHP. Są one jednak na tyle ogólne, że możliwości ich wykorzystania w procesie transformacji JHP do postaci SOW opartego na sieci jawnie specyfikowanych relacji są ograniczone.

W KABA, podobnie jak w Répertoire d'autorité matière encyclopédique et alphabétique unifié (RAMEAU) (BnF, 2016), ustanowiono dziedzinowe listy określników. Jednak nie wszystkie kategorie (listy) KABA występują we francuskim JHP. W NUKAT wyodrębniono osobną listę określników stosowanych po nazwach obozów i więzień (Ćwikowski, 2013, 80). Na stronie internetowej NUKAT napisano, że listy są „narzędziem ułatwiającym stosowanie określników w hasłach przedmiotowych rozwiniętych” (NUKAT-a, b.d.). Obecnie udostępniono następujące listy określników:

- |                              |                              |
|------------------------------|------------------------------|
| (1) Lista podstawowa (ORZF)  | (15) Muzyka                  |
| (2) Choroby                  | (16) Nazwy geograficzne      |
| (3) Ciała niebieskie         | (17) Nazwy osobowe           |
| (4) Ciała zbiorowe (CZB)     | (18) Obozy i więzienia (OBW) |
| (5) Czasopisma               | (19) Organizmy żywe          |
| (6) Części ciała             | (20) Parlamenty              |
| (7) Dyscypliny sportowe      | (21) Pojazdy                 |
| (8) Działy gospodarki (GOSP) | (22) Prawo                   |
| (9) Grupy etniczne           | (23) Produkty i materiały    |
| (10) Instrumenty muzyczne    | (24) Religie                 |
| (11) Języki                  | (25) Siły zbrojne            |
| (12) Kategorie osób          | (26) Substancje chemiczne    |
| (13) Księgi święte           | (27) Tytuły dzieł            |
| (14) Literatury              | (28) Wojny                   |

Dla każdej listy określono rodzaje tematów, które można stosować w połączeniu z zawartymi na niej określnikami. Na liście podstawowej wymieniono określniki stosowane po różnych rodzajach tematów. Na pozostałych określono dokładnie, po jakich rodzajach tematów mogą być stosowane. Na przykład, na liście „Działy gospodarki” zapisano, że występujące na niej określniki stosuje się po „nazwach działów gospodarki, np.: Przemysł budowlany” oraz po „typach działalności gospodarczej i handlowej, np.: Centra handlowe”. W opisie tej listy zaznaczono również, że określników z tego wykazu nie „stosuje się po hasle Przemysł” (NUKAT-c, b.d.). Podobne uwagi znajdują się przy wszystkich listach określników zamieszczonych na stronie internetowej NUKAT. Informacje te, zapisane w rekordach wzorcowych, mogłyby posłużyć do automatycznej kontroli poprawności haseł rozwiniętych. Potencjalnie, ich zapisanie w rekordach wzorcowych mogłoby być także podstawą przekształcenia KABA, co zostanie przedstawione w dalszej części artykułu.

Do formalizacji zasad łączliwości określników z tematami mogą posłużyć pola 072 i 073 formatu MARC 21. Przykłady ich wykorzystania w Medical Subject Headings (MeSH) zaprezentowano na stronie Biblioteki Kongresu (BK) poświęconej wspomnianemu formatowi. W pierwszym z wymienionych pól można zamieścić kod kategorii, do której przypisano dany temat lub deskryptor, w drugim natomiast informacje na temat kategorii tematów lub deskryptorów, po których można stosować dany określnik lub modyfikator. Poniżej zamieszczono przykład użycia obu pól wraz z objaśnieniem zamieszczonym w dokumentacji formatu MARC 21 (LoC, 1999).

```
150  ##  $a Mental Health Services
072  ##  $a F4. $x 408
072  ##  $a N2. $x 421. $x 461

180  ##  $x utilization
073  ##  $a E1 $a E2 $a E3 $a E4 $a E6 $a E7 $a N2 $z mesh
```

W rekordzie modyfikatora „utilization” (wykorzystanie) w powtarzalnym podpolu \$a pola 073 zamieszczono oznaczenia siedmiu kategorii deskryptorów, po których może on być stosowany. W podpolu \$z wskazano źródło słownictwa, czyli MeSH. W rekordzie deskryptora „Mental Health Services” (Opieka psychiatryczna) w podpolu \$a pola 072 zapisano kategorie, do których należy ta jednostka leksykalna. Dokładne umiejscowienie deskryptora w hierarchii zostało przedstawione w podpolach \$x. Z zapisu wynika, że deskryptor ten należy do kategorii F4 (Dyscypliny i działania behawioralne) oraz N2 (Zakłady opieki zdrowotnej, służba zdrowia i usługi zdrowotne). Modyfikator „utilization” można stosować po drugiej z wymienionych kategorii deskryptorów, a zatem jego połączenie z deskryptorem „Mental Health Services” jest zgodne z metodyką tworzenia haseł MeSH (LoC, 1999).

W Library of Congress Subject Headings (LCSH) wyróżniono 34 kategorie tematów. Kody dla tych kategorii do niedawna były zapisywane wyłącznie w polu 073 rekordów wzorcowych określników. W BK jest obecnie prowadzony projekt polegający na dodaniu odpowiednich kodów do wybranych rekordów wzorcowych tematów (pole 072). Autorzy tego przedsięwzięcia oczekują, że umożliwi to (LoC, 2012):

- (1) zwiększenie użyteczności LCSH w środowisku Sieci Semantycznej,
- (2) dostarczenie osobom katalogującym podpowiedzi poprzez udostępnienie listy określników, które można zastosować po danym temacie,
- (3) automatyczne tworzenie i walidację haseł rozwiniętych LCSH.

Poniżej zamieszczono przykład zastosowania pól 072 i 073 w rekordach LCSH (LoC, b.d.; LoC, 2016, 10).

|        |                                                                                                    |
|--------|----------------------------------------------------------------------------------------------------|
| 072 #7 | <i>\$a H 1195 \$2 lcsb</i>                                                                         |
| 150 ## | <i>\$a Mazda 6 automobile</i>                                                                      |
| 073 ## | <i>\$a H 1195 \$2 lcsb</i>                                                                         |
| 180 ## | <i>\$x Bumpers</i>                                                                                 |
| 680 ## | <i>\$i Use as a topical subdivision under individual land vehicles and types of land vehicles.</i> |

Obecnie sytuacja w języku KABA jest podobna do tej w LCSH sprzed rozpoczęcia wspomnianego projektu. Informacje o łączności poszczególnych okreśników KABA z tematami są wyłącznie częścią rekordów wzorcowych okreśników tego języka (pole 073). W rekordach wzorcowych tematów nie są uwzględnione informacje na temat ich przynależności kategorialnej (pole 072). Tym samym nie jest możliwa automatyczna kontrola poprawności haseł rozwiniętych. Jak podkreślono powyżej, najszerszy zakres zastosowania mają okreśniki zamieszczone na liście podstawowej. Część z nich można używać po wszystkich rodzajach tematów, np. okreśnik „analiza i indeksowanie dokumentów”, który „stosuje się po nazwach pospolitych i własnych do prac dotyczących analizy dokumentów z dziedziny określonej przez temat”. W ich przypadku kontrola wydaje się zbędna ze względu na uniwersalne zastosowanie. Inne okreśniki z tej listy mają jednak o wiele węższy zakres stosowania. Przykładowo, okreśnik „aparatura i sprzęt” „stosuje się po nazwach instytucji i organizacji, gałęzi gospodarki, obozów i więzień oraz po właściwych nazwach przedmiotów (dziedzin, procesów technologicznych, itp.)”. Informacje te są uwzględnione w rekordzie wzorcowym okreśnika:

|        |                                          |
|--------|------------------------------------------|
| 073 ## | <i>\$a CZB \$a GOSP \$a OBW \$a ORZF</i> |
| 180 ## | <i>\$x aparatura i sprzęt.</i>           |

W celu zapewnienia pełnej kontroli poprawności haseł należałoby doprecyzować granice kategorii opisanej ogólnie jako „właściwe nazwy pospolite”. Wspomniano o braku informacji zapisywanych w polu 072 rekordów wzorcowych tematów KABA. W pełni automatyczne, szybkie dodanie tych danych jest obecnie możliwe dla tematów osobowych poprzez wykorzystanie etykiety 100 formatu MARC 21. Istnieje osobna lista okreśników stosowanych po nazwach osób, rodów, rodzin i dynastii. W przypadku innych kategorii tematów podpowiedź mogą stanowić ich etykiety oraz znajomość ich budowy, np. wszystkie tematy identyfikujące księgi biblijne rozpoczynają się od słowa „Biblia”, a jednostki identyfikujące języki od słów „Język” bądź „Języki”. Potencjalnie znaczenie mogą mieć również relacje hierarchiczne między tematami. Jednak w tym przypadku część tematów może zostać zaliczona do błędnej kategorii, np. przyjęcie, że wszystkie tematy KABA węższe znaczeniowo w stosunku do hasła „Biblioteki – personel” przynależą do klasy „Kategorie osób” spowoduje zaliczenie do niej m.in. tematu „Duszpasterstwo bibliotekarzy”. Uwzględnienia wymagają wyjątki zapisane na listach, np. okreśniki z listy „Ciała zbiorowe” nie są stosowane po nazwach różnego rodzaju obozów i więzień, dla których utworzono osobną listę (NUKAT-b, b.d.).

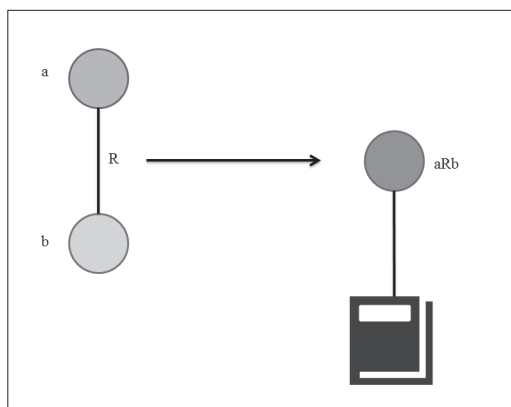
Istotnym zagadnieniem jest możliwość wykorzystania połączeń między okreśnikami a tematami w procesie przekształcania języka KABA w mapę tematów. Po uzupełnieniu rekordów wzorcowych tematów KABA o pola 072, uzyskane połączenia między tematami

a okreśnikami miałyby postać par jednakowych oznaczeń, np. „nazwy osobowe” – „nazwy osobowe”. Informacje te nie są wystarczające do zapisu ograniczeń w mapie tematów, nie wskazują bowiem kategorii, do których należą pojęcia reprezentowane przez okreśniki bądź ich elementy odpowiadające argumentom relacji. Niezbędne jest wyodrębnienie kategorii, do których przynależą.

### 3. Okreśniki rzeczowe KABA jako podstawa powiązań w mapie tematów – analiza wybranych przykładów

W języku KABA, podobnie jak w JHP BN, istnieje możliwość wyszczególnienia relacji między jednostkami leksykalnymi wchodzącymi w skład jednego hasła bądź też całego opisu przedmiotowego. W mapie tematów pojedyncze tematy oraz tematy połączone powiązaniami mogą posłużyć do opisu treści lub formy dokumentów. W celu reprezentacji powiązań w postaci odrębnych tematów należałoby zastosować reifikację. W mapach tematów tylko tematy mogą być połączone powiązaniami lub wystąpieniami z opracowywanymi dokumentami. Poniżej zamieszczono rysunek (Rys. 1) prezentujący zastosowanie reifikacji w stosunku do powiązania. Powstały temat może zostać połączony z dokumentem bądź jego reprezentacją.

Rysunek 1. Zastosowanie reifikacji w celu stworzenia tematu służącego do opisu treści dokumentu.



Źródło: opracowanie własne z wykorzystaniem „Book” by ANS (<http://aiconica.net/books>). CC0 1.0 Universal.

W tej części artykułu zostaną omówione możliwości przekształcania wybranych haseł i całych opisów przedmiotowych, zawierających wybrane okreśniki KABA do postaci powiązań w mapie tematów. Wybrane jako przykłady okreśniki rzeczowe identyfikują relacje między przedmiotami. Uwzględnione zostaną wyłącznie hasła i opisy, w których wszystkie argumenty relacji są zapisane w sposób jawny. W przykładach pominięto lokalizację przestrzenną i czasową przedmiotów.

Pierwszy okreśnik – „stosunki wojskowe”, stosuje się w celu opisanego polityki wojskowej jednego kraju wobec drugiego lub wobec grupy państw. Użycie okreśnika geograficznego po okreśniku „stosunki wojskowe” wskazuje na kraj lub region odniesienia. Zgodnie

z pragmatyką JHP KABA należy zawsze utworzyć drugie hasło zaczynające się od nazwy występującej w określniku geograficznym. Relacja ta jest związkiem symetrycznym. Schemat prezentujący sposób zastosowania określnika w opisie przedmiotowym oraz przykład opisu zawierającego omawiany określnik zaprezentowano poniżej. Oznaczenie „a” identyfikuje pierwszy człon relacji, oznaczenie „R” określnik wskazujący na rodzaj relacji, a oznaczenie „b” jednostkę identyfikującą drugi człon relacji.

*aRb 651 ## \$a Finlandia \$x stosunki wojskowe \$z Polska*  
*bRa 651 ## \$a Polska \$x stosunki wojskowe \$z Finlandia*

W oparciu o hasła zawierające ten określnik można utworzyć powiązanie typu „państwo-pozostające-w-stosunkach-wojskowych/państwo-pozostające-w-stosunkach-wojskowych”, w którym powinny wystąpić dwa tematy pełniące rolę typu „państwo-pozostające-w-stosunkach-wojskowych”. Poniżej zaprezentowano zapis ograniczenia dotyczącego tego powiązania w języku ograniczeń map tematów – Topic Maps Constraint Language (TMCL) (ISO, 2011). Zgodnie z tym zapisem „państwo-pozostające-w-stosunkach-wojskowych/państwo-pozostające-w-stosunkach-wojskowych” jest typem powiązania zawierającym dokładnie dwie role typu „państwo-pozostające-w-stosunkach-wojskowych”.

*państwo-pozostające-w-stosunkach-wojskowych/państwo-pozostające-w-stosunkach-wojskowych isa*  
*tmcl:association-type;*  
*has-role(państwo-pozostające-w-stosunkach-wojskowych, 2, 2).*

Poniżej przedstawiono przykład konkretnego powiązania tego typu.

*państwo-pozostające-w-stosunkach-wojskowych/państwo-pozostające-w-stosunkach-wojskowych(pa*  
*ństwo-pozostające-w-stosunkach-wojskowych:finlandia,*  
*państwo-pozostające-w-stosunkach-wojskowych:polska).*

Należy podkreślić, że w języku KABA występują również hasła rozwinięte zawierające określnik „stosunki wojskowe”, ale bez następującego po nim określnika geograficznego. Hasła takie wskazują na sytuację, w której omawiane są ogólne stosunki wojskowe z różnymi krajami.

*651 ## \$a Polska \$x stosunki wojskowe*

W przedstawionym wyżej przykładzie można zazwyczaj wskazać drugi człon relacji, jednak z różną precyzją. Potencjalnie dokument na temat stosunków wojskowych z wieloma państwami można opracować, dodając jako drugi argument relacji temat „Świat”. Obecnie w KABA nazwa ta nie występuje w funkcji określnika. Przypadki te wymagają jeszcze dalszej analizy. W artykule wskazano na hasła lub opisy przedmiotowe, w których wszystkie argumenty relacji są zapisane w sposób jawny.

Nieco inaczej należy postąpić z określnikiem „gramatyka porównawcza”, który może być stosowany po nazwach języków i grup języków samodzielnie bądź z dodatkiem drugiego określnika rzeczowego, wskazującego na „język będący obiektem badań komparatystycznych”. Określnik ten identyfikuje relację porównania języków z punktu widzenia ich gramatyki. Podobnie jak w przypadku wyżej omówionego określnika, należy zawsze utworzyć dwa hasła. Schemat budowy opisu jest więc identyczny jak w przypadku określnika „stosunki wojskowe”. Przykład zastosowania określnika „gramatyka porównawcza” podano poniżej.

```
650  ##  $a Język litewski $x gramatyka porównawcza $x język polski
650  ##  $a Język polski $x gramatyka porównawcza $x język litewski
```

Utworzone powiązanie reprezentuje relację porównywania dwóch przedmiotów i powinno zawierać dwie role typu „przedmiot porównywany”. Dodatkowo można dodać zakres wskazujący problematykę, z punktu widzenia której są rozpatrywane języki, w tym przypadku będzie to gramatyka. Nie jest to jednak konieczne, ponieważ zakres można również ustanowić między zreifikowanym tematem a wystąpieniem lub tematem reprezentującym zasób informacyjny.

```
przedmiot-porównywany/przedmiot-porównywany isa tmcl:association-type;
has-role(przedmiot-porównywany, 2, 2).
```

```
przedmiot-porównywany/przedmiot-porównywany(przedmiot-porównywany:
język-polski, przedmiot-porównywany: język-litewski).
```

Kolejnym przykładem jest określnik „dzieła”, który jest zawsze uzupełniany drugim określnikiem rzeczowym identyfikującym konkretną dziedzinę twórczości. Reprezentuje on relację między osobą a dziedziną działalności. Nie jest to związek symetryczny. Pierwszy człon relacji reprezentuje temat identyfikujący konkretną osobę, a drugi – określnik identyfikujący dziedzinę twórczości. W rekordzie wzorcowym określnika „dzieła” zaznaczono, że musi to być dziedzina niebędąca główną sferą działalności danej osoby. Schemat opisu zawiera jedno obowiązkowe hasło. Został zaprezentowany poniżej wraz z przykładem zastosowania.

```
aRb 600 1# $a Wyczółkowski, Leon $d (1852–1936) $x dzieła $x grafika
```

Relację identyfikowaną przez hasło rozwinięte z określnikiem „dzieła” można przedstawić jako powiązanie typu „osoba/dziedzina-działalności”. Ograniczenie dotyczące tego typu powiązania wraz z przykładem wykorzystania przedstawiono poniżej za pomocą konwencji TMCL.

```
osoba/dziedzina-działalności isa tmcl:association-type;
has-role(osoba, 1, 1);
has-role(dziedzina-działalności, 1, 1).
```

```
osoba/dziedzina-działalności(osoba: wyczółkowski, dziedzina-działalności: grafika).
```

Wśród wariantów nazw określników o budowie „dzieła – [dziedzina]” znajdują się nazwy o konstrukcji „wkład do [dziedziny]”. Warianty te są równokształtne z określnikami występującymi w KABA. Różnica w budowie polega na tym, że określnik ten identyfikuje zarówno rodzaj relacji, jak i jej drugi człon. W związku z tym w znajdującym się poniżej schemacie drugi argument relacji został zapisany w nawiasie okrągłym. Pod schematem zamieszczono również przykład zastosowania określnika i powiązanie powstałe w wyniku jego przekształcenia.

```
aR(b) 600 1# $a Pol, Wincenty $d (1807–1872) $x wkład do archeologii
```

```
osoba/dziedzina-działalności(osoba: pol, dziedzina-działalności: archeologia).
```

Kolejny przykład dotyczy całej grupy określników o budowie „wpływ [przymiotnik od nazwy cywilizacji, religii, obszaru geograficznego, grupy etnicznej]”. Służą one do wskazania

wpływu wymienionych w nawiasie kwadratowym kategorii przedmiotów na grupy etniczne oraz przedmioty z zakresu sztuki, kultury, literatury i cywilizacji. Są to relacje asymetryczne, czyli zachodzą tylko w jednym kierunku. W części przypadków oba człony relacji znajdują się w obrębie jednego hasła rozwiniętego.

*aR(b) 650 ## \$a Literatura niemiecka \$x wpływ amerykański.*

Powyższe hasło można odczytać jako wpływ Stanów Zjednoczonych na literaturę niemiecką. W takich przypadkach drugi człon relacji często nie jest precyzyjnie prezentowany w opisie. W podanym przykładzie może bowiem chodzić o wpływ kultury amerykańskiej na literaturę niemiecką. W opisie brakuje jednak hasła identyfikującego kulturę amerykańską. Niekiedy argumenty takiej relacji mogą występować w opisie jako osobne tematy uzupełnione określnikami. Poniżej zaprezentowano schemat opisu i przykład jego wykorzystania zaczerpnięty z katalogu NUKAT. Oznaczenie „x” identyfikuje elementy nie wchodzące w skład relacji.

*ax 650 ## \$a Literatura amerykańska \$x wpływ.*

*bR(x) 650 ## \$a Literatura angielska \$x wpływ amerykański.*

Do reprezentacji tego rodzaju relacji w mapie tematów może służyć powiązanie składające się z dwóch typów ról: „przedmiot-wpływający” i „przedmiot-na-który-wywierany-jest-wpływ”:

*przedmiot-wpływający/przedmiot-na-który-wywierany-jest-wpływ isa tmcl:association-type;*

*has-role(przedmiot-wpływający, 1, 1);*

*has-role(przedmiot-na-który-wywierany-jest-wpływ, 1, 1).*

*przedmiot-wpływający/przedmiot-na-który-wywierany-jest-wpływ(przedmiot-wpływający: literatura-amerykańska, przedmiot-na-który-wywierany-jest-wpływ: literatura-angielska).*

Wspomniany już schemat „aR(b)” występuje też w wielu innych hasłach, m.in. zawierających określniki o konstrukcji „skutki oddziaływania [przedmiotu]”. Poniżej zaprezentowano przykład hasła z jednym z określników należących do tej grupy.

*650 ## \$a Śliwa domowa \$x skutki oddziaływania wapnia.*

Hasło o takiej budowie również można przekształcić zgodnie z przedstawionym powyżej wzorem.

Inny schemat występuje w przypadku opisów zawierających określnik „deportacje z [nazwa kraju lub regionu większego niż kraj]”. Zaprezentowano go poniżej wraz z przykładem użycia.

*xa 650 ## \$a Deportacja \$z Związek Radziecki \$y 1900–1945.*

*xR(b) 650 ## \$a Wojna światowa (1939–1945) \$x deportacje z Litwy.*

W mapie tematów powyższą relację można przedstawić zgodnie z zaprezentowanym schematem:

*państwo-źródłowe-deportacji/państwo-docelowe-deportacji isa tmcl:association-type;*

*has-role(państwo-źródłowe-deportacji, 1, 1);*

*has-role(państwo-docelowe-deportacji, 1, 1).*

*państwo-źródłowe-deportacji/państwo-docelowe-deportacji(państwo-źródłowe-deportacji: litwa, państwo-docelowe-deportacji: związek-radziecki).*

Schemat podobny do przedstawionego powyżej obowiązuje w opisach z określnikiem „więźniowie i więzienia [nazwa narodowości]”. Różnica dotyczy drugiego elementu relacji, który nie jest reprezentowany przez określnik, ale przez temat.

*xR(a) 650 ## \$a Wojna światowa (1939–1945) \$x więźniowie i więzienia Niemców.  
bx 650 ## \$a Więźniowie polityczni polscy \$z Litwa \$z Wilno \$y 1900-1945.*

Relacja ponownie zachodzi między dwoma przedmiotami, które odgrywają odmienne role.

*więziony/osoby-które-więźą isa tmcl:association-type;  
has-role(więziony, 1, 1);  
has-role(osoby-które-więźą, 1, 1).*

*więziony/osoby-które-więźą(więziony: więźniowie-polityczni-polscy, osoby-które-więźą: niemcy).*

Na relację między przedmiotami wskazują także określniki o konstrukcji „w [medium]”. Identyfikują one relację przedstawienia przedmiotu m.in. w filmie, teatrze oraz literaturze. Schemat opisu z zastosowaniem tych określników obejmuje jedno lub dwa hasła. Poniżej został podany pierwszy z nich wraz z przykładem zastosowania.

*aR(b) 650 ## \$a Sacrum \$x w kinie*

W drugim schemacie występują dwa hasła.

*ax 650 ## \$a Kino polskie \$x tematy, motywy.  
bR(x) 650 ## \$a Siły zbrojne \$x w kinie.*

Ten opis przedmiotowy można przedstawić z wykorzystaniem typu powiązania „przedmiot/medium”.

*przedmiot/medium isa tmcl:association-type;  
has-role(przedmiot, 1, 1);  
has-role(medium, 1, 1).*

*przedmiot/medium(przedmiot: siły-zbrojne, medium: kino-polskie).*

Sytuację komplikuje nieco występowanie określników rzeczowych takich jak np. „kino i wojna”. Zgodnie z notą zapisaną w rekordzie wzorcowym, określnik ten „stosuje się po nazwach wojen do prac omawiających sytuację kina w okresie danej wojny, jak również do prac na temat tej wojny w produkcji filmowej współczesnej danej wojnie lub późniejszej”. Jest to jeden z określników, które identyfikują jednocześnie kilka rodzajów związków między przedmiotami. Nie można wyodrębnić w sposób automatyczny szczegółowych relacji między przedmiotami z haseł zawierających te określniki. Poniżej zaprezentowano schemat i przykład opisu dokumentu na temat przedstawienia II wojny światowej w filmie amerykańskim.

*ax 650 ## \$a Kino amerykańskie \$x tematy, motywy.  
bR 650 ## \$a Wojna światowa (1939–1945) \$x kino i wojna.*

Zaprezentowany powyżej opis można przedstawić z wykorzystaniem powiązania „przedmiot/medium”.

Kolejną jednostką identyfikującą jednocześnie kilka rodzajów związków między przedmiotami jest określnik „i [...]”, występujący w dwóch wariantach: „i [przedmiot]” oraz „i [jednostka geograficzna]”. Pierwszy z wariantów jest stosowany „po nazwach osobowych, nazwach rodów i rodzin oraz po nazwach odpowiednich instytucji i organizacji (partii politycznych, związków zawodowych, itp.), w odniesieniu do prac na temat ich postaw, opinii lub znajomości danego przedmiotu określonego wyłącznie nazwą pospolitą”, drugi zaś „do prac na temat ich postawy, opinii i znajomości danego obszaru (miasta, kraju lub regionu większego niż kraj)”. Próba wyodrębnienia bardziej szczegółowych związków między przedmiotami jest w tym przypadku utrudniona, ponieważ znaczenie określnika „i [...]” jest odmienne w różnych opisach przedmiotowych. W rekordach wzorcowych języka KABA zapisywane są odpowiedniki jednostek leksykalnych z RAMEAU oraz LCSH. Określnik KABA „i [...]” jest skonstruowany analogicznie jak jego odpowiednik w RAMEAU. W LCSH występuje pięć odpowiedników tego określnika wskazujących na trzy rodzaje relacji:

- (1) „Knowledge \$x [specific topic]” (znajomość danego przedmiotu),
- (2) „Relations with [specific class of persons or ethnic group]” (relacje z określoną kategorią osób lub grupą etniczną),
- (3) „Views on [specific topic]” (opinia o danym przedmiocie),
- (4) „Knowledge \$x [America, etc]” (znajomość danego obszaru) oraz
- (5) „Views on \$x [America, etc]” (opinia o danym obszarze).

Rozwiązanie przyjęte w BK ułatwia ewentualne przekształcenie LCSH w mapę tematów. Poniżej zamieszczono schemat zawierający określnik „i [...]” oraz przykład fragmentu opisu przedmiotowego zgodnego z językiem KABA. Podano również jego odpowiednik w LCSH.

*Modernist poetics of history: Pound, Eliot, and the sense of the past / James Longenbach. Princeton, cop. 1987.*

Hasło przedmiotowe KABA:

*aR(b) 600 1# \$a Eliot, Thomas Stearns \$d (1888–1965) \$x i historia.*

Hasło przedmiotowe LCSH:

*600 10 \$a Eliot, T. S. \$q (Thomas Stearns), \$d 1888–1965 \$x Knowledge \$x History.*

Jedynie opis sporządzony z użyciem LCSH umożliwia automatyczne utworzenie dokładnego powiązania oddającego treść dokumentu. W przypadku opisu w KABA należy utworzyć bardziej ogólny typ powiązania lub przeanalizować opisy przedmiotowe zawierające określnik „i [...]”:

*osoba/dziedzina-poznania isa tmcl:association-type;  
has-role(osoba, 1, 1);  
has-role(dziedzina-poznania, 1, 1).*

*osoba/dziedzina-poznania(osoba: eliot, dziedzina-poznania: historia).*

W KABA występują również tematy o konstrukcji „[przedmiot] i [przedmiot]”, wskazujące relacje między przedmiotami. Ich przekształcenie byłoby podobne jak w przypadku

hasel rozwiniętych z określnikiem „i [...]”. Tematy takie mogą identyfikować różne rodzaje relacji między przedmiotami.

Innym określnikiem wskazującym różne relacje jest określnik „aspekt społeczny”. Zastosowany po temacie „Wojna” lub temacie nazywającym konkretny konflikt zbrojny, może identyfikować zarówno sytuację społeczną w czasie wojny, jak i wpływ wojny na społeczeństwo lub społeczeństwa na wojnę (Ćwikowski, 2013, 61). Brak jawnego wyodrębnienia tych znaczeń ponownie utrudnia przekształcenie.

Tabela 1 prezentuje wykaz określników KABA identyfikujących relacje między przedmiotami, uporządkowanych zgodnie z wyodrębnionymi wcześniej schematami.

Tabela 1. Podział określników rzeczowych KABA wskazujących relacje między przedmiotami zgodnie ze schematami ich użycia w opisach przedmiotowych

| Schemat zastosowania określnika | Określniki rzeczowe wskazujące relacje między przedmiotami |                                                                                                                  |                                                       |                                                                                               |
|---------------------------------|------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------|-----------------------------------------------------------------------------------------------|
| 1                               | 2                                                          | 3                                                                                                                | 4                                                     | 5                                                                                             |
| <b>aRb<br/>bRa</b>              | fonologia porównawcza                                      | gramatyka porównawcza                                                                                            | stosunki                                              | stosunki gospodarcze z zagranicą                                                              |
|                                 | stosunki wojskowe                                          | stosunki zagraniczne                                                                                             |                                                       |                                                                                               |
| <b>aRb</b>                      | dzieła – [dziedzina]                                       | kariera – [dziedzina]                                                                                            | rola – [nazwa postaci]                                | studia i nauczanie – [grupa osób posługujących się określonym językiem/dziedzina]             |
| <b>aR(b)</b>                    | aspekt społeczny                                           | i [przedmiot / nazwa geograficzna]                                                                               | konsumpcja alkoholu                                   | konsumpcja energii                                                                            |
|                                 | odkrycia i badania [przymiotnik od nazwy narodowości]      | odporność na [przedmiot]                                                                                         | opinia publiczna [przymiotnik od nazwy geograficznej] | [przedmiot] i [powstanie / rewolucja / wojna]                                                 |
|                                 | relacja do [księga święta]                                 | roszczenia wobec [państwo]                                                                                       | skutki oddziaływania [przedmiotu]                     | studia i nauczanie [szczebel kształcenia]                                                     |
|                                 | transliteracja [przymiotnik od nazwy języka]               | udział [nazwa grupy etnicznej / nazwa kategorii osób / przymiotnik od nazwy geograficznej lub nazwy narodowości] | w [medium]                                            | wpływ [przymiotnik od nazwy cywilizacji / religii / obszaru geograficznego / grupy etnicznej] |

| 1                   | 2                                                                                             | 3                                                         | 4                                                                                                                | 5                       |
|---------------------|-----------------------------------------------------------------------------------------------|-----------------------------------------------------------|------------------------------------------------------------------------------------------------------------------|-------------------------|
| <b>aR(b)</b>        | wkład do [dziedziny]                                                                          | wpływ na człowieka                                        | wpływ na [nazwa dialektu / nazwa języka]                                                                         | wpływy z kosmosu        |
|                     | wykorzystanie [przymiotnik od nazwy przedmiotu]                                               | wykorzystanie dla celów [przymiotnik od nazwy przedmiotu] | zależność od [nazwa geograficzna]                                                                                | zapotrzebowanie na wodę |
|                     | zapożyczenia [przymiotnik od nazwy języka]                                                    | zastosowanie w [przedmiocie]                              | zawartość [nazwa substancji / nazwa grup substancji]                                                             |                         |
| <b>ax<br/>bR(x)</b> | wpływ [przymiotnik od nazwy cywilizacji / religii / obszaru geograficznego / grupy etnicznej] | w [medium]                                                | udział [nazwa grupy etnicznej / nazwa kategorii osób / przymiotnik od nazwy geograficznej lub nazwy narodowości] |                         |
| <b>xa<br/>xR(b)</b> | deportacje z [konkretnej jednostki geograficznej]                                             |                                                           |                                                                                                                  |                         |
| <b>xR(a)<br/>bx</b> | więźniowie i więzienia [nazwa narodowości]                                                    |                                                           |                                                                                                                  |                         |
| <b>ax<br/>bR</b>    | [przedmiot] i [powstanie / rewolucja / wojna]                                                 |                                                           |                                                                                                                  |                         |

Źródło: opracowanie własne.

#### 4. Kontrola typów ról i tematów w mapie tematów opartej na języku KABA

W tej części artykułu przedstawiono ograniczenia dotyczące instancji poszczególnych typów tematów i ról, w jakich mogą one występować w powiązaniach. Odwołano się do przykładów wcześniej zaprezentowanych w tekście. Pierwszy z przykładów zawierał dwa hasła z określnikiem „stosunki wojskowe”. W nawiasach ostrokatnych podano kategorie z listy KABA, do której należą argumenty relacji.

*<nazwy geograficzne> stosunki wojskowe <nazwy geograficzne>*

Wykorzystując te informacje, można stworzyć ograniczenie wskazujące, że instancje typu tematu „Jednostki geograficzne” (nazwy geograficzne) mogą, ale nie muszą, odgrywać dowolną liczbę ról typu „państwo-pozostające-w-stosunkach-wojskowych” w powiązaniach typu „państwo-pozostające-w-stosunkach-wojskowych/państwo-pozostające-w-stosunkach-wojskowych”:

*jednostki-geograficzne isa tmcl:topic-type;  
plays-role(państwo-pozostające-w-stosunkach-wojskowych, państwo-pozostające-w-stosunkach-wojskowych/państwo-pozostające w stosunkach wojskowych, 0, \*).*

W JHP KABA określnik „stosunki wojskowe” jest stosowany po nazwach państw i regionów większych. W mapie tematów można zapisać, że tematy występujące w powiązaniu reprezentującym stosunki wojskowe powinny należeć do typów: „Państwa” lub „Regiony”. Wymaga to ich wyodrębnienia z ogólniejszej kategorii tematów reprezentujących jednostki geograficzne.

Podobnie można postąpić w drugim przypadku (z określnikiem „gramatyka porównawcza”).

*<języki> gramatyka porównawcza <języki>*

Wykorzystując zaprezentowane informacje można utworzyć następujące ograniczenie:

*języki isa tmcl:topic-type;  
plays-role(przedmiot-porównywany, przedmiot-porównywany/przedmiot-porównywany, 0, \*).*

Kolejny określnik – „dzieła”, łączy osobę z dziedziną działalności. Na zaprezentowanej w pierwszej części artykułu liście brakuje odpowiedniej kategorii odpowiadającej drugiemu elementowi relacji. Najpierw została ona zapisana jako „Przedmioty”, czyli najogólniejsza kategoria gromadząca wszystkie przedmioty. W nawiasie klamrowym została przedstawiona propozycja bardziej szczegółowej kategorii.

*<nazwy osobowe> dzieła <przedmioty> {dziedziny}*

Ograniczenia dotyczą w tym przypadku dwóch typów tematów, które mogą grać w omawianym typie powiązania dwie osobne role.

*osoby isa tmcl:topic-type;  
plays-role(osoba, osoba/dziedzina-działalności, 0, \*).*  
*dziedziny isa tmcl:topic-type;  
plays-role(dziedzina-działalności, osoba/dziedzina-działalności, 0, \*).*

Tę samą relację identyfikuje również hasło o innej budowie:

*<nazwy osobowe> wkład do <przedmioty> {dziedziny}*

Podczas jego przekształcenia można zastosować ograniczenia przedstawione powyżej.

W przypadku bardziej złożonej transformacji hasła z określnikiem „wpływ [...]”, wśród kategorii zamieszczonych na liście określników, również brakuje odpowiedniej nazwy wskazującej grupę przedmiotów.

*<przedmioty> {przedmioty-z-dziedziny-sztuki-kultury-literatury} wpływ <przedmioty> {przedmioty-z-dziedziny-sztuki-kultury-literatury}*

Zapis TMCL wskazuje, że w powiązaniu „przedmiot-wpływający/przedmiot-na-który-wywierany-jest-wpływ” role odgrywają tematy należące do typu „Przedmioty z dziedziny sztuki, kultury i literatury”.

*przedmioty-z-dziedziny-sztuki-kultury-literatury isa tmcl:topic-type;*  
*plays-role(przedmiot-wpływający, przedmiot-wpływający/przedmiot-na-który-wywierany-*  
*-jest-wpływ, 0, \*);*  
*plays-role(przedmiot-na-który-wywierany-jest-wpływ, przedmiot-wpływający/przedmiot-na-*  
*-który-wywierany-jest-wpływ, 0, \*).*

Powyżej podano jedynie przykład jednego z możliwych ograniczeń. Na inne możliwości zapisu wskazuje nota stosowania określnika „wpływ [przymiotnik od nazwy cywilizacji, religii, obszaru geograficznego, grupy etnicznej]”. W powiązaniu tego typu mogą występować wspólnie tematy należące do wielu typów.

Relację wpływu identyfikują także inne określniki o konstrukcji „skutki oddziaływania [przedmiotu]”. Podobnie jak w przypadku określnika „wpływ [przymiotnik od nazwy cywilizacji, religii, obszaru geograficznego, grupy etnicznej]”, pierwszy argument relacji może należeć do różnych kategorii, na co wskazuje nota stosowania określnika. W tym miejscu pokazano jedynie przykładowe zestawienie typów tematów:

*<organizmy żywe> skutki oddziaływania <substancje chemiczne>*

Poniższe ograniczenia odnoszą się do przedstawionego przykładu.

*organizmy-żywe isa tmcl:topic-type;*  
*plays-role(przedmiot-na-który-wywierany-jest-wpływ, przedmiot-wpływający/przedmiot-na-*  
*-który-wywierany-jest-wpływ, 0, \*).*  
*substancje-chemiczne isa tmcl:topic-type;*  
*plays-role(przedmiot-wpływający, przedmiot-wpływający/przedmiot-na-który-wywierany-*  
*-jest-wpływ, 0, \*).*

Obydwa argumenty relacji identyfikowanej przez określnik „deportacje z [nazwa kraju lub regionu większego niż kraj]” należą do tej samej kategorii. Zaprezentowano ją poniżej.

*<nazwy geograficzne> deportacje z <nazwy geograficzne>*

W ograniczeniu występuje zatem jeden typ tematu:

*jednostki-geograficzne isa tmcl:topic-type;*  
*plays-role(państwo-źródłowe-deportacji, państwo-źródłowe-deportacji/państwo-docelowe-*  
*-deportacji, 0, \*);*  
*plays-role(państwo-docelowe-deportacji, państwo-źródłowe-deportacji/państwo-docelowe-*  
*-deportacji, 0, \*).*

W kolejnym typie powiązania opartym na opisie zawierającym określnik „więźniowie i więzienia [nazwa narodowości]”, argumenty należą do odmiennych kategorii.

*<kategorie osób> więźniowie i więzienia <grupy etniczne>*

Niezbędne są dwa ograniczenia:

*kategorie-osób isa tmcl:topic-type;*  
*plays-role(więziony, więziony/osoby-które-więżą, 0, \*).*

*grupy-etniczne isa tmcl:topic-type;*  
*plays-role(osoby-które-więźq, więziony/osoby-które-więźq, 0, \*).*

Następny przykład reprezentuje relację przedstawienia przedmiotu w medium. Na dziedzinowych listach określników ponownie brakuje odpowiedniej kategorii. W nawiasie klamrowym zaproponowano szczegółową propozycję.

*<przedmioty> w <przedmioty> {media}*

W różnych mediach mogą być prezentowane rozmaite przedmioty, co uwzględniono w notach stosowania takich określników jak „w kinie”. Niezbędne są dwa ograniczenia odnoszące do dwóch typów tematów:

*przedmioty isa tmcl:topic-type;*  
*plays-role(przedmiot, przedmiot/medium, 0, \*).*

*media isa tmcl:topic-type;*  
*plays-role(medium, przedmiot/medium, 0, \*).*

W poprzedniej części artykułu wskazano, że określnik „i [...]” może identyfikować różne rodzaje relacji, które w przeciwieństwie do LCSH nie są rozróżniane w KABA. Bez analizy znaczenia tych połączeń nie można określić szczegółowego rodzaju powiązania w mapie tematów.

*<osoby> i <przedmioty> {dziedziny}*

Ograniczenia prezentują wariant, w którym wyróżniono szczegółowy rodzaj powiązania – relację między osobą a dziedziną poznania, przedstawiającą znajomość dziedziny przez daną osobę.

*osoby isa tmcl:topic-type;*  
*plays-role(osoba, osoba/dziedzina-poznania, 0, \*).*

*dziedziny isa tmcl:topic-type;*  
*plays-role(dziedzina-poznania, osoba/dziedzina-poznania, 0, \*).*

## 5. Zakończenie

W artykule przedstawiono możliwości i problemy przekształcenia wybranych relacji, występujących w pojedynczych hasłach rozwiniętych KABA i całych opisach przedmiotowych do postaci powiązań w mapie tematów. W JHP KABA, oprócz podziału tematów i określników zgodnego z etykietami pól formatu MARC 21, dokonano kategoryzacji określników w celu kontroli połączeń między jednostkami leksykalnymi w hasłach rozwiniętych. Dane na ten temat są zapisane w rekordach wzorcowych określników (pole 073). Brakuje ich jednak w rekordach wzorcowych tematów (pole 072), co uniemożliwia automatyczną kontrolę poprawności hasel rozwiniętych. Pokazano, że uzupełnienie tych informacji tylko w ograniczonym zakresie może być pomocne przy przekształceniu języka do postaci mapy tematów. W przypadku niektórych jednostek identyfikujących argumenty relacji konieczne było zaproponowanie dodatkowych kategorii, które pozwoliły na zapis ograniczeń dotyczących powiązań. Wydzielono m.in. typy tematów „Dziedziny” i „Media”. Podczas tworzenia systemu organizacji wiedzy powinno się zwracać uwagę na możliwości

jego formalizacji, czyli opisu łączliwości jednostek leksykalnych i zasad pragmatycznych w sposób umożliwiający automatyczny nadzór nad tworzonymi charakterystykami wyszukiwawczymi. W języku KABA nie dokończono tego procesu.

W artykule zaprezentowano przykłady przekształcenia wybranych haseł i opisów przedmiotowych do postaci powiązań w mapie tematów. Kryterium doboru było występowanie w ich składzie określników rzeczowych reprezentujących relacje między przedmiotami. Podział całości zasobu leksykalnego KABA byłby dużo bardziej złożony i wymagałby utworzenia licznych ograniczeń. Zalecane byłoby przynajmniej częściowe zautomatyzowanie tej pracy.

Przedstawione rozwiązanie cechuje się większymi możliwościami wielowymiarowego opisu treści dokumentów w porównaniu z przekształceniami JHP odwołującymi się do tradycyjnych języków deskryptorowych. Dzięki zachowaniu kontekstu semantycznego umożliwia ono opracowanie rzeczowe zarówno za pomocą pojedynczych tematów, jak i tematów połączonych określonymi powiązaniami. W systemie opartym na mapie tematów, zapytania mogłyby być konstruowane w oparciu o jawnie specyfikowane relacje między tematami zapewniając większą precyzję wyszukiwania. Dzięki zaproponowanemu rozwiązaniu, nawiązującemu do teorii indeksowania relacyjnego Jasona Farradane'a (zob. np. Chmielewska-Gorczyca, 1981), możliwe jest wykorzystanie informacji dotyczących treści i formy dokumentów zawartych w dotychczas utworzonych opisach przedmiotowych.

## Bibliografia

- BnF (2016). *Guide d'indexation RAMEAU* [online]. Bibliothèque nationale de France, [07.07.2016], <http://guiderameau.bnf.fr>
- Cichoń, M.; Kalinowski, J.; Federowicz, G. (2014). Katalogowanie oparte na encjach. *Rocznik Biblioteki Narodowej*, 45, 151–200.
- Chmielewska-Gorczyca, E. (1981). Metoda indeksowania relacyjnego Farradane'a. *Zagadnienia Informacji Naukowej*, 38(1), 83–97.
- Ćwikowski, P. (2013). *Opis przedmiotowy dokumentów na temat wojen i wojskowości*. Warszawa: Wydaw. SBP.
- ISO (2006). ISO/IEC 13250–2:2006. Information technology – Topic Maps – Part 2: Data model.
- ISO (2011). ISO/IEC 19756:2011. Information technology – Topic Maps – Constraint Language (TMCL).
- Język deskryptorowy (2002). W: *Słownik encyklopedyczny informacji, języków i systemów informacyjno-wyszukiwawczych*. Warszawa: Wydaw. SBP, 102.
- LoC (b.d.). *Library of Congress Authorities* [online]. Library of Congress, [07.07.2016], <http://authorities.loc.gov/>
- LoC (1999). 073 – *Subdivision Usage (NR)* [online]. Library of Congress, [07.07.2016], <https://www.loc.gov/marc/authority/ad073.html>
- LoC (2012). *Library of Congress to Add Subject Category Codes to Selected Subject Authority Records* [online]. Library of Congress, [07.07.2016], [http://loc.gov/catdir/cpsd/field\\_072\\_announcement.pdf](http://loc.gov/catdir/cpsd/field_072_announcement.pdf)
- LoC (2016). *Free-floating Subdivisions* [online]. Library of Congress, [07.07.2016], <https://www.loc.gov/aba/publications/FreeLCSH/SUBDIVISIONS.pdf>
- NUKAT-a (b.d.). *Katalog NUKAT* [online]. Centrum NUKAT, [07.07.2016], <http://katalog.nukat.edu.pl/search/query?theme=nukat>
- NUKAT-b (b.d.). *Listy dziedzinowe określników KABA* [online]. Centrum NUKAT, [07.07.2016], <http://centrum.nukat.edu.pl/pl/warsztat/slovnictwo-jhp-kaba?view=okreslniki>
- NUKAT-c (b.d.). *Listy dziedzinowe określników KABA – rzeczowe* [online]. Centrum NUKAT, [07.07.2016], <http://centrum.nukat.edu.pl/pl/warsztat/slovnictwo-jhp-kaba?view=okreslniki&layout=rzeczowe>

Włodarczyk, B. (2016). *Mapy tematów jako podstawa semantyzacji języka haseł przedmiotowych*. Warszawa: Wydaw. SBP.

---

## The Transformation of Relationships in KABA Subject Heading Strings into Associations in Topic Maps

### Abstract

**Purpose/Thesis:** The purpose of the paper is to present the transformation of selected KABA subject heading strings into associations in topic maps.

**Approach/Research methods:** The possibility of introducing more detailed categorization of KABA Subject Headings was assessed based on the documentation published by the NUKAT Center. The analysis of the scope notes of KABA topical subdivisions was used to present the types of constructions occurring in the system and the options of their transformation into associations in topic maps.

**Results and conclusions:** KABA Subject Headings subdivision authority records already are had been populated with categories (073 field). However, the authority records for subject headings lack these data (expected to appear in field 072) and consequently the automatic heading string validation is impossible. Yet the addition of the categories can be helpful in transforming subject headings into topic map only to a limited extent. There is a need to create additional categories which enable the creation of constraints on topic map ontology. The paper presents examples of the transformation of selected subject heading strings. The selected strings had to include subdivisions representing relationships between subjects. The division of the whole vocabulary would be far more difficult and requiring numerous constraints. It would be advisable to automate this process, at least some of its aspects.

**Originality/Value:** The author puts forward the proposal to transform the relationships within KABA subject heading strings into associations in topic maps. There has not been any similar analysis recorded in library and information science literature before. Owing to the compatibility of KABA, RAMEAU and LCSH systems, the results of the research can be at least partly applied in the transformation of those tools.

### Keywords

KABA Subject Headings. Relationships. Subject heading strings. Topic Maps.

---

Dr BARTŁOMIEJ WŁODARCZYK – adiunkt w Katedrze Bibliografii i Dokumentacji na Wydziale Dziennikarstwa, Informacji i Bibliologii Uniwersytetu Warszawskiego. Jego zainteresowania badawcze dotyczą problematyki organizacji informacji i wiedzy. Najważniejsze publikacje: B. Włodarczyk: *Mapy tematów jako podstawa semantyzacji języka haseł przedmiotowych*. Warszawa 2016; B. Włodarczyk: *Deskryptory formalne w Bibliotece Kongresu a tematy formalne w języku haseł przedmiotowych Biblioteki Narodowej i w języku KABA*. *Przegląd Biblioteczny* 2014, 4, 518–536; B. Włodarczyk: *Topic Map as a Method for the Development of Subject Headings Vocabulary: An Introduction to the Project of the National Library of Poland*. *Cataloging & Classification Quarterly* 2013, 7, 816–829.

Kontakt z autorem:

bm.wlodarczyk@uw.edu.pl

Katedra Bibliografii i Dokumentacji

Wydział Dziennikarstwa, Informacji i Bibliologii

Uniwersytet Warszawski

ul. Nowy Świat 69

00–927 Warszawa

# Credibility of Information: an Overview of the Concept in the Digital Age

Magdalena Shishenkova

*GERiiCO Laboratory, University of Lille, France*

---

## Abstract

**Purpose/thesis:** The primary objective of this paper is to review the concept of credibility of information in the context of the digital environment and in the light of research over the last decade. Its purpose is to synthesize related concepts and trends in perceptions of credibility in interdisciplinary dialogue.

**Approach/methods:** A review of the relevant scholarly literature. The study sketches how the concept of credibility has emerged in philosophy, psychology, communication and information sciences.

**Results and conclusions:** A number of constructs associated to credibility have been identified and some related notions in information research in the last decade have been summarized. There is presented a short overview of elements of the digital context, that influence this notion. The author of the article suggests that recent research provides grounds for more than a communicative approach to authority construct and acknowledgment of diversity.

**Originality/value:** The article extends the dialogue by outlining the scope of the concept, its complexity, together with its multi-disciplinary value.

## Keywords

Credibility of information. Digital content. Digital age. User perception.

*Received: 2 November 2015. Reviewed: 8 December 2015. Accepted: 12 July 2016.*

---

## 1. Introduction

Information credibility is a central consideration in information science when developing a collection or engaging in reference interaction. Whilst information professionals and scientists are trained in how to critically analyse information sources, non-professionals are not armed with the necessary tools to methodologically identify credible information. Nowadays, when digital collections tend to form as miscellaneous masses of information, communication platforms or open access repositories, the user is challenged with multiple issues to resolve. Amongst them are the credibility features of the information they obtain.

Digital, open access environment is defined as democratic by nature. Weinberger noticed, that “The Web (...) breaks the traditional publishing model” (Weinberger, 2002, X). Everyone, having or not having official recognition can express her/his viewpoint, and use and forecast information freely. The latter represents a challenge not only for the mass users, but it also invites information specialists and scientists of all disciplines to rethink the way they trust sources. The knowledge out there might be worthy of knowing according to Surowiecki, especially considering the argument that Google was built and it functions

upon it (Surowiecki, 2004, 16). So does Wikipedia. Collective intelligence can bring an important contribution to science and we cannot afford to ignore it for the sake of the benefits it can bring. The critical issues are, whom and how do we trust? Who and what do we consider credible? These are difficult questions, yet there are attempts to automate the process of selection. This is, for example, reflected in the latest work of Google towards refining search results on the basis of facts and truthfulness (Kottasova, 2015).

Making sense of the concept in the digital age can lead to a better understanding of how the environment influences users' behaviour and vice versa. It can also contribute to the improved management of data quality and the development of information literacy approaches.

In this context, in the present paper I will firstly discuss my motivation to explore the topic. Secondly, I will seek to make sense of the concept of information credibility. Then I will trace how the concept has evolved and been transformed over the last decade. Finally, drawing upon the latter, I will discuss the limitations of and directions for future work.

## **2. Motivation: Why credibility in digital age?**

The attempt to understand the real world has always been the core intellectual preoccupation of science. Today science faces the urge to explore multiple phenomena in a world of human-driven changes. The Web is used extensively by many people and it is important to acknowledge that the collaborative media offers new and, at times, surprising insights about human thinking and actions. In the present time, not only science but also politics and marketing utilize certain instruments to extend the understanding about the user's information seeking behaviour. For example, there is big data at the tips of marketers' fingers to mathematically model expected purchases based on information users' behaviour data. Data only represents figures until it is interpreted in the light of human analysis and judgment. Both scientist and practitioners are still at the very beginning of analysis of big data beyond pure calculations. The starting point in this article is that in information science the concept of credibility in the digital world should be based on a better understanding of how our current reality of interaction in digital environment shapes our thinking in the area of which information is credible. My motivation to seek an understanding of credibility as a concept in the digital age has been inspired by several discussions in interdisciplinary space. Firstly, a number of authors have addressed the impact of the information found on the Web on everyday life. Research in the last decade suggests that the information found in the virtual medium has had a strong impact on real-life decisions such as finances and health (e.g. Betsch et al. 2012; Doty, 2015; Zhao et al. 2015). Individual choices of the source as well as the selected information not only matter, but strongly influence personal life. Therefore, it is important to understand how credibility is viewed by users in the context of the digital environment and to explore how the concept is shaped by the very presence of digital content. Another important catalyst for this article are the discussions both in the scientific and public space concerning the approaches to developing models facilitating choices of information. An example of such a tool is the work in the field of automated extraction of facts from Web pages. Among the many contributors to the topic are a number of Google scientists (Dong, 2015). Their work will certainly have important influence

on information users. How the scientists involved in the field see information credibility could potentially shape the virtual environment and become an essential underlying factor influencing users' information choices. Last, but not least, an inspiration for this paper is the intangible nature of the networked environment. On the Web many choices can dissipate our focus and question our beliefs. The latter, combined with the perplexed and subjective construct of credibility, offers in my view an opportunity to open further discussions in a world of human-driven changes.

### 3. Making sense of the concept “information credibility”

#### 3.1. *Theoretical framework: philosophy, psychology, communication science*

Credibility is an interdisciplinary concept. It is impossible to encompass everything said about it over time in variety of disciplines. However, I will attempt to situate the concept by briefly outlining some of the viewpoints on it.

We find interpretations of the credibility concept in the work of important philosophers even before epistemology was defined as a field (Goldman et al., 2015). The key questioning, related to credibility in the work of philosophers was related to the judgment of how we know that our knowledge is true and objective or how to distinguish truth from non-truth. Philosophy offers answers to the latter within different frameworks. Generally, these can be divided in three epochs: pre-modernism, modernism, and postmodernism. Pre-modernism regards the ultimate truth as held by God and thus only revealed to humans. Modernism encompasses the 17<sup>th</sup> to the 20<sup>th</sup> century and it emphasises reasoning, empiricism and methodology as the only path to truth (Locke, Hume, Descartes, Kant). Relativism regards the truth as a function of the experience of the message perceiver (Hume). The pragmatic approach to knowledge also falls into this time-frame (Peirce, Dewey). In optics of pragmatism, “trust” is contextual, and we trust an assumption/belief unless we have a positive reason to doubt it. On the democratic platform of Web 2.0 the subjective facet has strong foundations as everyone can share, comment, build and argue (Weinberger, 2002; Surowiecki, 2004). Credibility is also associated with power, influence and the source position in society or a professional domain (Foucault, Bourdieu). One can be prevented from knowing due to lack of education/access to knowledge (Flicker). In psychology credibility is often interpreted from a cognition perspective, but also as “believability” (Metzger & Flanagin, 2015). The communication model of Shannon & Weaver (message-information-environment-receiver) remains an important framework when interpreting credibility in communication science. Communication studies focus on the media choices and perception of the source of the message, interpersonal transactions, authority and message (Rieh, 2014).

#### 3.2. *Dimensions of credibility in information science: mapping the concept to related terms*

At the core of information and library science is the democratisation of knowledge through assuring access to numerous sources. The latter might not be obtainable in their completeness by the user for multiple reasons such as expense, limited editions, geographic

constraints, language, etc. Libraries by definition select, collect, systemise and make information available. According to Rieh, in information and library science the notion of credibility is heavily “centered on the notion of relevance” (Rieh 2010, XX). In the digital environment, the relevance is only one side of the coin as many of the library procedures and processes are not present. Balancing content, evaluating credibility and value of the sources is an essential process in libraries. In contrast, the digital world concentrates upon accumulating rather than collecting. Thus, the evaluation is left to the user to analyse.

Table 1. Concepts related to credibility

| The level of the message                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | The level of authority/the source                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | The level of the information user                                                                                                                                                                                                                                                                                       | The level of media (features of a website, database, blog, or other digital platform)                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Information quality <sup>4, 10, 16, 19</sup><br>Quality of argumentation <sup>10, 16</sup><br>Factuality <sup>10, 16, 20</sup><br>Quality of the language<br>Information/data integrity <sup>12</sup><br>Expression of first-hand experience (anecdotes) <sup>3</sup><br>Presentation errors <sup>7</sup><br>Believability <sup>6, 10, 16</sup><br>Accuracy <sup>1, 10, 16</sup><br>Coherence/internal consistency <sup>5</sup><br>Freedom from bias, objectivity <sup>6, 13</sup><br>Informativeness, depth <sup>10</sup><br>Authenticity <sup>9</sup><br>Currency, relevance <sup>1, 4, 10, 15</sup><br>Usefulness <sup>7, 10, 16</sup><br>Clarity <sup>16</sup><br>Structure <sup>16</sup><br>Specificity <sup>16</sup><br>Completeness <sup>6, 16</sup> | Authoritativeness <sup>12</sup><br>Cognitive Authority vs Regulatory <sup>1, 3, 11</sup><br>Reliability <sup>10</sup><br>Competence <sup>15</sup><br>Experience (first-hand) <sup>3</sup><br>Expertise <sup>7, 10, 14</sup><br>Formal education; scholarliness <sup>7</sup><br>Objectivity <sup>9</sup><br>Officialness <sup>10</sup><br>Other characteristics, that may trigger stereotyping (ex. sex) <sup>9</sup><br>Reputation <sup>10</sup><br>Degree of formal recognition <sup>2</sup><br>Trust <sup>1, 6</sup><br>Trustfulness, freedom from bias <sup>6, 8, 9, 10, 12, 14</sup> | Knowledge <sup>10, 12</sup><br>Culture <sup>10, 16</sup><br>Expertise <sup>10, 16</sup><br>Experience <sup>10, 16</sup><br>Stereotypes (religion, political) <sup>9, 17</sup><br>Information need/purpose or research <sup>10, 16</sup><br>Other characteristics, that may trigger stereotyping (ex. sex) <sup>17</sup> | Appearance (Design/interface/organization) <sup>2, 13, 18</sup><br>Accessibility <sup>13</sup><br>Safety <sup>8</sup><br>Information quality<br>Information/data integrity <sup>6, 10, 16, 18</sup><br>Trustworthiness <sup>14</sup><br>Authority <sup>3, 11</sup><br>Formal recognition <sup>2</sup><br>Reputation with regard to fact coverage<br>Believability <sup>9</sup><br>Usefulness <sup>10, 16</sup><br>Impartiality/freedom from bias/neutrality/objectivity <sup>10, 16</sup><br>Popularity <sup>15</sup> |
| Sources:<br>1. Beldad, A., et al., 2010<br>2. Bergson-Michelson, 2012<br>3. Doty, 2015<br>4. Metzger & Flanagan, 2013<br>5. Batini et al. In : Floridi et. al., 2014<br>6. Metzger & Flanagan, 2015<br>7. Metzger & Flanagan, 2013<br>8. Nielek et al. 2013<br>9. Rieh, 2009<br>10. Rieh, 2010<br>11. Wilson quoted in Rieh, 2010<br>12. Choi & Stvilia, B., 2015<br>13. Lackaff & Cheong, 2008.<br>14. Rieh et al., 2014<br>15. Nurse et al., 2014<br>16. Rieh and Danielson, 2007<br>17. Savolainen, 2011<br>18. Flanagan & Metzger, 2007<br>19. Dong, 2015                                                                                                                                                                                               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                                                                                                                                                                                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |

In the first place, although it is an intuitive concept, credibility is not a simple construct. Traditionally, in information and library science credibility is seen in relation to retrieval and the preservation of information. Therefore, it is described in terms such as relevance, accuracy, currency, objectivity, usefulness and reliability, etc. (Rieh, 2014). Being a multi-layer construct, the definition of credibility in literature is related to other notions, which are also referred to as dimensions (Choi, 2015) or “sub-factors” (Nurse et al., 2014). Table 1 summarises the related concepts as discussed in literature in the last decade. The proposed categorisation of the levels of credibility is based on the distinction that is elaborated in communication studies and which is influential for the information science as well the ones offered in the table below. The related terms are selected across sources predominantly from the information science literature, but some of them are used also in computer science and communication studies.

Several of the concepts in this summary represent cues for evaluation of credibility in the assessment process. Alternatively they represent characteristics of the authority, source, and message (Rieh & Danielson, 2007). Further, literature suggests that credibility is a subjective perception and the information user perspective has to be taken into consideration (Metzger & Flanagin, 2015; Rieh, 2010; Nielek et al., 2013). The understanding, experience, values, attitudes, culture and other characteristics of the user are crucial for the interpretation of facts.

It is evident, that there is no firm definition of the concept of credibility. However there are number of concepts that stand out. It is interesting for example, that authenticity is one of the pillars for determining the credibility of the message. The concept of credibility is heavily interrelated with its authorship. It is also associated with archival science in its endeavour to preserve unique documents where attribution is crucial. Authenticity is becoming a major concern for Google as well, in replacing popularity, together with the factual and accuracy dimensions:

we define the trustworthiness or accuracy of a Web source as the probability that it contains the correct value for a fact [...] assuming that it mentions any value for that fact (Dong et al., 2015).

#### **4. Concept transformation in the context of digital environment over the last decade**

The digital medium sets up a new framework for the notion of credibility. Authors often describe the concept as a function of trust and authority (Choi, 2015; Jiong & Rieh, 2014; Metzger & Flanagan, 2015; Rieh, 2010; Slavolainen, 2011). The Web challenges the judgment of authority and thus leaves the users with less or no basis for judgment. Yet, there is evidence that authority is still presumed to be an important element weighed upon the dichotomy expert – non-expert. According to Doty (2015), the interaction online implies establishing Cognitive authority of the Self and a Cognitive authority of Others. For example, users tend to claim recognition for their opinions by publishing comments such as “Being a nurse I find this article very disturbing. ...” (Doty, 2015, 4). The phrase “Being a nurse” aims to strengthen the health-related statement in the eye of the reader. However, in the virtual world there are limited cues to support such a declaration of competence. Therefore users find ways to handle the uncertainty and deal with it using a variety of strategies with

regard to reliability of information. There are a number of studies regarding the strategies and a summary of typologies of these, concerning credibility is offered by Rieh (2009) (see table 2). According to that research, users also tend to avoid complex cross-checking for their validation of information. While deciding which resource to trust, users pick the information that comes up first as a relevant result. The reason given is that the obtained information satisfies the minimum requirements necessary to meet a particular information need. In the decision making theory of Herbert Simon this phenomenon is known as “satisfying” (Goodrich et al, 2000). In communication and information sciences it is referred to as the heuristics (Metzger & Flanagin, 2013).

Table 2. Typologies of credibility

| Tseng & Fogg typology                                                                                                                                                                                                                         | Flanagin & Metzger typology                                                                                                                                                                                                        |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Presumed credibility (based on general assumptions or stereotype)<br>Reputed credibility (endorsement from people, media, source)<br>Surface credibility (from simple inspection)<br>Experienced credibility (based on first-hand experience) | Conferred credibility (recommended or produced by well-regarded entities)<br>Tabulated credibility (influenced by other individual's ratings or recommendations)<br>Emergent credibility (arises from group and social engagement) |

Source: (Rieh, 2009)

The digital environment is complex. Firstly, it offers a shared experience to which probably users would never have access to (in a library for example). In this sense the Internet is a place where “repressed testimony is at last being heard” (Doty, 2015, 7). This is a knowledge users may benefit from (see Surowiecki, 2004, XX). Secondly, according to Jeon and Rieh, in digital content there is a “lack of quality control mechanisms and a limited number of available cues” to assess information (Jeon & Rieh, 2014, 1). That implies a complex situation: a necessity both to judge the credibility of strangers who disseminate information and the reliability of their message while it is impossible to weigh the cognitive authority (Rieh et al, 2014). The digitized environment is a platform that allows equal possibility for each and every participant in the discourse to express oneself. It is egalitarian by nature. It is not yet stratified in the way that societies are. Although the user is independent while making decisions online, she/he is prone to certain biases. Scientists discuss number of phenomena on the Web such as echo chambers or reinforcing believes inside an “enclosed” system; networking while interacting for obtaining information online; and the bandwagon effect or the tendency to follow the beliefs of others (Weinberger, 2002; Pentland, 2014). Along the same lines, Lankes suggests that in the digital environment we observe a shift from “cognitive authority” to “reliability” (coherence across sources) (Lankes, 2008, 220). Lankes also supports Dervin (2011) in describing knowledge as created through conversation. The notion of “expertise” in its traditional sense is no longer applicable with the expansion of the “wisdom of crowds” (Surowiecki, 2004). Thirdly, discussing Wikipedia, Garfinkel notices that “within this framework it is not about >>truth<<, but rather about >>verifiability<< and the >>neutral point of view<<” (Garfinkel, 2008). The results of research by Zhao et al. which focused on elderly population financial decisions suggested that access to “neutral” sources benefits decision-making even in the circumstances of uncertainty (Zhao et al., 2015).

Further, the ease of discovery of information in digital environment challenges the credibility notion in information science in the part of its relevancy. How fast and easy the users find information in linked environment becomes an essential feature of credibility. Weinberger (2015, XX) gives a useful example with regard to that, referring to the libraries “as [library] information becomes harder to find, it becomes less relevant”.

Thus, the users apply the heuristics strategy in order to “shortcut” through the information for the sake of efficacy (Metzger et al., 2010). In these terms, technology sets up a completely different framework with regard to the conceptualising credibility: the Web implies operational approaches to credibility. For example, the influential Google literacy lessons on the Web offer a paradigm for judgment based on “how much you find X believable/consistent with your beliefs” (Bergson-Michelson, 2012, 8:00”).

The linked environment is another aspect of digital content that contributes to the transformation of the notion of “credibility”. For example, Google PageRank algorithm is based extensively on linking potentiality of the Internet and, thus, until lately, it has been built upon the popularity of sources (Brabazon, 2006; Dong, 2015). However, Google has reconsidered its mechanism in the view of credibility. A Google Research team has estimated factuality as important when extracting information from the Web and they have recently made a strong contribution to the automatic evaluation of accuracy. Knowledge Based Trust (KBT) is a concept that contributes to the understanding of credibility in the digital age. KBT has been interpreted as correspondence of the data present on a Web page to the real-world facts (Dong, 2015). Even with imperfections of the developed probabilistic model, the work suggests that in the future ranking will no longer be matter of pure linking or popularity (Dong et al., 2015). Of course, this contribution only in part covers the complex notion of credibility. It is definitely a major shift on the Internet that is worthy of acknowledgment.

Finally, features of media/source such as visual representation have an important role in perceiving a message as being reliable (Metzger et al., 2010; Lankes, 2008; Nurse et al., 2014). If in the past the visual features of an encyclopaedia were of less importance to judging its credibility, unless probably the weight of a volume, today the design matters.

The credibility in digital environment differs from the credibility in traditional systems in certain aspects. More specifically, the concept is more dynamic, time-sensitive and open to dialogue. It is subjective, relative and situational, it is a function of user’s knowledge, goal of the search and other people’s judgments in the networked environment (Rieh, 2010).

## 5. Conclusion

A person or a medium is judged credible because the information coming from her/him/it is unbiased, balanced (neutral), relevant to the topic researched, independent and soundly grounded on facts. Recent research suggests that in the digital world the credibility of the message has gained prevalence over the “authority of the source”. As authority features are often harder to be determined, because of the substantial limits of the environment, they no longer can be leading credibility factors. Personal testimonials are easy to access and dialog between users is a standard. Attribution of the message still matters yet the role of the users’ personal experience is now more important than ever. In addition, elements

of the message and its structure are emerging as more relevant to users' judgments. Dialogue and communicative approach to information is therefore becoming a standard in credibility determination as it is a prerequisite set by the networked environment. In this context future research in information science can focus individual strategies to credibility assessment through communicative approach (Dervin, 2011). In the upcoming research on information quality, researchers can concentrate on how the digital environment sharpens users' sensitivity for science-based evidence and facts-based argumentations. Complexity of the credibility construct implies constant acknowledgment of diversity and therefore a cross-cultural orientation of future research is suggested.

## References

- Batini, C.; Palmonari M.; Viscusi, G. (2014). Opening the closed world: A survey of information quality research in the wild. In: Floridi, L.; Illari, P. (eds.). *The Philosophy of Information Quality*. Heidelberg: Springer, 43–74.
- Beldad, A.; De Jong, M.; Steehouder, M. (2010). How shall I trust the faceless and the intangible? A literature review on the antecedents of online trust. *Computers in Human Behavior*, 26 (5), 857–869.
- Bergson-Michelson, T. G. I. (2012). *Believe it or not: Authority and credibility of sources on the Web*, Google, [video file], [12.07.2016], <https://www.youtube.com/watch?v=vUp3P-fPmBM>
- Betsch, C.; Brewer, N. T.; Brocard, P.; Davies, P.; Gaissmaier, W.; Haase, N.; ... Stryk, M. (2012). Opportunities and challenges of Web 2.0 for vaccination decisions. *Vaccine*, 30(25), 3727–3733.
- Brabazon, T. (2006). The Google effect: Googling, blogging, wikis and the flattening of expertise. *Libri*, 56, 157–167.
- Choi, W.; Stvilia, B. (2015). Web credibility assessment: Conceptualization, operationalization, variability, and models. *Journal of the Association for Information Science and Technology*, 66(12), 2399–2414.
- Dervin, B. (2011). *Connecting with Specific Publics: Treating Communication Communicatively. Eastern spotlight*, [video file], [12.07.2016], <https://www.youtube.com/watch?v=foyH6eoIseQ>
- Dong, X. L. (2015). Knowledge Vault and Knowledge-Based Trust – Xin Luna Dong of Google. Stanford University, [video file], [12.07.2016], <https://www.youtube.com/watch?v=Z6tmDdrBnpU>
- Dong, X. L.; Gabrilovich, E.; Murphy, K.; Dang, V.; Horn, W.; Lugaresi, C.; Sun, S.; Zhang, W. (2015). *Knowledge-Based Trust: Estimating the Trustworthiness of Web Sources*, [online], arXiv preprint arXiv:1502.03519, [12.07.2016], <http://arxiv.org/pdf/1502.03519.pdf>
- Doty, C. (2015). Social Epistemology and Cognitive Authority in Online Comments about Vaccine Safety. *iConference 2015 Proceedings*, [online], [12.07.2016], [https://www.ideals.illinois.edu/bitstream/handle/2142/73664/210\\_ready.pdf](https://www.ideals.illinois.edu/bitstream/handle/2142/73664/210_ready.pdf)
- Flanagin, A. J.; Metzger, M. J. (2007). The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. *New Media & Society*, 9(2), 319–342.
- Garfinkel, S. L. (2008). Wikipedia and the Meaning of Truth. *Technology Review*, 111(6), 84–86.
- Goldman, A.; Blanchard T. (2015). Social Epistemology. In: E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2015 ed.). Stanford: The Metaphysics Research Lab.
- Goodrich, M. A.; Stirling, W. C.; Boer, E. R. (2000). Satisficing Revisited. *Minds and Machines: Journal for Artificial Intelligence, Philosophy and Cognitive Science*, 10(1), 79–109.
- Jeon, G. Y. R.; Reih, S. Y. (2014). Answers from the crowd: How credible are strangers in social Q&A? [online] *iConference 2014*, 663–668, [12.07.2016], <https://www.ideals.illinois.edu/handle/2142/47266>

- Kottasova, I. (2015). The truth, according to Google [online]. *CNN Money*, [12.07.2016], <http://money.cnn.com/2015/03/04/news/google-truth-ranking/>
- Lackaff, D.; Cheong, P. H. (2008). Communicating authority online: Perceptions and interpretations of Internet credibility among college students. *Open Communication Journal*, 2, 143–155.
- Lankes, R. D. (2008). Credibility on the internet: shifting from authority to reliability. *Journal of Documentation*, 64(5), 667–686.
- Metzger, M. J.; Flanagin, A. J.; Medders, R. B. (2010). Social and heuristic approaches to credibility evaluation online. *Journal of Communication*, 60(3), 413–439.
- Metzger, M. J.; Flanagin, A. J. (2013). Credibility and trust of information in online environments: The use of cognitive heuristics. *Journal of Pragmatics*, 59, 210–220.
- Metzger, M. J.; Flanagin, A. J. (2015). Psychological approaches to credibility assessment online. In: S. Shyam Sundar (ed.) *The Handbook of the Psychology of Communication Technology*, New Jersey: Wiley-Blackwell, 445–466.
- Nielek, R.; Wawer, A.; Jankowski-Lorek, M.; Wierzbicki, A. (2013). Temporal, cultural and thematic aspects of web credibility. *Social Informatics*, 419–428.
- Nurse, J. R. C.; Agraftotis, I.; Goldsmith, M.; Creese, S.; Lamberts, K. (2014). Two sides of the coin: measuring and communicating the trustworthiness of online information. *Journal of Trust Management*, [online], 1(5), [12.07.2016], <http://journaloftrustmanagement.springeropen.com/articles/10.1186/2196-064X-1-5>
- Pentland, A. (2014). Social Physics: How Good Ideas Spread-The Lessons from a New Science. Talks at Google, [video file], [12.07.2016], <https://www.youtube.com/watch?v=HMBI0ttu-Ow>
- Rieh, S.Y. (2009). Credibility and Human Information Behavior [online]. Presentation at Information Ethics Roundtable Misinformation and Disinformation April 3–4, 2009. University of Arizona, Tucson. [12.07.2016], [http://www.powershow.com/view/3d0dbf-NmIwZ/Credibility\\_and\\_Human\\_Information\\_Behavior\\_powerpoint\\_ppt\\_presentation](http://www.powershow.com/view/3d0dbf-NmIwZ/Credibility_and_Human_Information_Behavior_powerpoint_ppt_presentation)
- Rieh, S. Y. (2010). Credibility and Cognitive Authority of Information. In M. Bates & M. N. Maack (Eds.) *Encyclopedia of Library and Information Sciences*, 3rd Ed. New York: Taylor & Francis, 1337–1344.
- Rieh, S. Y.; Danielson, D. R. (2007). Credibility: A multidisciplinary framework. *Annual review of information science and technology*, 41(1), 307–364.
- Rieh, S. Y.; Morris, M. R.; Metzger, M. J.; Francke, H.; Jeon, G. Y. (2014). Credibility perceptions of content contributors and consumers in social media. *MEETProceedings of the American Society for Information Science and Technology*, 51(1), 1–4.
- Savolainen, R. (2011). Judging the quality and credibility of information in Internet discussion forums. *Journal of the American Society for Information Science and Technology*, 62(7), 1243–1256.
- Surowiecki, J. (2004). *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies, and Nations*. New York: Doubleday.
- Weinberger, D. (2002). *Small Pieces Loosely Joined: A Unified Theory of the Web*. Cambridge, MA: Perseus.
- Weinberger, D. (2015). *Is there a library-sized hole in the internet?* [online], [12.07.2016], [http://www.researchinformation.info/news/news\\_story.php?news\\_id=1847](http://www.researchinformation.info/news/news_story.php?news_id=1847)
- Zhao, J. C.; Fu, W. T.; Zhang, H.; Zhao, S.; Duh, H. B. L. (2015). *To Risk or Not to Risk? Improving Financial Risk-Taking of Older Adults by Online Social Information*, [online], [12.07.2016], [http://home.uchicago.edu/hanzhe/research/1208peereffects\\_WP.pdf](http://home.uchicago.edu/hanzhe/research/1208peereffects_WP.pdf)

## Wiarygodność informacji w epoce cyfrowej: przegląd wybranych koncepcji

### Abstrakt

**Cel/teza:** Podstawowym celem artykułu jest prezentacja wybranych koncepcji wiarygodności informacji w odniesieniu do środowiska cyfrowego z ostatnich dziesięciu lat. Dodatkowym celem badań była próba syntetycznego ujęcia pojęć związanych z zagadnieniem wiarygodności informacji w ujęciu interdyscyplinarnym.

**Koncepcja/metody badań:** W badaniach zastosowano metodę analizy i krytyki piśmiennictwa. Uwzględniono również próby interpretacji pojęcia wiarygodności informacji w ujęciu filozoficznym, psychologicznym oraz z punktu widzenia nauki o komunikacji oraz informacji naukowej.

**Wyniki i wnioski:** W artykule wskazano na kilka konstruktów pojęciowych związanych z pojęciem wiarygodności oraz scharakteryzowano próby ich interpretacji w kontekście badań nad informacją w ostatnich dziesięciu latach. Przedstawiono również zagadnienie wpływu środowiska cyfrowego na postrzeganie zjawiska wiarygodności informacji. W artykule stwierdzono, że dotychczasowe wyniki badań dają podstawę do dalszych dociekań nad zjawiskiem wiarygodności informacji wykraczających poza ujęcie komunikacyjne.

**Oryginalność/wartość poznawcza:** Zaprezentowane wyniki badań stanowią głos w dyskursie naukowym, którego przedmiotem jest zagadnienie wiarygodności informacji poprzez charakterystykę jego zakresu pojęciowego i wewnętrznej złożoności oraz jego multidyscyplinarny charakter.

### Słowa kluczowe

Wiarygodność informacji. Treści cyfrowe. Epoka cyfrowa. Percepcja użytkownika informacji.

---

*MAGDALENA SHISHENKOVA holds Master of Information Science degree from University of Montreal, Canada and works as a Clinical Research Associate. She is a doctoral candidate at the University of Lille. Her interests are in the field of knowledge organization and credibility of information.*

*Contact to the Author:*

*magdalena.shichenkova@etu.univlille3.fr*

*GERiiCO Laboratory*

*Université de Lille 3*

*Domaine universitaire du «Pont de Bois», BP 60149, 59653 – Villeneuve d'Ascq Cédex, France*

# Pole semantyczne terminu „książka elektroniczna”

Jarosław Pacek

*Biblioteka Narodowa*

---

## Abstrakt

**Cel/teza:** Celem artykułu jest ustalenie aktualnego znaczenia terminu książka elektroniczna na podstawie jego definicji w źródłach encyklopedycznych i słownikowych.

**Koncepcja/metody badań:** W badaniach wykorzystano analizę pola semantycznego. Zebrany korpus definicji prezentowanych (w j. polskim i angielskim) przeanalizowano w poszukiwaniu kluczowych elementów deskrypcji i wskazano natężenie występowania poszczególnych cech (znaczeń).

**Wyniki/wnioski:** Artykuł prezentuje wyniki badania, sprawdzającego w jaki sposób książka elektroniczna rozumiana jest w źródłach informacyjnych, oraz która grupa cech formalnych lub abstrakcyjnych, odgrywa istotniejszą rolę w rozumieniu książki.

**Oryginalność/wartość poznawcza:** Językoznawcza teoria pola semantycznego zakładająca, że każde słowo współwystępuje z innymi w pewnych wyodrębnionych kategoriach znaczeniowych, może znaleźć zastosowanie w badaniach terminologicznych z zakresu informacji naukowej i bibliotekoznawstwa. Zaprezentowane badania przyczynią się do rozwoju terminologii z zakresu tych dyscyplin.

## Słowa kluczowe

Bibliologia. Informatologia. Książka elektroniczna. Pole semantyczne. Terminologia.

Otrzymano: 24 maja 2016. Zrecenzowano: 6 lipca 2016. Zaakceptowano: 22 lipca 2016.

---

## 1. Wprowadzenie

Książka elektroniczna stanowi interesujące zjawisko na tle dotychczasowej historii książki i czytelnictwa. Jednak dokumenty cyfrowe utraciły część cech uznawanych dotychczas przez badaczy za najważniejsze, przede wszystkim kodeksową formę wiążącą zadrukowane, papierowe arkusze. Książki elektroniczne wyróżnia głównie sposób utrwalenia i odczytu, materialnym nośnikiem dla treści staje się przestrzeń nośnika danych cyfrowych. Właściwy nośnik może często znajdować się poza urządzeniem, za pomocą którego wyświetlamy książkową treść, gdy korzystamy z dostępu metodą online. Środowisko cyfrowe sprzyja ponadto zmienności wizualizacji takich książek. Ich forma prezentacji może być różna, zależnie od cyfrowego formatu zapisu i urządzenia, na którym są wyświetlane. Ostateczna postać dokumentów wyświetlanych na ekranie jest również kwestią aktualnie dostępnych rozwiązań technologicznych. Jak się zaś wydaje, zasadniczą cechą książki tradycyjnej jest stabilność i niezmiennność. Tymczasem książka elektroniczna może być pewnym konstruktem pojęciowym, którego znaczenie rozciąga się na inne typy elektronicznych dokumentów. Zdaniem Małgorzaty Góralskiej (1996, 37) termin ten

stanowić może również pewien nadrzędny termin dla różnych publikacji cyfrowych, nawet jeśli dzieła te pod względem formy i realizacji typograficznej znacznie odbiegają od swych drukowanych odpowiedników.

Pytanie zasadnicze, które można postawić przy okazji podobnych rozważań brzmi: czy obiekty, które nazywamy książkami elektronicznymi, można zakwalifikować do kategorii książek w ogóle? Sebastian Kotuła (2013, 63, 64). wskazał, że

elektroniczna kopia książki drukowanej, także rękopiśmiennej, charakteryzuje się mniejszym stopniem natężenia cechy książkowości<sup>1</sup> niż oryginał, na podstawie którego powstała

oraz

należy wyraźnie podkreślić, iż doświadczenie książki oznacza de facto doświadczenie obcowania z czymś materialnym, tj. okładką, określonym kształtem, teksturą, wagą itp. itd. Są to podstawowe elementy i aspekty, które ulegają zatraceniu, gdy książka zostanie przetransformowana do postaci digitalnej. Ważnym elementem książki jest jej typograficzność wiążąca się z faktem, że dana książka została wydrukowana.

Tabela 1. Pytania o uczestnictwo w kulturze książki w ciągu 12 miesięcy poprzedzających badanie (reprezentatywna ogólnopolska próba 3049 respondentów w wieku co najmniej 15 lat, procentowe wartości danych)

| Czy w ciągu ostatnich 12 miesięcy...                                                                                                        | Nie | Tak, raz | Tak, co najmniej kilka razy |
|---------------------------------------------------------------------------------------------------------------------------------------------|-----|----------|-----------------------------|
| ...czytał(-a) Pan(-i) książkę w formacie e-booka?                                                                                           | 95  | 4        | 2                           |
| ...słuchał(-a) Pan(-i) jakiegoś audiobooka, czyli książki w pliku dźwiękowym, czytanej przez lektora, nagranej na płytę CD lub inny nośnik? | 95  | 4        | 2                           |
| ...ściągnął(-ęła) Pan(-i) jakąś książkę z internetu lub czytał(a) ją on-line?                                                               | 94  | 4        | 1                           |
| ...wypożyczył(-a) Pan(-i) jakąś książkę z biblioteki?                                                                                       | 84  | 7        | 9                           |
| ...czytał(-a) Pan(-i) w prasie lub w internecie jakiś artykuł, recenzję, wywiad z autorem albo inny wpis lub opinię o książce?              | 82  | 12       | 6                           |
| ...czytał(-a) Pan(-i) dłuższe teksty w formie elektronicznej (czyli z ekranu i bez użycia papieru)?                                         | 81  | 9        | 9                           |
| ...dostał(-a) Pan(-i) jakąś książkę w prezencie?                                                                                            | 81  | 15       | 5                           |
| ...pożyczył(-a) Pan(-i) komuś jakąś książkę?                                                                                                | 80  | 13       | 7                           |
| ...podarował(-a) Pan(-i) komuś książkę w prezencie?                                                                                         | 79  | 16       | 6                           |
| ...pożyczył(-a) Pan(-i) od kogoś jakąś książkę?                                                                                             | 77  | 15       | 8                           |
| ...polecił(-a) Pan(-i) komuś jakąś książkę?                                                                                                 | 76  | 16       | 9                           |
| ...rozmawiał(-a) Pan(-i) z kimś o książkach?                                                                                                | 68  | 18       | 14                          |

Źródło: (Michalak et al., 2016, 76)

<sup>1</sup> *Książkowość* to termin (odpowiednik angielskiego *bookness*) zaproponowany przez S. Kotulę, na oznaczenie cechy, jaką musi posiadać obiekt, aby można go było uznać za książkę.

Mamy zatem do wyboru dwie drogi. Można uznać, że termin „książka” powinno się zarezerwować wyłącznie dla drukowanych, tradycyjnych form, jako wcieleń idealnych. Można też przyjąć interpretację „otwartą” i uznać, że cyfrowa zmiana stanowi naturalny krok w ewolucji i jest kolejną ze znanych nam form książki, opowiadam się za tym właśnie, otwartym ujęciem rozumienia książki. Pamiętamy, że kodeksowa postać również wyłoniła się w wyniku procesów rozwoju i udoskonalania. Książki elektroniczne, cyfrowe, multimedialne, czy nawet wykorzystujące możliwości rzeczywistości rozszerzonej (zob. np. Gmiterek, 2015), można traktować właśnie jako kolejną postać książkowego dokumentu utrwalającego myśl ludzką. Należy jednakże przyznać, że właśnie na materialnym aspekcie książki nadal skupiają się definicje naukowe, encyklopedyczne, słownikowe (Pacek, 2015a; 2015b).

Uzupełnieniem powyższych rozważań dotyczących elektronicznej książki mogą być wyniki badania czytelnictwa zaprezentowane przez Instytut Książki i Czytelnictwa Biblioteki Narodowej w opracowaniu *Stan czytelnictwa w Polsce w 2015 roku*. Okazuje się, że 8% zbadanych respondentów posiada w swoich kolekcjach oprócz tradycyjnych, papierowych woluminów, również książki elektroniczne lub książki audio. Z badania obejmującego 2015 r. dowiadujemy się, że 6% respondentów czytało w ciągu minionego roku książkę w formacie elektronicznym (4% jeden raz, 2% kilkakrotnie). Identyczny wynik procentowy dało pytanie o słuchanie książek audio, 5% przyznało, że pobrało jakąś książkę z Internetu lub czytało ją online (4% jeden raz, 1% kilkakrotnie) (Michalak et al., 2016). Te dane ilustruje tabela 1.

## 2. Cel i metoda badań

Celem badań prezentowanych w niniejszym artykule było ustalenie, w jaki sposób rozumiany jest termin „książka elektroniczna”<sup>2</sup> w źródłach o charakterze informacyjnym (słowniki i encyklopedie, polsko- i angielskojęzyczne). Do realizacji tego celu zastosowano metodę analizy pola semantycznego. Teoria pola semantycznego zakłada, że język jest systemem składającym się z elementów – słów, połączonych z sobą siecią relacji semantycznych. W ten sposób tworzą one określone kategorie, inaczej grupy, lub mapy znaczeniowe. Dla potrzeb przeprowadzonego badania przyjęto rozumienie pola semantycznego zgodne z lingwistyczną definicją

jako uporządkowanej grupy wyrazów powiązanych przynależnością do wspólnej kategorii znaczeniowej (Tokarski, 2012, 362).

Podobne do tego ujęcia jest rozumienie zaprezentowane w *Słowniku encyklopedycznym informacji, języków i systemów informacyjno-wyszukiwawczych*, zgodnie z którym pole semantyczne to

zbiór wyrazów pozostających względem siebie w określonych relacjach semantycznych (Bojar, oprac. 2002, 197).

<sup>2</sup> W zbadanych źródłach, na oznaczenie tego typu książki, używane są terminy: e-book, e-książka, electronic book, elektroniczna książka, książka elektroniczna. W niniejszym artykule stosowany jest jednolicie termin książka elektroniczna.

Zastosowana w prezentowanym badaniu metoda zbliża się do założeń analizy pola semantycznego Régine Robin. Według tej autorki

tekst nie jest przezroczysty. Szukanie znaczenia tekstu, zdania, wyrazu, oznacza przeanalizowanie wszystkich jego zastosowań i kontekstów, wymaga ciągłej pracy nad tekstem, pozornego rozłożenia ciągu wypowiedzeniowego i porządku wypowiedzi po to, by złożyć je na powrót zgodnie z czytelnością znaczącą (Robin, 1980, 252).

Zgodnie z metodą R. Robin dokonuje się wyboru słów kluczowych z badanych tekstów, a następnie uporządkowuje zebrany materiał w ramach sześciu, wskazanych przez autorkę, kategorii: określenia podmiotu; jego asocjacji, opozycji, ekwiwalentów (synonimów), działań podmiotu, działań wobec podmiotu. Dla celów prezentowanego badania, zdecydowano się na zastosowanie odmiennego zestawu kategorii (zostaną one scharakteryzowane w trzeciej sekcji). Książka postrzegana jest tutaj z perspektywy bibliologii, bibliotekarstwa i informatologii. Współczesne standardy opisu dokumentów wypracowały własne propozycje kategorii porządkujących opis cech książki. Właśnie one zostały wykorzystane do uporządkowania pozyskanego słownictwa. Jak zauważyła Zina Jarmoszuk (2001, 19):

Poszczególne pola, ich zakres i wewnętrzne relacje między wyrażeniami odzwierciedlają rozumienie i widzenie świata charakterystyczne dla danej grupy użytkowników języka.

Należy przywołać także wiadomości o pracach nad językowym obrazem świata podejmowanych w środowisku językoznawców i etnolingwistów lubelskich, działających pod kierunkiem Jerzego Bartmińskiego. W efekcie tych działań publikowany jest *Słownik stereotypów i symboli ludowych*. Opracowanie ma na celu zbadanie formy i znaczenia języka ludowego oraz rolę stereotypii (topiki, formuł, idiomów) w tekstach folkloru, odtworzenie ludowego obrazu świata (Bartmiński, 1988b, 16). Cechą szczególną tego projektu jest zastosowanie nowatorskiego rodzaju definicji kognitywnej, która polega na zebraniu wszystkich cech ustabilizowanych w językowym obrazie danego przedmiotu (Bartmiński, 1988a, 177). Realizowane jest to poprzez zestawienie fragmentów tekstów źródłowych, zawierających odniesienia do badanych desygnatów rzeczywistości pozajęzykowej. Do uporządkowania zdań definicyjnych zawartych w tak zbudowanej eksplikacji, stosowane są kategorie semantyczne, określane również fasetami (pod wpływem prac Anny Wierzbickiej, a także prac z zakresu informacji naukowej, por. Bartmiński, 2012, 84).

Siatka owych faset pozwala na rekonstrukcję struktury poznawczej haseł, na oddanie ich podmiotowej konceptualizacji, wewnętrznej sieci relacji (UMCS, 2016).

Można więc uznać, że prezentowane w niniejszym artykule badania z zakresu pola semantycznego książki elektronicznej częściowo zbliżają się do metodyki zastosowanej do badań językowego obrazu świata. Głównym punktem styku jest operacja wyboru elementów reprezentatywnych dla desygnatu (fragmentów tekstów i w naszym przypadku słów kluczowych), zestawienie ich ze sobą i zastosowanie uporządkowania wyodrębnionych aspektów (cech, reprezentowanych tutaj przez słowa kluczowe) według pewnych kategorii znaczeniowych.

Badając sposoby definiowania terminów, warto również wyjaśnić relację pomiędzy terminem a pojęciem. Rozważania nad istotą pojęcia są dość częste w literaturze z zakresu informacji naukowej i bibliotekoznawstwa. Jak zauważyła Jadwiga Woźniak-Kasperek (2015, 177):

Pojęcia są jednostkami mentalnymi, łączącymi się w rodziny, grupy, kategorie. Żeby mogły stać się obiektem jakichkolwiek działań poza umysłem osoby, muszą zostać zeksternalizowane, na przykład w postaci terminu.

W odniesieniu do działalności terminologicznej Jacek Tomaszczyk (2014, 48) scharakteryzował dwa sposoby analizy terminologicznej przeprowadzanej na poziomie systemowym, który obejmuje badanie pojęć, tworzenie ich systemu, definiowanie pojęć, przyporządkowywanie im odpowiednich terminów. Analizę terminologiczną można przeprowadzać w drodze od pojęcia do terminu (podejście onomazjologiczne) oraz od terminu do pojęcia (podejście semazjologiczne). Jak wskazuje autor, podejście semazjologiczne stosowane jest najczęściej dla tworzenia słowników ogólnych (filologicznych) i polega na:

- (1) Wyborze dziedziny i analizie źródeł z jej zakresu, w celu identyfikacji i selekcji terminów.
- (2) Wstępnym określeniu znaczenia terminów; ustaleniu relacji semantycznych między nimi.
- (3) Sformułowaniu definicji oraz opcjonalnie, na wyborze ekwiwalentów obcojęzycznych.

Podejście onomazjologiczne również rozpoczyna się od wyboru dziedziny, natomiast kolejnym etapem jest analiza obiektów w celu ustalenia ich właściwości, których podzbiory, odzwierciedlające naturę badanych obiektów (tzw. cechy wystarczające i konieczne) zostaną wykorzystane do opisu pojęć tych obiektów, czyli do procesu konceptualizacji. Kolejne kroki to przyznanie pojęciom nazw roboczych, czyli preterminów, systematyzacja i definiowanie pojęć, następnie przyporządkowanie ostatecznych nazw, czyli terminów, oraz ewentualnie odpowiedników w obcych językach. W tak zarysowanym kontekście metoda badawcza pola semantycznego wydaje się łączyć elementy obydwu dróg analizy terminologicznej. Polega bowiem, zgodnie z ujęciem semazjologicznym, na analizie źródeł i wyborze terminów, ustaleniu relacji semantycznych i na ich podstawie kwalifikowaniu do wspólnych kategorii znaczeniowych. Ta operacja pełni jednocześnie zadania wyznaczone przez podejście onomazjologiczne, ma na celu odzwierciedlić złożoność natury badanego obiektu i ostatecznie prowadzić do konceptualizacji pojęcia. Jacek Tomaszczyk stwierdza, że metoda onomazjologiczna jest w praktyce trudna do realizacji, gdyż wymaga zbudowania schematu struktury badanego pojęcia (można jednak uznać, że pole semantyczne jest właśnie takim schematem) i te etapy konceptualizacyjne muszą poprzedzać dopasowanie właściwych terminów. Tymczasem trudno taki proces zrealizować bez stosowania terminów (przynajmniej w ich postaci roboczej – preterminów), bez których komunikacja, zwłaszcza wykorzystująca język naukowy, jest w zasadzie niemożliwa:

Jak w sposób systematyczny wybierać obiekty do analizy, aby zapewnić kompletność i spójność zbioru pojęć i terminów? Wykorzystanie w tym celu literatury przedmiotowej, najważniejszego źródła w pracach leksykograficznych, wydaje się być niezgodne z ideą tego podejścia, ze względu na fakt, iż w tekstach nie występują pojęcia, a reprezentujące je terminy. Nawet najlepszy specjalista dziedziny nie zna wszystkich pojęć wchodzących w obszar pojęciowy swojej dziedziny, więc wykorzystanie literatury naukowej i fachowej oraz innych źródeł leksykalnych wydaje się niezbędne. Teoretycznie mogłaby rozwiązać ten problem współpraca wielu specjalistów, ale nie wiadomo, w jaki sposób mieliby oni bez używania terminów komunikować rezydujące w ich umysłach pojęcia. Czy w ogóle możliwe jest identyfikowanie pojęć i opisywanie ich cech bez używania nazw (terminów) pojęć i nazw cech, które również mogą być terminami? (Tomaszczyk, 2014, 49).

Jak również zauważa ten autor, pojęcie jest pierwotne w stosunku do terminu, powstaje wcześniej w umyśle, następnie otrzymuje nazwę. Jednak w rzeczywistości komunikacyjnej niemożliwe jest opisywanie pojęć bez zastosowania odnoszących się do niego terminów. Wydaje się zatem, że można w tym miejscu zaproponować właśnie wykorzystanie techniki badania pola semantycznego do działalności terminologicznej, uwzględniając indeksowanie i grupowanie słownictwa w celu odtworzenia siatki kategorii semantycznych w strukturze pojęcia.

Na potrzeby badania zebrano korpus 23 definicji<sup>3</sup>, których źródła sytuują się w obszarze problemowym bibliologii i informatologii oraz informatyki. Kilka z nich ma charakter ogólnych opracowań encyklopedycznych. Wykorzystane publikacje pochodzą z ostatnich 20 lat. Jest to okres kiedy problematyka związana z książką elektroniczną zajęła już stałe miejsce w dyskursie naukowym. Zebrany materiał został poddany indeksowaniu swobodnemu, w wyniku czego wybrano 57 słów kluczowych, reprezentujących zawartość definicji. Niektóre słowa kluczowe zostały utworzone doraźnie, choć nie pojawiły się bezpośrednio w tekstach definicji, lecz najtrafniej wyrażają treść analizowanego fragmentu<sup>4</sup>. Następnie słowa te zostały pogrupowane w ramach utworzonych kategorii, w których porządek słów wyznaczony został częstotliwością ich występowania w definicjach. W efekcie uzyskano obraz znaczenia książki elektronicznej i wiedzę o tym, jakie jej cechy są uznawane za najistotniejsze. Swoisty obraz, mapa lub pole semantyczne książki elektronicznej rysuje się tu dzięki obserwacji, które z cech są przywoływane w definicjach najczęściej i które są najliczniej reprezentowane za pomocą repertuaru zastosowanych słów kluczowych. Przedstawiane badanie można również zaklasyfikować do obszarów problemowych dziedzin infolingwistyki oraz bibliolingwistyki (zob. Migoń & Skalska-Zlat, 1995).

### 3. Analiza słownictwa

Wiedzę o znaczeniu terminu książka elektroniczna uzyskano poprzez wykonanie specyficznych operacji na zebranych słownictwie. Zostało ono podzielone na sześć grup, z których pięć odnosi się do fizycznych cech książki elektronicznej oraz związanych z jej obiegiem wydawniczo-księgarskim. Ze względu na wspólną kategorię cech książki wskazywanych przez słowa kluczowe grupy te nazywane są dalej wspólnie „Formalnymi”. Szósta grupa zawiera wyrażenia związane z obszarem, który można nazwać przestrzenią społeczną lub informacyjną książki (Krawczyk, 2009, 16), dlatego określona jest jako „Abstrakcyjna”. W tym miejscu istotne jest także wyjaśnienie, czym charakteryzują się dwa typy informacji statystycznych uzyskanych w badaniu. Pierwszy typ to liczebność, wskazująca jak duża jest grupa słów kluczowych odnoszących się do danej cechy książki elektronicznej, a więc ile słów udało się zakwalifikować do określonej grupy (kategorii). Tego rodzaju statystyka pozwala wnioskować jak szeroki jest repertuar słów (nazw) odnoszących się do danej kategorii cech książki elektronicznej, a więc w jakim stopniu dany obszar został

<sup>3</sup> W jednym przypadku z jednego źródła zaczerpnięto dwie definicje, stąd 22 a nie 23 pozycje w wykazie źródeł na końcu artykułu.

<sup>4</sup> Zgodnie z metodą indeksowania swobodnego za pomocą języka swobodnych słów kluczowych, „[...] elementarnymi jednostkami leksykalnymi są słowa kluczowe pochodzące z testu indeksowanego dokumentu lub wybrane przez indeksatora z zasobu jego słownika czynnego [...]” (Bojar, oprac. 2002, 111).

rozpoznany i stanowi przedmiot badań wykorzystywany do opisu książki elektronicznej. Wyższa liczebność danej grupy, w porównaniu z innymi, jest efektem szerszego rozpoznania określonej kategorii cech książki elektronicznej i prawdopodobnie wyższej transparentności tych cech i ich reprezentatywności dla danej kategorii. Drugi typ danych to częstotliwość słów, przejawiająca się liczbą przywołań określonych cech książki elektronicznej za pomocą słów kluczowych w definicjach. Indeksowane były pojedyncze wskazania wybranej cechy w danej definicji, a nie każde wystąpienie określonego słowa/znaczenia. W przeciwnym przypadku zostałaby utworzona lista frekwencyjna występowania danych słów, tymczasem celem było ustalenie ile ze zbadanych definicji przywołuje określone cechy książki elektronicznej. Częstotliwość danego słowa oznacza ile definicji (spośród zebranych 23) odniosło się do danej cechy wyrażonej za pomocą słowa kluczowego. Im wyższa częstotliwość tym wyższa „popularność” danej cechy, wyższa jej ranga w rozumieniu terminu książka elektroniczna. Scharakteryzowane dwa typy danych można rozpatrywać oddzielnie oraz komplementarnie. Postrzegane razem tworzą pole semantyczne książki elektronicznej obecne w źródłach encyklopedycznych i słownikowych.

Cechy formalne książki elektronicznej wskazywane w definicjach uszeregowano w ramach pięciu podgrup. Odzwierciedlają one kategorie cech wskazywane również w bibliotekarskich opracowaniach standaryzacyjnych zajmujących się identyfikacją i opisem książek. Wzorowane są na terminologii wykorzystanej w zasadach katalogowania *International Standard Bibliographic Description* (ISBD), *Resource Description and Access* (RDA) oraz dokumencie specyfikującym format danych bibliotecznych *MARC 21 Format for Bibliographic Data*. W ten sposób zebrane słownictwo zakwalifikowano do następujących podgrup: „Budowa nośnika”, „Typ nośnika”, „Typ zawartości”, „Tryb odczytu”. Ponadto na potrzeby właściwej kategoryzacji słownictwa związanego z zagadnieniami obiegu książki na rynku wydawniczo-księgarskim oraz biblioteczno-czytelniczym utworzono podgrupę „Obieg”. Dla aspektów abstrakcyjnych utworzono tylko jedną grupę. Podgrupa „Budowa nośnika” gromadzi słowa związane ze strukturą książki. Słowa kluczowe w podgrupie „Typ nośnika” odzwierciedlają fizyczny format, który przybiera książka; podzbiór „Tryb odczytu” zawiera słowa związane z cechami książki, warunkującymi technikę jej odbioru; „Typ zawartości” wskazuje na formę wyrazu, w której zawartość książki jest komunikowana.

Jak podano wyżej, kategoryzacji poddano 57 słów kluczowych charakteryzujących treść definicji. Ich sumaryczna częstotliwość występowania (dokładniej, częstotliwość powołań na określone cechy książki w definicjach) wyniosła 156 (Tab. 2).

Tabela 2. Porównanie liczebności i częstotliwości słów w podgrupach

|                           | Pole semantyczne książki elektronicznej |                      |
|---------------------------|-----------------------------------------|----------------------|
|                           | Grupa „Formalna”                        | Grupa „Abstrakcyjna” |
| Liczba podgrup            | 5                                       | 1                    |
| Łączna liczba słów        | 56                                      | 1                    |
| Łączna częstotliwość słów | 155                                     | 1                    |

#### 4. Charakterystyka podgrup

Podgrupa „Tryb odczytu” pojawia się w poniższym zestawieniu na początku, ponieważ pod względem liczebności zajmuje pierwsze miejsce *ex aequo* z podgrupą „Typ zawartości”. Każda z nich zgromadziła 17 słów. Podgrupa „Tryb odczytu” uzyskała natomiast najwyższą częstotliwość (64). W tej podgrupie znalazły się następujące słowa (liczba w nawiasie oznacza częstotliwość, czyli liczbę definicji, w których słowo się pojawiło): odczyt za pomocą komputera (14), odczyt za pomocą czytnika książek (8), odczyt na urządzeniu przenośnym (7), wyświetlana (6), odczyt za pomocą oprogramowania (5), odczyt za pomocą urządzenia elektronicznego/cyfrowego (5), odczyt za pomocą telefonu (3), zaznaczanie (3), adnotowanie (2), odczyt za pomocą tabletu (2), podświetlanie (2), wymaga uniwersalnego formatu zapisu (2), czytelna maszynowo (1), na wyświetlaczu podobna do drukowanej (1), odczyt na różnych platformach sprzętowych (1), odczyt za pomocą palmtopa (1), przeszukiwanie (1). Specyfika tej podgrupy polegająca na zgromadzeniu dużej liczby słów, również o wysokiej częstotliwości przywoływania w definicjach, dowodzi, że cechą różnicującą, najbardziej charakterystyczną dla książki elektronicznej, najczęściej wskazywaną w definicjach, jest oryginalny sposób odczytu, polegający na zastosowaniu do tej czynności komputera, lub innych typów urządzeń elektronicznych. „Tryb odczytu” to zdecydowanie najistotniejsza, różnicująca kategoria cech książki tego typu.

Na miejscu drugim znalazła się podgrupa „Typ zawartości”. Jak napisano wyżej, choć zawiera ona również 17 słów kluczowych, jej łączna częstotliwość wyniosła 43. Na tę podgrupę składają się następujące słowa: elektroniczna wersja drukowanej (7), cyfrowa wersja drukowanej (6), tekst (6), dokument elektroniczny (4), efekt konwersji druku do formatu cyfrowego (3), born-digital (2), multimedialna (2), obraz (2), skan (2), tekst cyfrowy (2), animowana (1), atrament elektroniczny (1), dźwięk (1), interaktywna (1), plik tekstowy (1), nie druk (1), tworzenie za pomocą procesora tekstu (1). Identyfikowalny typ zawartości książki elektronicznej pozostaje pod wpływem środowiska cyfrowego, w którym funkcjonuje książka elektroniczna. Jak można zauważyć jest to głównie tekst, ale zapisany elektronicznie/cyfrowo. Zawartość jest wzbogacana również obrazem, dźwiękiem, animacjami. Rozszerza się więc tu oferta wykorzystywanych kodów, czyniąc ten typ książki, w odróżnieniu od książki tradycyjnej, multimedialną.

Podgrupa „Obieg” jest trzecią pod względem liczebności. Składa się na nią 13 słów kluczowych i uzyskała łączną częstotliwość 34. Tworzą ją słowa kluczowe: pobierana (8), internetowa (5), obsługiwana przez biblioteki (3), w obrocie na rynku książki (3), wymaga systemu dostarczania (3), elektroniczna dystrybucja (2), chroniona prawem (2), ładowalna (2), publikacja (2), darmowa (1), dodatek do książki papierowej (1), efektywna w przechowywaniu i transporcie (1), kopiowana (1). W granicach procesów związanych z obiegiem zauważalna jest również wyraźna zmiana w stosunku do obiegu książki tradycyjnej. W przypadku książki elektronicznej podstawową metodą uzyskania do niej dostępu jest pobieranie, ładowanie, dystrybucja w środowisku elektronicznym. Istotne jest podkreślenie, że taka postać książki została zaakceptowana przez instytucje biblioteczne, jest ona przez nie obsługiwana, stanowi także przedmiot ochrony prawnej.

Kolejna podgrupa „Budowa nośnika” zawiera pięć słów kluczowych, zaś jej łączna częstotliwość wynosi sześć. Znalazły się tutaj następujące słowa: wielkość podobna do książki tradycyjnej (2), posiada elementy sterujące (1), mniejsza waga i objętość niż drukowanej (1), mniejsza od książki tradycyjnej (1), lekka (1). Statystyki tej grupy dowodzą, że trudno jest odnosić się

do budowy i struktury książki elektronicznej, która sama w sobie jest zmienna, tzn. może uzyskiwać różną formę i nośniki. Trudno też w definicji książki elektronicznej podejmować czysto informatyczne zagadnienia struktury plików, aplikacji lub urządzeń elektronicznych.

Ostatnia z tego obszaru podgrupa „Typ nośnika” zawiera już tylko cztery słowa kluczowe, a jej łączna częstotliwość to osiem. Znalazły się tu następujące słowa: czytnik (5), nowa generacja książki (1), nie papier (1), dysk optyczny (1). Może nieco zaskakiwać nieliczna reprezentacja słów tej grupy. Wydawać się mogło, że odniesienia do sposobu utrwalenia treści zawartych w elektronicznej książce, „rodzaj materii” wykorzystanej do stworzenia takiego obiektu, będzie również istotną sferą podkreślaną i różnicującą taki typ książki. Zauważalne jest jednak dość częste utożsamianie książki elektronicznej z urządzeniem – czytnikiem cyfrowych treści. Częstotliwość tego słowa zaważyła o wyższej łącznej częstotliwości podgrupy „Typ nośnika” od „Budowy nośnika”, pomimo mniejszej liczby słów.

Oddzielna grupa utworzona została w celu zaszeregowania słów, które można uznać za abstrakcyjne. W istocie ta sfera jest reprezentowana w uzyskanym zbiorze tylko przez jedno słowo kluczowe odnoszące się do „Zawartości książki”. Jest to słowo „dzieło”, które pojawiło się tylko jeden raz w badanych definicjach i uzyskało częstotliwość 1, czyli tylko jedna z definicji wykorzystała je do wyjaśnienia istoty książki. Początkowa interpretacja niskich wyników statystycznych tej grupy podpowiada, że znaczenie abstrakcyjnych sfer zawartości książki jest niewielkie. Wyrażone jest zaledwie jednym słowem, przywołanym w jednej na 23 definicje. Wiemy dzięki temu, że cechy abstrakcyjne nie są najłatwiejszymi do zdefiniowania, opisanie i wykorzystania dla eksplikacji książki elektronicznej. Jednakże sytuacja szczególnie słabego odzwierciedlenia cech abstrakcyjnych na mapie semantycznej książki elektronicznej, może mieć jeszcze inne, również interesujące uzasadnienie, które wymaga odpowiedniego dochodzenia. Będzie o tym mowa w kolejnym rozdziale artykułu prezentującym wnioski z przeprowadzonej analizy.

## 5. Wnioski

Przedstawione w artykule wyniki badań pozwalają na wyciągnięcie konkluzji dotyczących jednostkowego typu książki, jaką jest jej elektroniczna forma, ale też w zestawieniu z wcześniejszymi ustaleniami, dowodzą prawdziwości pewnych tez odnoszących się do znaczenia terminu „książka” jako takiego.

Tabela 3. Kategorie formalne pola semantycznego książki elektronicznej  
(kolejność według częstotliwości)

| Kategoria      | Pole semantyczne książki elektronicznej.<br>Kategorie formalne |            |
|----------------|----------------------------------------------------------------|------------|
|                | Częstotliwość                                                  | Liczebność |
| Tryb odczytu   | 64                                                             | 17         |
| Typ zawartości | 43                                                             | 17         |
| Obieg          | 34                                                             | 13         |
| Typ nośnika    | 8                                                              | 4          |
| Budowa nośnika | 6                                                              | 5          |

Z zestawienia kategorii cech książki elektronicznej wynika, że najważniejszą cechą różnicującą, wskazywaną przez twórców definicji, jest specyficzny sposób odczytu książki elektronicznej, który wymaga zastosowania urządzenia elektronicznego. To właśnie stanowi jądro znaczenia terminu książka elektroniczna. Inne cechy, szczególnie takie jak abstrakcyjna zawartość, nie stanowią obszaru wyróżniającego w badanym polu semantycznym. Zawartość części abstrakcyjnej nie skupia cech zmiennych, charakterystycznych dla określonego typu książki. Statystyki cech formalnych książki elektronicznej zebrano w tabeli 3.

Mamy zatem do czynienia z dość niezwykłą sytuacją. Z jednej strony obfitość terminów odnoszących się do fizycznej i utylitarnej funkcjonalności książki i częstotliwość ich przywoływania mogą dowodzić istotnego znaczenia dla wyodrębnienia książki elektronicznej, jako obiektu informacyjnego określonego typu. Z drugiej zaś strony cechy skupione w tym obszarze pola semantycznego są jednocześnie najbardziej zmienne na szerszej mapie znaczenia książki i różnią się w zależności od typu książki, z którą mamy do czynienia. Dla uzyskania pełnej podstawy formułowanych spostrzeżeń należy przywołać wyniki wcześniejszych badań podobnego typu (Pacek, 2015a; 2015b) wykorzystujących teorię pola semantycznego dla ustalenia rozumienia książki jako takiej (w źródłach encyklopedycznych i słownikowych z zakresu bibliologii, informatologii i bibliotekoznawstwa) oraz znaczenia książki w piśmiennictwie o charakterze normalizacyjnym. Wyniki tych badań wykazały zdecydowaną dominację grup formalnych w rozumieniu istoty książki, choć w obydwu przypadkach obszarem pola semantycznego najliczniej reprezentowanym w definicjach i jednocześnie uzyskującym najwyższą częstotliwość była „Budowa nośnika”. W przypadku analizowanej tutaj książki elektronicznej zauważalny jest spadek znaczenia tej podgrupy pola semantycznego na rzecz „Trybu odczytu”. Świadczy to oczywiście o tym, że najistotniejszą różnicą pomiędzy zbadanymi rodzajami książek jest właśnie sposób komunikowania i odbioru ich treści<sup>5</sup>.

Prawdziwa wydaje się zatem teza, że wbrew pozorom, obszar cech fizycznych nie jest najistotniejszy dla zrozumienia czym jest książka jako taka. Cechy te nie są stałe, nie przenoszą wiedzy pozwalającej na odkrycie tego, co jest elementem wspólnym dla książki bez względu na jej typ. Wobec tego należy zadać pytanie: gdzie jest ukryta cecha/zespół cech stanowiących o idei książki, pozwalających odróżnić książkę jako taką od innych obiektów informacyjnych, bez względu na to, jaką jej „odmianę” będziemy w danej chwili analizowali? Odpowiedzią na to pytanie będzie zarazem kolejna teza: wspólną, najważniejszą cechą książki (w ujęciu naukowym, definicyjnym) jest jej abstrakcyjna zawartość. Jest to zbiór konotacji wywołujących u odbiorcy książki wyobrażenie wiedzy, myśli, komunikatów i pojęć. Grupa słów związanych z abstrakcyjną zawartością, choć stosunkowo nieliczna, uzyskała we wcześniejszych badaniach wysoką łączną częstotliwość. Była często przywoływana w definicjach, co świadczy o wysokim uznaniu dla tego składnika znaczenia książki (zob. Pacek, 2015a). Dla aspektów abstrakcyjnych wspólnym określeniem pojawiającym się w dotychczasowych badaniach pola semantycznego książki jest słowo „dzieło”.

Termin „dzieło” jest wieloznaczny i również nie poddaje się łatwo ujęciom definicyjnym. Bywa różnie rozumiany, w zależności od kręgu kulturowego, obszaru nauki czy kultury,

<sup>5</sup> Jednocześnie warto przywołać tutaj obserwację, zgodnie z którą w ogólnych definicjach książki, w bardzo niewielkim stopniu nastąpiła recepcja książki elektronicznej (w 24 definicjach encyklopedycznych i słownikowych ten typ przywołano zaledwie czterokrotnie, zaś w siedmiu definicjach obecnych w piśmiennictwie normalizacyjnym zaledwie jednokrotnie).

w którym jest stosowany. Dokument porządkujący jednostki i relacje uniwersum bibliograficznego – *Functional Requirements for Bibliographic Records* (FRBR) wprowadza szereg nowych dla bibliotekarstwa jednostek opisu i umieszcza dzieło (ang. *work*) jako jedną z jednostek naczelných, lecz przyznaje jej dość niejednoznaczne rozumienie, zgodnie z którym nie istnieje żaden materialny obiekt, który stanowiłby dzieło. Rozpoznawane i postrzegane jest jedynie poprzez swe indywidualne realizacje (czyli formę intelektualną lub artystyczną jaką przyjmuje dzieło za każdym razem kiedy jest wykonywane). Z kolei standard katalogowania RDA definiuje dzieło jako odrębny utwór intelektualny lub artystyczny (tzn. o intelektualnej lub artystycznej zawartości). W normie terminologicznej PN-ISO 5127:2005 *Informacja i dokumentacja – Terminologia* zamieszczono również dość enigmatyczną definicję dzieła:

Zawartość dokumentu będąca wynikiem działalności twórczej, dzięki której podlega on ochronie copyrightu lub tytułu własności przemysłowej ze względu na indywidualny charakter lub nowość (PN-ISO, 2005, 150).

Pozornie nieco bardziej precyzyjna jest definicja zamieszczona w *Nowym słowniku języka polskiego*, w którym dzieło jest rozumiane jako

utwór literacki, naukowy, muzyczny, artystyczny, zwykle dużej wartości (Sobol, red. 2002, 172).

Jednak podobnie jak w przypadku zawartości, aby zrozumieć dzieło trzeba dookreślić rozumienie utworu. Uwidacznia się zatem trudność definiowania zjawisk abstrakcyjnych, choć ich znaczenie dla zrozumienia także materialnych obiektów środowiska informacyjnego i kulturowego jest bardzo istotne. Ten fakt stanowi barierę w badaniu fenomenu książki. Analizy pozwalają dostrzec, że to właśnie sfera abstrakcyjna ma prawdopodobnie najistotniejsze znaczenie w materiale definicyjnym dla rozumienia istoty książki jako takiej. Ta sfera znaczenia książki jest trudna do badań, a przez to pozostaje słabo rozpoznana.

## Bibliografia

- Bartmiński, J. (1988a). Definicja kognitywna jako narzędzie opisu konotacji słowa. W: J. Bartmiński (red.) *Konotacja*. Lublin: Wydaw. UMCS, 169–183.
- Bartmiński, J. (1988b). Słownik ludowych stereotypów językowych. Założenia ogólne. *Etnolingwistyka* 1, 11–34.
- Bartmiński, J. (2012). *Punkt widzenia, perspektywa, językowy obraz świata*. Lublin: Wydaw. UMCS, 76–88.
- Bojar, B. oprac. (2002). *Słownik encyklopedyczny informacji, języków i systemów informacyjno-wyszukiwawczych*. Warszawa: Wydaw. SBP.
- Gmiterek, G. (2015). Rzeczywistość rozszerzona a książka i prasa. W: G. Gmiterek, M. Ochmański, M. Roszkowski (red.) *LaTel. Z badań nad wykorzystaniem technologii informacyjnych w bibliologii i informatologii*. Warszawa: Wydaw. SBP, 41–66.
- Górska, M. (1996). Elektroniczne książki. *Zagadnienia Informacji Naukowej*, 2, 33–40.
- Jarmoszuk, Z. (2001). Pole semantyczne terminu kultura w definicjach kultury. *Zagadnienia Informacji Naukowej*, 1, 17–34.
- Kotula, S. (2013). *Komunikacja bibliologiczna wobec World Wide Web*. Lublin: Wydaw. UMCS, 2013.
- Krawczyk, A. (2009). Przestrzeń jako kategoria poznawcza przy wyjaśnianiu funkcjonowania książki. W: J. Konieczna, S. Kurek-Kokocińska, H. Tadeusiewicz (red.) *Przestrzeń informacyjna książki*. Łódź: Wydaw. Biblioteka, 15–26.

- Michalak, D.; Koryś, I.; Kopeć, J. (2016). *Stan czytelnictwa w Polsce w 2015 roku* [online]. Biblioteka Narodowa, [18.04.2016], <http://bn.org.pl/download/document/1459845698.pdf>
- Migoń, K.; Skalska-Zlat, M. (1995). Bibliolinguistics and bibliometrics – connections and interdependence: theoretical and quantitative analysis. In: M.E.D. Koenig, A. Bookstein (eds.) *Fifth International Conference of the International Society for Scientometrics and Informetrics: Proceedings – 1995: June 7–10 1995*. Medford: Learned Information, 331–347.
- Pacek, J. (2015a). Pole semantyczne książki. *Przegląd Biblioteczny*, 4, 546–562.
- Pacek, J. (2015b). Pole semantyczne książki w tekstach normalizacyjnych i prawnych. *Praktyka i Teoria Informacji Naukowej i Technicznej* 2, [w druku].
- PN-ISO (2005). PN-ISO 5127. *Informacja i dokumentacja. Terminologia*. Warszawa: PKN.
- Robin, R. (1980). Badanie pól semantycznych. Doświadczenia Ośrodka Leksykologii Politycznej w Saint-Cloud. W: M. Głowiński (red.) *Język i społeczeństwo*. Warszawa: Czytelnik, 205–282.
- Sobol E. red. (2002). *Nowy słownik języka polskiego*. Warszawa: Wydaw. Naukowe PWN.
- Tokarski, R. (2012). Słownictwo jako interpretacja świata. W: J. Bartmiński (red.) *Współczesny język polski*. Lublin: Wydaw. UMCS.
- Tomaszczyk, J. (2014). *Model systemu informacji terminologicznej*. Katowice: Wydaw. UŚ.
- UMCS (2016). *Słownik stereotypów i symboli ludowych* [online]. Uniwersytet Marii Curie-Skłodowskiej, [18.04.2016], <http://www.umcs.pl/pl/sownik-stereotypow-i-symboli-ludowych,5561.htm#page-2>
- Woźniak-Kasperek, J. (2015). Terminologia jako mapa świata. W: G. Gmiterek, M. Ochmański, M. Roszkowski (red.) *LaTeL. Z badań nad wykorzystaniem technologii informacyjnych w bibliologii i informatologii*. Warszawa: Wydaw. SBP, 174–185.

### Wykaz źródeł definicji

- Adamczewski, P. (2005). *Słownik informatyczny*. Warszawa: Helion.
- ALA (2016). *Instruction for Diverse Populations Multilingual Glossary. Definitions* [online]. ALA American Library Association [29.03.2016], <http://www.ala.org/acrl/sites/ala.org.acrl/files/content/aboutacrl/directoryofleadership/sections/is/iswebsite/projpubs/idpdefinitions.pdf>
- Bangia, R. (2010). *Dictionary of Information Technology. Second Edition*. New Delhi: Laxami Publications, Ltd.
- Cooper, M. ed. (2005). *Illustrated Dictionary of Library and Information Science*. New Delhi: Lotus Press.
- Czapnik, G.; Gruszka Z. oprac. (2011). *Podręczny słownik bibliotekarza*. Warszawa: Stowarzyszenie Bibliotekarzy polskich, Uniwersytet Łódzki.
- Drake, M. red. (2003). *Encyclopedia of Library and Information Science. Second Edition*. Vol. 2. New York, Basel: Merce Dekker, Inc., 803–1590.
- Encyclopaedia Britannica (2016). *Electronic book* [online]. Encyclopaedia Britannica [29.03.2016], <http://www.britannica.com/technology/e-book>
- Henderson, H. (2009). *Encyclopedia of Computer Science and Technology*. New York: Facts On File.
- Kubiak, M. J. red. (1999). *Słownik technologii informacyjnej*. Warszawa: Mikom.
- Levine-Clark, M.; Carter, T. M. eds. (2013). *ALA Glossary of Library and Information Science. Fourth Edition*. Chicago: American Library Association.
- Oxford Dictionaries (2016). *E-book* [online]. Oxford Dictionaries. Language matters [29.03.2016], <http://www.oxforddictionaries.com/definition/english/e-book>
- Oxford Dictionaries (2016). *E-book* [online]. Oxford Advanced Learner's Dictionary [29.03.2016], <http://www.oxforddictionaries.com/definition/learner/e-book>
- PCMag (2016). *E-book* [online]. PC Magazine. Technology Product Reviews, News, Prices. Encyclopedia [29.03.2016], <http://www.pcmag.com/encyclopedia/term/42214/e-book>
- Prytherch, R. J. ed. (2012). *Harrod's Librarians' Glossary and Reference Book: A Directory of Over 10,200 Terms, Organizations, Projects and Acronyms in the Areas of Information Management, Library Science, Publishing and Archive Management. Tenth Edition, Revised*. Burlington: Ashgate Publishing, Ltd.

- PWN (2016). *E-book* [online]. Słownik języka polskiego PWN [29.03.2016], <http://sjp.pwn.pl>
- Quigley, M. ed. (2008). *Encyclopedia of Information Ethics and Security*. New York: Information Science Reference, cop.
- Reitz, J. M. (2016). *ODLIS Online Dictionary for Library and Information Science* [online]. [29.03.2016], [http://www.abc-clio.com/ODLIS/odlis\\_A.aspx](http://www.abc-clio.com/ODLIS/odlis_A.aspx)
- Suraj V. K. ed. (2005). *Encyclopaedic dictionary of library and information science*. Vol. 1. Delhi: Isha Books.
- Szaniawski, J. red. (1997). *Duży słownik informatyczny*. Warszawa: ArsKom.
- Tomei, L. A. ed. (2008). *Encyclopedia of information technology curriculum integration*. Hershey, New York: IGI Global.
- Wikipedia, English (2016). *E-book* [online]. Wikipedia. The Free Encyclopedia [29.03.2016], <https://en.wikipedia.org/wiki/E-book>
- Wikipedia, polska (2016). *E-book* [online]. Wikipedia. Wolna encyklopedia [29.03.2016], <https://pl.wikipedia.org/wiki/E-book>

## Semantic Field of the Term “electronic book”

### Abstract

**Purpose/Thesis:** The purpose of this paper is to study the current meaning of the term “electronic book” based on selected contemporary encyclopaedias and dictionaries.

**Approach/Research methods:** The study is conducted with semantic field analysis. Definitions were collected from encyclopaedias and dictionaries (in Polish and English) and then studied in search for key elements of descriptions and frequency of particular feature occurrence.

**Results and conclusions:** The paper presents the results of the analysis of collected definitions and aims to investigate the meaning of the term “electronic book” based on two groups of extracted features – formal and abstract ones.

**Originality/Value:** The results of the research may be used in terminology studies in the field of Library and Information Science. The research will contribute to the development of terminology in the field of information science and book studies.

### Keywords

Book studies. Information science. Electronic books. Semantic field. Terminology.

*Dr JAROSŁAW PACEK jest kierownikiem Instytutu Bibliograficznego Biblioteki Narodowej. Publikuje artykuły naukowe z zakresu problemów bibliografii w środowisku cyfrowo-sieciowym, wpływu Internetu na organizację działalności informacyjnej, zmian nośników informacji, terminologii nauki o informacji. Najważniejsze publikacje: Bibliografia w zmieniającym się środowisku informacyjnym, Warszawa, 2010; Bibliografia w XXI wieku. W: Bibliologia i informatologia, Warszawa 2011, s. 267–279; Pole semantyczne książki. Przegląd Biblioteczny 2015, nr 4, s. 546–562. Jest twórcą i redaktorem serwisów sieciowych związanych z działalnością informacyjną (m.in. Sternik. Słownik terminologiczny z zakresu bibliografii i katalogowania – <http://sternik.bn.org.pl/>). Jest członkiem Rady Naukowej Biblioteki Narodowej. Należy do Polskiego Towarzystwa Informacji Naukowej oraz do International Society for Knowledge Organization.*

*Kontakt z autorem:*

*j.pacek@bn.org.pl*

*Instytut Bibliograficzny*

*Biblioteka Narodowa*

*al. Niepodległości 213*

*02–086 Warszawa*

# Sieciowe sposoby społecznościowego zarządzania informacją o dokumentach

Grzegorz Gmiterek

*Katedra Bibliografii i Dokumentacji,  
Wydział Dziennikarstwa, Informacji i Bibliologii, Uniwersytet Warszawski*

---

## Abstrakt

**Cel/teza:** Celem artykułu jest przedstawienie zjawisk i przykładów powiązanych z nimi narzędzi organizacji informacji o dokumentach i zasobach cyfrowych (zwłaszcza sieciowych), głównych funkcji tych narzędzi, jak również podstawowych kierunków i tendencji ich rozwoju. Pod uwagę wzięto zarówno przykłady społecznościowego katalogowania i współtworzenia metadanych przez użytkowników, jak i przedsięwzięcia odnoszące się do informacyjnej działalności bibliotek, w ramach której oferowane są dodatkowe metody porządkowania informacji o dokumentach oraz podejmowane próby kontroli współtworzonych przez internautów metadanych.

**Koncepcja/metody badań:** Artykuł opracowany został na podstawie analizy i krytyki literatury przedmiotu. Posłużono się również wynikami obserwacji i analizy poszczególnych zjawisk oraz możliwości wykorzystania wybranych narzędzi (oprogramowania desktopowego, aplikacji mobilnych oraz usług sieciowych) do porządkowania danych o zasobach cyfrowych.

**Wyniki i wnioski:** Od kilkunastu lat widoczna jest dynamiczna ewolucja różnego typu sieciowych narzędzi społecznościowych, które ułatwiają zarządzanie informacją o dokumentach, jej gromadzenie i udostępnianie. Serwisy i usługi, które z różnych powodów tracą popularność i znikają są zastępowane nowymi, bardziej interaktywnymi i przyjaznymi dla użytkowników. Mamy też coraz częściej do czynienia z możliwościami wykorzystania w tym zakresie urządzeń i aplikacji mobilnych.

**Oryginalność/wartość poznawcza:** W artykule syntetycznie przedstawiono problem ewolucji oprogramowania i usług sieciowych służących zarządzaniu informacjami o dokumentach z jednoczesnym ukazaniem efemeryczności niektórych z wykorzystywanych przez użytkowników typów narzędzi. Zagadnienie to nie było do tej pory przedmiotem szerszych rozważań. Artykuł może być w przyszłości punktem odniesienia do badań szczegółowych oraz stanowić materiał porównawczy dla podobnych analiz.

## Słowa kluczowe

Folksonomia. Tagi. Hashtagi. Serwisy zakładkowe. Menedżery bibliografii. Generatory bibliografii.

*Otrzymano: 27 czerwca 2016. Zrecenzowano: 27 lipca 2016. Zaakceptowano: 18 sierpnia 2016.*

---

## 1. Wprowadzenie

Bogactwo świata informacji i wiedzy XXI w. sprawia, że coraz częściej użytkownicy stają przed problemem zarządzania metadanymi będącymi częścią zgromadzonej wcześniej informacji osobistej dotyczącej różnych typów dokumentów, z których korzystają w danej chwili, ale także chcieliby ponownie z nich skorzystać w przyszłości. Chodzi zwłaszcza o dokumenty cyfrowe dostępne online, nie mniej jednak, w kontekście problemu analizowanego w niniejszym artykule, w niektórych przypadkach będą też brane pod uwagę

dokumenty zapisane na tradycyjnych nośnikach, których znalezienie w World Wide Web nastęrcza dużych trudności. W celu szybkiego dostępu do wcześniej już znalezionych w sieci, zachowanych i pogrupowanych zasobów użytkownik może skorzystać z różnych metod i narzędzi ułatwiających porządkowanie oraz organizację informacji o tych dokumentach i ich zawartości treściowej. Mowa zwłaszcza o usługach sieciowych i programach komputerowych umożliwiających łatwą organizację i zarządzanie informacją osobistą w postaci metadanych dotyczących zasobów dostępnych w sieci (np. treści wchodzących w skład stron internetowych), jak również usługach wykorzystywanych do gromadzenia i zarządzania opisami bibliograficznymi. Mamy więc z jednej strony do czynienia z narzędziami ułatwiającymi organizację informacji o zasobach cyfrowych i tworzenie tzw. form bibliografiopodobnych, np. hotlist, zestawień załadek czy linków (zob. Woźniak-Kasperek, 2015, 528), z drugiej natomiast – obserwujemy trend zyskiwania na popularności oprogramowania, którego podstawową, ale nie jedyną funkcją jest ułatwienie użytkownikowi przygotowania i zarządzania danymi bibliograficznymi oraz cytowaniami.

Celem artykułu jest przedstawienie zjawisk i powiązanych z nimi narzędzi służących do organizacji informacji o dokumentach i zasobach cyfrowych (w tym zwłaszcza sieciowych), ich głównych funkcji, jak również podstawowych tendencji rozwoju. Pod uwagę brane są zarówno przykłady społecznościowego katalogowania i współtworzenia metadanych przez użytkowników, jak i przedsięwzięcia odnoszące się do informacyjnej działalności bibliotek, w ramach której mamy do czynienia z dodatkowymi metodami porządkowania informacji o dokumentach, jak również kontroli współtworzonych przez internautów metadanych. Przedstawiony został także wpływ sieci drugiej generacji (Web 2.0) na charakteryzowane narzędzia. Uwaga została skupiona na usługach i aplikacjach, których fundamentem jest tworzenie treści online i dzielenie się nimi z innymi użytkownikami. Omówiono stan obecny rozwoju tych narzędzi i ich wykorzystania, a także wskazano kierunki ich ewolucji w kontekście społecznościowego tworzenia przez użytkowników informacji o dokumentach i zasobach cyfrowych.

## 2. Folksonomia i tagowanie

Od ponad dziesięciu lat jesteśmy świadkami rozkwitu zjawiska swobodnego opisywania przez użytkowników zasobów sieciowych przy wykorzystaniu swobodnie dobieranych słów kluczowych. Nie ulega wątpliwości, że jest to jedna z najbardziej rozpowszechnionych społecznościowych form organizowania informacji o dokumentach dostępnych w WWW. Praktyka ta jest określana jako folksonomia<sup>1</sup>. Nazwę tę stosuje się w odniesieniu do

oddolnego klasyfikowania leksji, bez wyraźnej odgórnej hierarchicznej struktury narzucającej schemat klasyfikacyjny, za to z silnie obecnym pierwiastkiem społecznym i publicznym (Kotula, 2009, 143).

Termin „folksonomia” powstał z połączenia dwóch słów: „folk” (z ang. *ludowy*) i „taksonomia”<sup>2</sup> (ewentualnie greckiego „nomia” dla określenia systemu praw rządzących daną

<sup>1</sup> Twórcą terminu „folksonomia” jest Thomas Vander Wal, który w 2005 r. scharakteryzował to zjawisko podczas obrad amerykańskiej konferencji na temat architektury informacji.

<sup>2</sup> Taksonomia rozumiana jest jako klasyfikacja obiektów jednorodnych z pewnego punktu widzenia (Roszkowski, 2007). Zob. także Woźniak-Kasperek, 2011, 153–164; Tomaszczyk, 2007; Kotula, 2009; Stępień, 2010, 58–59; Gilton, 2016, 97; Lambe, 2007, 249–253.

dziedzina) (Pacek, 2010, 203). Dla nazwania tego zjawiska używane są także m.in. takie terminy jak demokratyzacja indeksowania, obywatelska taksonomia, społeczna kategoryzacja czy społeczna klasyfikacja (Gmiterek, 2012, 179). Justyna Hofmokl zaproponowała też określenie „kumplonomia” (Hofmokl, 2009, 164). Użytkownicy, dzięki narzędziom umożliwiającym prosty opis sieciowych treści, tworzą interaktywny system kategoryzacji wiedzy oparty na swobodnej osobistej ocenie, oznaczaniu, interpretowaniu i porządkowaniu informacji o poszczególnych dokumentach. Opisy te są później w większości przypadków udostępnione innym użytkownikom, co służy współpracy internautów, współtworzeniu cyfrowych treści i ich organizacji. Po kilkunastu latach ewolucji różnych narzędzi ułatwiających społecznościową organizację informacji o dokumentach, utwierdzamy się w przekonaniu, że funkcjonują obok siebie dwa równorzędne typy folksonomii – szeroka i wąska, ograniczona (zob. Vander Wal, 2005). W pierwszym przypadku mamy do czynienia z narzędziami i usługami sieciowymi umożliwiającymi opisywanie jednego obiektu przez dużą liczbę osób za pomocą dowolnych zestawów słów kluczowych, np. w serwisach zakładkowych (ang. *social bookmarking*). Wąska folksonomia natomiast jest związana z sytuacją, kiedy użytkownicy oznaczają dokumenty, z których poza nimi rzadko ktoś inny korzysta, np. gromadzą dokumenty przechowywane w serwisach społecznościowych współdzielone z niewielką liczbą użytkowników. Wydaje się, że za wąską folksonomię można też uznać sytuację, gdy użytkownicy tworząc osobistą kolekcję dokumentów cyfrowych (np. rozbudowując swoją wirtualną półkę w bibliotece cyfrowej) opisują poszczególne obiekty lub ich fragmenty dodając do nich swobodnie wybierane słowa kluczowe. Oczywiście zdarza się, że opisy takie mogą być później udostępnione przez użytkownika innym.

Przypomnijmy, że owo oznaczanie czy opisywanie poszczególnych obiektów za pomocą słów kluczowych, umieszczanie odesłań do konkretnych bloków danych i odwoływanie się do nich według pewnych cech, nazywane jest tagowaniem (ang. *tagging*), a termin „folksonomia” używany jest też jako nazwa rezultatu tego działania. Tagiem określa się słowny znacznik, etykietę, czy właśnie słowo kluczowe przypisane do konkretnej, opisywanej przez użytkownika informacji (Pacek, 2010, 203). Sebastian Dawid Kotuła, charakteryzując folksonomię słusznie więc stwierdził, że możemy o niej mówić

dopiero w momencie, gdy ktokolwiek użyje dowolnego taga na opisanie elementu strony bądź całej strony WWW. Tagowanie służy deskrybowaniu dowolnymi słowami kluczowymi, a nie poprzez wykorzystanie kontrolowanego słownictwa (Kotuła, 2009, 143).

Tagi wykorzystywane przy oznaczaniu dokumentów odgrywają tutaj zasadniczą rolę. To one nadają dokumentom dodatkowe znaczenie, ale z punktu widzenia ich twórców:

Opisują obiekt w kategoriach uznanych przez jego odbiorców za relewantne – pożądane i użyteczne (Woźniak-Kasperek, 2011, 193).

Dlatego należy zgodzić się z Lawrence’em Lessigiem, że nie tylko pomagają organizować informacje o dokumentach, ale stanowią swoiste ślady wędrówki internautów,

dzięki którym następni użytkownicy mogą lepiej zrozumieć albo znaleźć te same treści (Lessig, 2009, 67–68).

Tagi to zatem proste metadane w postaci dowolnie dobranych przez użytkowników słów kluczowych. Tagowanie – jak zauważa J. Woźniak-Kasperek

jest w pewnym sensie powrotem do klasyki indeksowania manualnego za pomocą swobodnych słów kluczowych, z tym że tymi, którzy to robią, nie są specjaliści w zakresie indeksowania, ale użytkownicy Internetu (Woźniak-Kasperek, 2011, 192–193).

Należy jednak pamiętać, że tagi mogą dotyczyć i często dotyczą swoistej relacji pomiędzy użytkownikiem i opisywanym dokumentem. Nie zawsze więc odnoszą się do tematu, tytułu czy zawartości, ale mogą mieć np. związek z osobistą opinią użytkownika, kontekstem wykorzystania dokumentu czy jego powiązań z innymi zasobami. W tym przypadku mamy więc do czynienia z tagowaniem na własny użytek, a udostępnienie tych znaczników innym użytkownikom raczej nie ma większego sensu.

Za najczęstsze zalety tagowania uznawane są zwłaszcza: jego globalny zasięg, użyteczność, prostota i elastyczność w tworzeniu znaczników, możliwość odkrywania wcześniej otagowanych dokumentów przez innych użytkowników, współdzielenie się własnymi opisami czy zwyczajowy brak ingerencji w dowolność w tworzeniu poszczególnych słów kluczowych. Najczęściej wymieniane wady to przede wszystkim: subiektywność, brak kontroli nad bliskoźnaczością i stosowanymi synonimami wykorzystywanymi przez użytkowników, błędy i brak precyzji w pisowni słów kluczowych czy brak hierarchii struktury, powiązań i relacji pomiędzy tagami. Notabene popularność poszczególnych znaczników nie zawsze idzie w parze z jakością w ich tworzeniu. Ponadto, często mamy w tym przypadku do czynienia z brakiem homogeniczności form gramatycznych oraz nieumiejętnym łączeniem kilku słów przez użytkowników<sup>3</sup>. Warto zauważyć, że obecnie niektóre serwisy pozwalają na tworzenie tagów złożonych, tzn. takich które składają się z kilku słów oddzielonych spacją. Stosowanie takich tagów jest możliwe na przykład w serwisie Flickr.com, a także w polskich bibliotekach cyfrowych przygotowanych w oparciu o oprogramowanie dLibra.

### 3. Projekt *The Commons* jako przykład efektywnego tagowania zbiorów bibliotecznych

Mówiąc o współpracy bibliotekarzy i użytkowników sieci w ramach katalogowania cyfrowych zasobów należy przynajmniej wspomnieć o jednym z najbardziej efektywnych sposobów wykorzystanie „ducha” zbiorowej inteligencji w kontekście udostępniania i opisywania zbiorów bibliotecznych. Chodzi o zainicjowany przez Bibliotekę Kongresu w 2008 r. projekt *The Commons*, który dał początek trwającemu do dziś przedsięwzięciu, w ramach którego ponad sto instytucji (bibliotek, muzeów i archiwów) promuje fotografie znajdujące się w ich zbiorach, udostępniając je w społecznościowym serwisie Flickr.com. Główne hasło skierowane do odbiorców *The Commons* brzmi: *Pomagaj nam katalogować światowe, publiczne archiwum fotografii!* Współtworzenie metadanych (tagów) dla poszczególnych zdjęć jest tutaj jednym z głównych atutów akcentowanych przez twórców tego projektu (Gmiterek, 2014, 29–30). Wszystkie dodane przez użytkowników tagi podlegają ocenie i komentarzom innych, ale także kontroli pracowników instytucji, które udostępniły poszczególne fotografie. Warto dodać, że w okresie pilotażowym projektu, od stycznia do października 2008 r., użytkownicy serwisu Flickr dodali ponad 67 tys. tagów do 4548 zdjęć, a tylko około

<sup>3</sup> O zaletach i wadach folksonomii zob. np. G. Gmiterek (2012, 184–185), A. Górską (2012), J. Pacek (2010, 208–209), K. Stępień (2010, 56–57).

25 tagów lub komentarzy zostało usuniętych ze względu na niezgodność z tematem zdjęć (Springer et. al., 2008). W marcu 2015 r. w ramach projektu *The Commons* udostępniano już ponad 4 mln zdjęć, a użytkownicy dodali do tego czasu 53 mln tagów (Roth, 2015). Informacje dodawane przez użytkowników to jednak nie tylko słowa kluczowe przypisane do poszczególnych fotografii. Internauci uzupełniają metadane także o dodatkowe adnotacje, które w wielu przypadkach pomagają bibliotekarzom poszerzyć opis udostępnianych dokumentów o nowe elementy (np. rok powstania fotografii, nazwiska osób na nich się znajdujących, autorów zdjęcia, lokalizacji miejsca) (Natanson, 2012).

#### 4. Autotagowanie (automatyczne tagowanie)

Warto zaznaczyć, że serwis Flickr i rozwijany za jego pośrednictwem projekt *The Commons* to przykład zastosowania folksonomii tam, gdzie wykorzystanie któregoś ze znanych narzędzi kontrolowanego słownictwa jest na razie raczej niemożliwe. Za ciekawostkę można uznać uzupełnienie sposobów porządkowania zdjęć w serwisie Flickr przez dodanie algorytmu tzw. autotagowania, polegającego na automatycznym rozpoznawaniu obrazu i nadawaniu zdjęciom znaczników w celu ich tematycznego sortowania. Dla rozróżnienia, automatycznie dodawane tagi są prezentowane obok poszczególnych dokumentów w Flickr.com na białym tle, a znaczniki nadawane przez internautów dostępne są na szarym<sup>4</sup>. Warto jednak zaznaczyć, że mechanizm automatycznego wstawiania znaczników generuje znaczną liczbę błędnych tagów. Często są to bardzo ogólne charakterystyki poszczególnych obiektów (np. informują, że zdjęcie zostało wykonane na zewnątrz lub w pomieszczeniu), ale zdarzają się także sytuacje, kiedy w żaden sposób nie nawiązują do tematu prezentowanych fotografii (np. fragment wieży Eiffła został opisany jako most<sup>5</sup>). Twórcy tego mechanizmu zwracają uwagę, że jeśli skasujemy niewłaściwy znacznik w żaden sposób niepowiązany z tematem opisywanego zdjęcia, algorytm to zauważy i w przyszłości nie będzie popełniał podobnych błędów (Flickr, b.d.). Jednak prawidłowe działanie tego przedsięwzięcia jest bez wątpienia związane z koniecznym dalszym rozwojem wdrożonego algorytmu i mechanizmów automatycznego rozpoznawania obrazu.

#### 5. Tagowanie a katalogi biblioteczne i cyfrowe biblioteki

Nawiązując do działalności bibliotek i oferowanych przez nie usług informacyjnych, warto zaznaczyć, że tagowanie coraz częściej jest także wykorzystywane w komputerowych katalogach bibliotecznych. Od kilku lat staje się ono dodatkowym narzędziem opisu kolekcji bibliotecznych, obok tradycyjnych systemów organizacji wiedzy i języków informacyjno-wyszukiwawczych. Nie zastępując ich oczywiście, ale raczej funkcjonując równolegle. Mamy więc tutaj do czynienia z przypadkami, kiedy z różnym stopniem skuteczności próbuje się zaimplementować dodatkowe systemy porządkowania informacji o dokumentach.

<sup>4</sup> Więcej na temat automatycznego tagowania można się dowiedzieć z artykułu G. Sriharee, 2015.

<sup>5</sup> Przykładem może być zdjęcie M. Davisa: Eiffel Tower at dusk, <https://www.flickr.com/photos/euto-uring/15898124420/> Zob. także Northrup, 2015.

Kilkunastoletnia praktyka pokazuje jednak, że systemy te mimo że z założenia mają pozwalać na swobodne tworzenie i porządkowanie treści, w niektórych przypadkach poddawane są bibliotecznej kontroli wprowadzanych przez użytkowników metadanych. Do dyspozycji są wprawdzie możliwości społecznego tagowania, ale obecny jest też mechanizm weryfikacji terminologii stosowanej przez internautów. Dlatego też za Anną Górską również dzisiaj możemy mówić o tagowaniu kontrolowanym jako o jednej z dwóch głównych tendencji rozwoju folksonomii w bibliotekach (Górska, 2012, 15). Trudno jest w tym miejscu wyrokować, jak mogłoby wyglądać najbardziej skuteczne rozwiązanie problemu weryfikacji stosowanego słownictwa. Być może, jak wskazuje A. Górska, prezentacja stosowanego w ten sposób słownictwa

powinna mieć postać tezauryusa lub słownika haseł przedmiotowych (Górska, 2012, 15).

W większości jednak przypadków wdrażane w bibliotekach rozwiązania umożliwiają czytelnikom i bibliotekarzom posługiwanie się dowolnymi słowami kluczowymi (tagami) opisującymi dostępne w bibliotece dokumenty. Oczywiście często, choć nie zawsze, istnieje tutaj również opcja dodawania przez nich własnych znaczników na potrzeby własne, jak i społeczności czytelników. W praktyce bibliotecznej znajdziemy dzisiaj znaczną liczbę narzędzi służących społecznościowemu tagowaniu gromadzonych treści – zarówno jeśli chodzi o ksiąznice akademickie i naukowe, jak i biblioteki publiczne. W przypadku tych ostatnich ciekawe są wyniki badania przeprowadzonego w Nowej Zelandii (Clements & Li Liew, 2016). Uczestniczący w nim bibliotekarze wyrażali raczej aprobujące opinie na temat zastosowania tagów w katalogach bibliotecznych. Jednak z badania wynika, że nadal część nowozelandzkich bibliotekarzy preferuje stosowanie tradycyjnych systemów organizacji wiedzy. Do tagów podchodzą z rezerwą, choć uważają, że mogą one wzbogacić rekordy bibliograficzne. Używają ich właściwie jako dodatkowej, a nie równorzędnej opcji przeglądania informacji o zbiorach. Co ważne, funkcja dodawania przez bibliotekarzy dowolnych słów kluczowych do gotowych opisów wykorzystywana jest raczej sporadycznie. Na podstawie wyników przeprowadzonych analiz można więc nadal mówić o pewnych oporach przed stosowaniem przez pracowników bibliotek tego typu znaczników, choć trzeba zaznaczyć, że w przypadku instytucji, w których zostało przeprowadzone badanie opcje tagowania zaczęły być udostępniane już w 2008 r.

Także w polskich bibliotekach publicznych tagowanie dokumentów jest wykorzystywane rzadko. Przykładem może być sytuacja w Bibliotece Publicznej m. st. Warszawy, udostępniającej multiwyszukiwarkę Primo, za pomocą której użytkownicy dodali dotychczas tylko trzy publiczne tagi<sup>6</sup>. Wydaje się, że uzasadnieniem takiego stanu rzeczy jest brak znajomości przez czytelników dostępnych mechanizmów i funkcjonalności zarządzania informacjami o dokumentach (w tym oczywiście tagowania) w katalogu bibliotecznym. Z rozmowy z Maciejem Dziubeckim, przedstawicielem Aleph Polska, wynika, że dzisiaj biblioteki implementujące dystrybuowaną przez tę firmę multiwyszukiwarkę Primo raczej rezygnują z opcji tagowania. Powodem jest właśnie rzadkie wykorzystanie jej przez użytkowników innych ksiąznic, w których została wcześniej udostępniona. Być może rozwiązaniem tego typu problemów byłoby przygotowanie przez bibliotekarzy dodatkowych informacji o zasadach

<sup>6</sup> Zob. Katalog główny Biblioteki Publicznej m. st. Warszawy. Tagi [on-line] [04.06.2016], [http://primo.koszykowa.pl/primo\\_library/libweb/](http://primo.koszykowa.pl/primo_library/libweb/)

i możliwościach tagowania. Informacja taka mogłaby być udostępniana zarówno w formie tradycyjnej – papierowej, jak i elektronicznej (w postaci powiadomień na głównej stronie WWW biblioteki, krótkich screenastów, webinarów czy szkoleń on-line użytkowników).

Ze zdecydowanie lepszą sytuacją mamy do czynienia w przypadku bibliotek uniwersyteckich. Chyba najlepszym tego przykładem może być wykorzystywanie tagowania w katalogu Biblioteki Bodlejańskiej w Oksfordzie, w której również stosuje się wcześniej wspomnianą multiwyszukiwarkę Primo. W tym przypadku tagi dodawane są przez czytelników niemal codziennie, a najbardziej aktywny użytkownik otagował do tej pory ponad 1160 dokumentów<sup>7</sup>. Większość z dodanych znaczników ma raczej charakter prywatny, a nie publiczny. Również w Bibliotece Uniwersyteckiej w Warszawie obserwujemy aktywność użytkowników, którzy do tej pory dodali około 850 publicznych tagów. Także i tutaj problemem jest niestety niska jakość dodawanych znaczników (pojedynczych terminów i złożonych wyrażień). Po pierwsze, znaczna część z nich to przypadkowe słowa (np. tagami są cyfry, numeracja wydawnictw ciągłych, sygnatura, pojedyncze litery). Po drugie, tagi często mają bardzo ogólny charakter (np. informatologia, Japonia, biblioteki). Zdarza się, że dodawane do publicznego zasobu słowa kluczowe są bezpośrednio związane z indywidualnymi potrzebami użytkowników (np. pojawiają się osobiste tagi: praca, praca roczna 1, praca roczna 2, przeczytać, przeczytane itp.). Problemem wydaje się być również pojedyncze wykorzystanie poszczególnych znaczników. Do rzadkości należy sytuacja, że dany tag został ponownie użyty przez innego czytelnika<sup>8</sup>.

Od kilku lat istnieje także możliwość organizacji i zarządzania przez użytkowników informacjami o dokumentach dostępnych w polskich bibliotekach cyfrowych. Taki mechanizm funkcjonuje w przypadku oprogramowania dLibra, a przykład jego wykorzystania znajdziemy m.in. w Śląskiej Bibliotece Cyfrowej czy Kujawsko-Pomorskiej Bibliotece Cyfrowej. dLibra umożliwia tworzenie przez użytkownika własnego zasobu tagów prywatnych, dotyczących interesujących go publikacji; tagi te widoczne są tylko dla danego użytkownika. Istnieje także możliwość dodawania propozycji tagów publicznych, które w przyszłości mogą być wykorzystane przy wyszukiwaniu dokumentów przez wszystkich czytelników biblioteki. Ta opcja wymaga jednak wcześniejszej oceny propozycji przez redaktorów bibliotek, którzy mogą modyfikować, odrzucać lub akceptować tagi przesyłane przez czytelników (dLibra, 2010). Zastosowany w dLibrze mechanizm tagowania obiektów cyfrowych może stanowić efektywne narzędzie kontroli procesu dodawania ogólnie dostępnych w wyszukiwaniu słów kluczowych przez użytkowników. Niemniej jednak wykorzystanie przez internautów mechanizmu tagowania jest w praktyce niezmiernie rzadkie. Według Marcina Werli z Poznańskiego Centrum Superkomputerowo-Sieciowego, użytkownicy tych bibliotek w zdecydowanej większości nie korzystają z opcji przeznaczonych dla posiadaczy kont. Zwraca także uwagę nikłą aktywność dotycząca publicznego tagowania dostępnych dokumentów. Do lutego 2012 r. w bibliotekach liczących ponad 60 tys. obiektów, czytelnicy dodali znaczniki do niewielkiej liczby publikacji (wahała się ona od dwóch do szesnastu dokumentów w zależności od danej biblioteki cyfrowej) (Werla, 2012). Dzisiaj niewiele się

<sup>7</sup> Zob. SOLO tags [on-line] [06.06.2016], [http://solo.bodleian.ox.ac.uk/primo\\_library/libweb/action/tagsAction.do?menuitem=5&fn=showTagsPage&fromTop=true&fromPreferences=false&fromEshelf=false&vid=OXVU1](http://solo.bodleian.ox.ac.uk/primo_library/libweb/action/tagsAction.do?menuitem=5&fn=showTagsPage&fromTop=true&fromPreferences=false&fromEshelf=false&vid=OXVU1)

<sup>8</sup> Na podstawie wykazu stosowanych przez czytelników tagów i informacji uzyskanych od pracowników Biblioteki Uniwersyteckiej w Warszawie.

w tym zakresie zmieniło. Generalnie mechanizmy tagowania nadal nie są używane przez użytkowników, dlatego twórcy oprogramowania dLibra odchodzą od tego typu rozwiązań w kierunku możliwości tworzenia prywatnych kolekcji przez czytelników<sup>9</sup>. Taką funkcję zastosowano już np. w serwisie Cyfrowe Zbiory Muzeum Narodowego w Warszawie, wykorzystującym inne oprogramowanie PCSS, tj. dMuseion (<http://cyfrowe.mnw.art.pl/dmuseion>).

Reasumując, możliwości publicznego tagowania dokumentów w polskich bibliotekach cyfrowych nie spotkały się dotychczas z większym zainteresowaniem użytkowników. Być może powodem są pewne ograniczenia w kategoryzacji treści z wykorzystaniem dowolnie wybranych słów kluczowych, zwłaszcza możliwość konfiguracji systemu, która ułatwia automatyczne odrzucanie zaproponowanych przez użytkowników tagów publicznych. Redaktor biblioteki cyfrowej zawsze ma możliwość ich akceptacji w ciągu 30 dni od przesłania przez internautę propozycji znacznika, ale jeśli tego nie robi tag zostaje definitywnie odrzucony (dLibra, 2010). Innym powodem małego zainteresowania tagowaniem, podobnie do jego nikłego wykorzystania w bibliotekach publicznych, może być niezajomość przez czytelników tego mechanizmu i zasad jego działania. Niemniej jednak zaproponowane przez twórców dLibry rozwiązanie stanowi przykład usługi, która zapewnia możliwość tworzenia społecznych klasyfikacji, równocześnie umożliwiając kontrolę wykorzystanego przez użytkowników słownictwa. Można wskazać też inne zastosowania narzędzi weryfikacji dodawanych przez czytelników/internautów słów kluczowych, np. w serwisie społecznościowym LibraryThing ([librarything.com](http://librarything.com)) zapewnia ją usługa LibraryThing for Libraries (Primann, 2012). Do zasobów LibraryThing użytkownicy dodali już prawie 130 mln tagów. W katalogach bibliotecznych, pod względem jakości wprowadzanych danych, tagi te są weryfikowane przez bibliotekarzy, którzy tworzą także powiązania pomiędzy poszczególnymi znacznikami<sup>10</sup>.

Biorąc pod uwagę próby znalezienia przez badaczy sposobu na efektywne kontrolowanie i porządkowanie słów kluczowych dodawanych do opisów dokumentów przez użytkowników sieci, warto przywołać słowa A. Górskiej, która zauważyła, że zmierzamy w kierunku społecznego tworzenia tezaurusów fasetowych. Jako przykład podała ona udostępniony w 2006 r. prototyp narzędzia FaceTag, służącego do społecznego tagowania z uwzględnieniem relacji semantycznych stosowanego słownictwa. Projekt ten rozwiązał wiele kluczowych kwestii dotyczących folksonomii, w tym m. in. problem polisemii, synonimii i zastosowania tagów złożonych (Górska, 2012, 16). Jak wskazuje A. Górska

bardzo szczegółowo zaprojektowane formularze do tworzenia nowych deskryptorów tezausa prowadzą użytkownika krok po kroku do dobrania poprawnej formy hasła i umieszczenia go w gnieździe semantycznym w odpowiedniej kategorii i w odpowiednich relacjach wobec innych haseł. [...] Zaproponowane terminy są weryfikowane przez bibliotekarzy specjalizujących się w językach informacyjnych i dopiero skontrolowane i ewentualnie poprawione trafiają do systemu (Górska, 2012, 16).

Podane rozwiązanie z pewnością może stanowić sposób na uporządkowanie tagów dotyczących udostępnianych przez bibliotekę dokumentów. Narzuca się jednak w tym miejscu

<sup>9</sup> Na podstawie korespondencji autora z M. Werłą z dn. 30.05.2016 r.

<sup>10</sup> Więcej na ten temat można dowiedzieć się bezpośrednio ze strony usługi. Zob. LibraryThing for Libraries. Catalog enhancement package. [on-line] [29.04.2016], <https://www.librarything.com/forlibraries/index.php?page=catalog>

pytanie: czy tego typu projekty nie prowadzą do zbyt daleko idącej kontroli dowolności w tworzeniu znaczników przez użytkowników bibliotek? Mechanizm kontroli uniemożliwia bowiem całkowitą swobodę zamieszczania słów kluczowych, które użytkownik uważa w danym momencie za najbardziej odpowiednie dla organizowania informacji o dokumentach na określony temat. W tym miejscu mówimy oczywiście o szerokiej folksonomii i przygotowywaniu tagów publicznych. Większy problem stanowi dodawanie znaczników prywatnych, które są tworzone przez użytkownika w celu osobistej organizacji danych. Czy te tagi także powinny być kontrolowane w podobny sposób? Ta kwestia pozostaje nadal nierozwiązana.

## 6. Inne przykłady sieciowej organizacji informacji o dokumentach

Kontrola wprowadzanych przez użytkowników metadanych to także ważny element towarzyszący funkcjonowaniu serwisów społecznościowych skierowanych do miłośników książek. Mam tutaj na myśli projekty takie jak *Lubimy Czytać* czy *Goodreads*, służące dzieleniu się informacjami i rekomendowaniu ciekawych publikacji. Serwisy te dają także możliwość tworzenia opisów bibliograficznych dokumentów dodawanych do ich baz oraz uzupełniania ich tagami. Czynność tę mogą jednak wykonywać tylko użytkownicy (nazywani zresztą bibliotekarzami), którzy uzyskali odpowiednie uprawnienia od administratorów serwisu. Prawo tworzenia opisów i tagowania można uzyskać po spełnieniu warunku wcześniejszego dodania do swojej wirtualnej półki odpowiedniej liczby książek, np. w serwisie *Goodreads* jest to 50 pozycji (*Goodreads*, b.d.; *Lubimyczytać*, b.d.). Możliwość katalogowania własnych książek z wykorzystaniem tagów zapewnia także serwis *LibraryThing*. Chodzi zwłaszcza o listy książek, które użytkownik już przeczytał i ocenił. Dzięki temu inne osoby mają okazję zobaczyć, kto czytał podobne publikacje, ale także znaleźć informacje o tytułach, których nie znają i które mogą być dla nich interesujące (na zasadzie pojawiających się w serwisie rekomendacji: *Osoby, które przeczytały tę książkę czytały również...*).

Tagowanie odgrywa istotną rolę również w przypadku serwisu *bibliosfera.net*. Tutaj także można uzupełniać znacznikami treści dodawane przez zarejestrowanych użytkowników. Tagi są elementem obowiązkowym w opisie zasobów. Towarzyszą zarówno linkom, jak i informacjom o poszczególnych wydarzeniach. Co ważne, w przypadku opisów dotyczących wydarzeń (dział „Kalendarz”) raz stworzony tag może być poddawany przez użytkownika zmianom i aktualizacji. Wszystkie dodawane do serwisu tagi mają charakter publiczny, nie ma możliwości zarządzania prywatnymi kolekcjami informacji o zasobach. Nie ma też tutaj żadnej kategoryzacji znaczników, choć funkcjonuje mechanizm podpowiadania tagów (wcześniej użytych przez użytkowników serwisu). Znacznikami automatycznie zostają także nazwy miast, organizatorów imprez czy użytkowników pojawiające się w dodawanych opisach. W 2012 r. średnio 8–9 tagów dodawano ręcznie do jednego linku lub wydarzenia. Długość pojedynczego tagu nie może przekraczać 120 znaków, a serwis pozwala na wprowadzanie znaczników złożonych z więcej niż jednego słowa. Redaktorzy mają możliwość ingerowania w treści wprowadzane przez użytkowników, choć korzystają z niej rzadko (Paleczna, 2013).

Tagowanie jest dzisiaj wykorzystywane w wielu serwisach internetowych, aplikacjach i usługach sieciowych. Znajdziemy je zarówno na platformach blogowych, serwisach typu

social news, ale także stronach typu wiki (np. tworzonych w ramach wikispaces.com) czy w serwisach umożliwiających udostępnianie różnego typu dokumentów cyfrowych (YouTube, Vcasmo, Slideshare, Slideboom). Dostępne są też narzędzia służące do organizacji informacji o posiadanych przez użytkowników dokumentach (plikach muzycznych, filmach, książkach itp.), a także automatycznego przetwarzania wprowadzonych metadanych, które ten proces wspiera. Przykładem może być aplikacja Mp3tag, służąca do zapisywania, kasowania i modyfikacji metadanych o plikach muzycznych, która automatycznie pobiera stworzone wcześniej przez innych znaczniki o poszczególnych plikach audio. Informacje te są pobierane z internetowych baz danych Amazon, freedb czy otwartej encyklopedii muzycznej MusicBrainz.

Odrębnym zjawiskiem jest udostępnianie otagowanych treści przez pracowników bibliotek. W sieci znajdziemy znaczną liczbę przykładów takich działań. Standardem stało się wykorzystanie przygotowywanych przez bibliotekarzy znaczników w porządkowaniu treści wchodzących w skład blogów bibliotecznych. Tutaj wzorem mogą być blogi tworzone przez pracowników Biblioteki Narodowej (<http://www.bn.org.pl/zasoby-cyfrowe-i-linki/blogi>) czy British Library (<http://www.bl.uk/blogs/>). Dodawane przez bibliotekarzy tagi zapewniają także przejrzystość treści na niektórych bibliotekarskich branżowych i instytucjonalnych stronach internetowych, np. na portalu Stowarzyszenia Bibliotek Amerykańskich (<http://www.ala.org/>) i Stowarzyszenia Bibliotekarzy Polskich ([www.sbp.pl](http://www.sbp.pl)), a także na stronie Biblioteki Uniwersytetu Rzeszowskiego (<http://bur.ur.edu.pl/>).

Warto zwrócić uwagę także na wykorzystywanie tagów w serwisach internetowych umożliwiających korzystanie z publikacji naukowych (w tym w czytelniach online). Dostęp do treści o określonej tematyce ułatwiają np. znaczniki dostępne w serwisie Otwórz Książkę, tworzonym przez pracowników Interdyscyplinarnego Centrum Modelowania Matematycznego i Komputerowego Uniwersytetu Warszawskiego. Na stronie głównej tego serwisu zostały one zaprezentowane w postaci chmury tagów. Również na platformie książek elektronicznych iBuk Libra znajdziemy dużą liczbę tagów przyporządkowanych do poszczególnych tytułów udostępnianych w ramach tej usługi. Także i tutaj wcześniej przygotowane znaczniki zapewniają łatwy i przejrzysty dostęp do publikacji.

## 7. Tagi i hashtagi w serwisach społecznościowych

Tworzenie tagów jest dzisiaj powszechnym, choć niepozbawionym wad sposobem na organizację informacji przez użytkowników również w kontekście zarządzania informacją osobistą. Mechanizmy dodawania znaczników znajdziemy w większości serwisów społecznościowych czy internetowych sklepach. Będąca wynikiem tagowania folksonomia stanowi podstawę specyficznego języka tworzonego przez i dla internautów, a jej sukces jest w największej mierze związany z prostotą tworzenia znaczników. Warto zaznaczyć, że dzisiaj użytkownicy tworzą tagi nie tylko w celu opisu konkretnych dokumentów i sieciowych treści. W przypadku serwisów Twitter, Instagram, Facebook czy Tumblr można dodatkowo wykorzystać taką funkcję do tagowania czy oznaczania samych użytkowników. Warto jednak zaznaczyć, że dodając w poście lub komentarzu tag dotyczący konkretnej osoby, internetowej strony czy grupy, tworzony jest automatycznie link do polecanej informacji. Widoczność hiperłącza jest uzależniona od ustawień prywatności konta czy

strony, o której wspominamy. W związku z tym, jeśli otagujemy na przykład nazwę konta grupy, która w serwisie społecznościowym ma charakter zamknięty (jest niedostępna dla użytkowników niebędących jej współuczestnikami), link będzie widoczny tylko dla jej członków. Oczywiście użytkownicy są automatycznie informowani, że wspomnieliśmy o nich w udostępnionej przez nas wiadomości bez konieczności pisania do nich dodatkowych prywatnych wiadomości<sup>11</sup>. Ten specyficzny sposób tagowania jest obecnie wykorzystywany do komunikacji między użytkownikami w coraz większej liczbie serwisów społecznościowych. Pozwala na szybkie powiadomienie danej osoby, aby zainteresowała się określonym tematem czy wiadomością. Inny przykład to tworzenie tagów dotyczących osób widocznych na zdjęciach udostępnianych za pośrednictwem niektórych serwisów społecznościowych. Facebook udostępnia mechanizm DeepFace, generujący automatyczne podpowiedzi dotyczące osób, które znajdują się na fotografiach publikowanych w tym serwisie. Mechanizm analizuje i rozpoznaje twarze osób znajdujących się na zdjęciach. Algorytm ten docelowo ma samodzielnie tagować analizowane wcześniej dokumenty, a propozycje znaczników przedstawiać użytkownikowi do akceptacji (Popielarczyk, 2015; zob. też. Barycki, 2015). Proces tagowania jest więc tu uzupełniony o dodatkowy element automatycznych podpowiedzi, a w przyszłości być może osobliwy mechanizm kontroli poprawności wprowadzanych danych.

Nieco inna jest rola hashtagów, tj. tagów wykorzystywanych w celu uporządkowania informacji w portalu społecznościowym, w którym mamy do czynienia z wieloma treściami tematycznymi i wątkami dyskusji prowadzonymi na jego łamach w czasie rzeczywistym. Hashtagi są pojedynczymi słowami lub wyrażeniami poprzedzonymi wpisywanym bez spacji symbolem „#” (ang. *hash*). Hashtagi w ostatnich latach zyskały dużą popularność dzięki ich wykorzystaniu najpierw w serwisie mikrobloggingowym Twitter, a później również w innych serwisach (np. Facebook, Instagram). Źródeł tworzenia hashtagów można doszukiwać się w sieciach IRC (Internet Relay Chat) i ich wykorzystaniu w ramach porządkowania poszczególnych treści (określania grup i tematów). Dzisiaj stanowią one nieodłączny element konwersacji w mediach społecznościowych i porządkowania wątków będących częścią dyskusji internautów. Mają za zadanie kategoryzować i grupować posty czy informacje poświęcone określonemu wydarzeniu lub tematowi. Dodawanie hashtagów odbywa się z panelu głównego serwisu, dzięki czemu ich wprowadzanie jest prostsze i szybsze niż w przypadku tradycyjnych znaczników zamieszczanych za pomocą dodatkowego okna danego serwisu WWW. W przypadku hashtagów, znaczniki znajdują się wewnątrz komunikatu przesyłanego za pośrednictwem mediów społecznościowych. Leszek Olszański zwraca uwagę, że mogą być one w prosty sposób wykorzystane do tworzenia miniserwisów informacyjnych poświęconych popularnemu tematowi (Olszański, 2012, 238). Można to zrobić na kilka sposobów, jednak

najprostszy, z powodzeniem używany także w automatycznych strumieniach, to po prostu tag z nazwą medium (Olszański, 2012, 238).

Rolę tych znaczników ilustrują następujące przykłady. Po wpisaniu w wyszukiwarce Twittera hashtagu #britishlibrary wyświetlone zostaną informacje i dyskusje dotyczące Biblioteki Narodowej Wielkiej Brytanii. Co ważne, nie są to treści pochodzące tylko

<sup>11</sup> Więcej informacji na ten temat znajduje na stronie WWW centrum pomocy serwisu Facebook. [on-line] [21.04.2016], <https://pl-pl.facebook.com/help/>

i wyłącznie od pracowników tej instytucji, ale przede wszystkim od użytkowników serwisu, którzy interesują bądź interesowali się jej działalnością. Z drugiej strony, po umieszczeniu powyższego hashtaga w naszej wiadomości, automatycznie dołączamy do dyskusji na temat tej instytucji, a nasz post staje się częścią komunikatów jej dotyczących. Podobnie sytuacja wygląda z informacjami i dyskusjami na temat konkretnych produktów – telefonów, tabletów, samochodów itp. Tutaj przykładem może być hashtag #kindlefire, który grupuje informacje na temat tabletu firmy Amazon, ale także powoduje włączenie się do dyskusji na jego temat. Natomiast jeśli chcielibyśmy uzyskać więcej informacji dotyczących konkretnego wydarzenia, którego organizacji towarzyszy udostępnianie postów mikrobloggingowych i włączyć się do dyskusji o nim z innymi użytkownikami, wystarczy skorzystać ze znacznika podanego przez jego organizatorów bądź hashtaga utworzonego przez osoby zainteresowane tym przedsięwzięciem. Np. aby znaleźć informacje o corocznej konferencji online dotyczącej zjawiska biblioteki 2.0 (*Library 2.0. The future of libraries in the digital age*) można użyć hashtag #lib2015, którym oznaczono to wydarzenie. Powyższe przykłady charakteryzują hashtagi jako społecznościowe znaczniki, które w zwięzły sposób opisują dostępny w sieci zasób. Umożliwiając współudział w globalnej dyskusji, są dzisiaj wykorzystywane także do aktualizacji wiedzy przez internautów oraz pozwalają na kreowanie wirtualnych społeczności skupionych wokół danego tematu, problemu czy produktu<sup>12</sup>. Należy przy tym podkreślić, że hashtagowanie to sposób na wzmocnienie przekazu, ale także natychmiastowy dostęp do predefiniowanej problematyki.

Na marginesie naszych rozważań warto także krótko scharakteryzować możliwości wykorzystania hashtagów w ramach tak zwanego wyszukiwania informacji w czasie rzeczywistym. Również i w tym przypadku odgrywają one nieocenioną rolę. Wyszukiwanie w czasie rzeczywistym (ang. *real-time search*) jest bezpośrednio związane z brakiem wynikowego indeksu hiperłączy, z którego byłyby generowane wyniki dotyczące zapytań użytkownika. Narzędzie wyszukiujące komunikuje się z wybranymi serwisami (np. Twitterem, który należy do kategorii serwisów *social networkingowych*, funkcjonujących w czasie rzeczywistym) (Lakomy, 2014, 153), a po otrzymaniu odpowiedzi łączy wyniki w jedną listę, ale też cały czas na bieżąco przedstawia treści, w których wystąpił poszukiwany hashtag. Są to więc też treści, które na bieżąco powstają i są zamieszczane w serwisie. W przeciwieństwie do powszechnie wykorzystywanych wyszukiwarek, nie mamy tutaj do czynienia ze statyczną listą wyników, ale z dynamicznie monitorowanym i aktualizowanym zbiorem informacji na dany temat. Dobrym przykładem takiego narzędzia, które możemy wykorzystać w celu śledzenia newsów w serwisie Twitter jest wyszukiwarka twazzup.com lub twitterfall.com. Narzędzia te oprócz monitoringu treści serwisu mikrobloggingowego wskazują także hashtagi najczęściej wykorzystywane w danym momencie, komentowane tweety czy dodatkowe popularne linki związane z wyszukiwanym słowem kluczowym. Ciekawym rozwiązaniem dotyczącym prezentacji hashtagów jest serwis hashtagify.com, którego głównym celem jest wskazanie pokrewnych słów kluczowych powiązanych ze znacznikiem, który aktualnie wykorzystujemy w trakcie naszego wyszukiwania. Narzędzie to prezentuje także najpopularniejsze wypowiedzi użytkowników Twittera, dotyczące interesującego nas hashtaga. Nieco inną funkcjonalność proponuje serwis trendsmap.com, który ma za zadanie przedstawiać na wirtualnej mapie świata najpopularniejsze w danym momencie

<sup>12</sup> Szczegółowe rozważania na ten temat zostały zaprezentowane w Yang, 2012.

hashtagi i profile użytkowników. Zapewnia on możliwość sprawdzenia czyje wypowiedzi są najczęściej komentowane, a także jakie słowa kluczowe aktualnie cieszą się największym powodzeniem wśród użytkowników danego regionu, państwa czy miasta. Warty wskazania jest również serwis tagboard.com, który gromadzi i grupuje aktualne informacje z kilku serwisów społecznościowych. Oczywiście również i w tym przypadku punktem wyjścia dla użytkownika jest hashtag powiązany z konkretnym tematem.

## 8. Delicious i kłopoty społecznościowego zakładkowania (social bookmarking)

Kontynuując rozważania na temat sposobów społecznościowego zarządzania i organizacji danych o dokumentach, konieczna wydaje się analiza dotycząca jednego z głównych zjawisk towarzyszących powstaniu i ewolucji funkcjonalności mechanizmów sieci drugiej generacji. Mam mianowicie na myśli zjawisko społecznościowego zakładkowania (ang. *social bookmarking*) i jego potencjalny wpływ na wspólne, sieciowe gromadzenie i udostępnianie informacji o ciekawych cyfrowych zasobach w postaci tzw. ulubionych (zakładek). Informacje te za pośrednictwem narzędzi dostępnych online są organizowane za pomocą metadanych w formie tagów współtworzonych przez internetowe społeczności. Gromadzone w ten sposób znaczniki są powiązane z konkretnymi treściami dostępnymi w sieci tworząc w ten sposób swoiste katalogi rekomendowanych stron. Stanowią one jednocześnie zbiory informacji o najciekawszych stronach w World Wide Web, często dodatkowo ocenionych i skomentowanych przez innych użytkowników.

Mimo że zjawisko tworzenia zakładek online sięga korzeniami do połowy lat 90. XX w., kiedy to udostępniono pierwsze narzędzia ułatwiające przenoszenie osobistych list zakładek do Internetu, to momentem przełomowym dla społecznościowego tworzenia sieciowej bazy „ulubionych” było utworzenie przez Joshuę Schachtera w 2003 r. serwisu Delicious, a następnie powstanie znacznej liczby innych podobnych do niego przedsięwzięć. Jednak to właśnie Delicious stał się synonimem zjawiska, które na kilka lat zdominowało wspólne tworzenie przez internautów globalnej bazy informacji o stronach WWW i ich zawartości. Był on projektem, który w ramach wspólnej organizacji informacji o dokumentach jako jeden z pierwszych wykorzystał właśnie społeczną kategoryzację treści. Na początku nowego milenium większość internautów indywidualnie tworzyło zasoby swoich ulubionych stron WWW za pomocą przeglądarki, z której na co dzień korzystali. Każdy z nich kreował swój własny zbiór informacji, z których w przyszłości mógłby skorzystać bez ponownego wyszukiwania w globalnej sieci znalezionych wcześniej treści. Taki sposób organizacji zakładek wiązał się z kilkoma ważnymi problemami, z których najważniejszym wydaje się dostępność informacji o zapisanych adresach tylko z poziomu konkretnego komputera, ale także brak możliwości łatwego oraz szybkiego dzielenia się nimi z innymi. Nie bez znaczenia była również konieczność regularnego eksportowania opisanych adresów URL do pliku, jeśli internauta chciał mieć pewność, że gromadzone dane są zabezpieczone przed awarią komputera, systemu operacyjnego czy przypadkowym skasowaniem. Usługa Delicious i podobne do niej narzędzia rozwiązywały te problemy przez udostępnienie funkcjonalnego narzędzia online, które umożliwiło użytkownikom szybkie tworzenie zakładek i towarzyszących im słów kluczowych (tagów) oraz dodatkowych adnotacji, ich łatwą organizację, a także – co ważne

– scentralizowane dzielenie się informacjami o zawartości treściowej stron internetowych z innymi użytkownikami (Gmiterek, 2012, 188). Warto dodać, że Delicious charakteryzował się także przez długi czas sprawnie funkcjonującym mechanizmem podpowiadania słów kluczowych, które były wcześniej wykorzystane przez innych internautów w kontekście opisywanych przez nas treści. Rekord danego zasobu sieciowego w tym serwisie zawierał zbiór tagów użytych wcześniej do jego opisanie (Roszkowski, 2015, 147). Mechanizm ten dawał pewną kontrolę nad stosowanym słownictwem, choć oczywiście występowanie czy podpowiadanie przez Delicious danego słowa kluczowego nie wiązało się z koniecznością jego wykorzystania. Użytkownik zawsze mógł zaproponować swój tag, który w późniejszym czasie z dużym prawdopodobieństwem byłby podpowiadany innym internautom.

Nie ulega wątpliwości, że w pierwszych latach funkcjonowania serwisów typu Delicious dynamika rozwoju zjawiska *social bookmarking* zaczęła być wiązana z ułatwieniem i przyspieszeniem wyszukiwania, gromadzenia, grupowania oraz współdzielenia zasobów informacyjnych w Internecie, które w dużej mierze można uznać za informację osobistą zarządzaną cyfrowo przez internautów. Usługi *social bookmarking* są też często wymieniane jako jeden ze sposobów usprawnienia i przyspieszenia pewnych czynności w kontekście tak zwanego lifehackingu i zarządzania danymi (Pash & Trapani, 2012, 79–82). Mówiąc w 2007 r. o Delicious, Joshua Schachter stwierdził wręcz, że narzędzie to może być rozbudowywane w celu przechowywania, kategoryzowania i udostępniania informacji o wszystkich aspektach naszego cyfrowego życia (Galanciak, 2007). Do niedawna mogło się wydawać, że narzędzia *social bookmarking* to efektywne rozwiązanie informacyjno-organizacyjnych potrzeb internautów. Jednak duża liczba funkcjonujących jeszcze kilka lat temu usług straciła na popularności i większość z nich niestety zniknęła z sieci. Przy czym samo zjawisko zamykania społecznościowych serwisów i sieciowych usług nie powinno dziwić. „Śmiertelność” to niestety jedna z podstawowych cech projektów Web 2.0. Sytuacja taka ma miejsce, kiedy dany projekt nie zgromadzi wokół siebie w odpowiednim czasie wystarczającej liczby użytkowników lub ich grono stale się pomniejsza i w końcu serwis jest likwidowany, a jego miejsce zajmują inne przedsięwzięcia (Zalewski, 2007). Powodów takiego stanu rzeczy jest dużo. W wielu przypadkach nowsze lub na bieżąco aktualizowane usługi są po prostu bardziej atrakcyjne dla użytkowników, np. z punktu widzenia przejrzystości, intuicyjności interfejsu, sposobu opisu dokumentów czy dzielenia się nimi z innymi, funkcjonalności mechanizmów współpracy w chmurze obliczeniowej, a dzisiaj także możliwości korzystania z nich na urządzeniach mobilnych. Z drugiej jednak strony, jesteśmy coraz częściej świadkami sytuacji, kiedy mając do dyspozycji szereg rozwiązań i usług o podobnych funkcjach, użytkownicy koncentrują się na serwisach najbardziej popularnych. Pozostałe usługi, po pewnym czasie są zamykane i znikają z sieci. Dodatkowo, dla użytkowników, którzy chcą organizować i dzielić się z innymi metadanymi dotyczącymi różnych typów dokumentów, do dyspozycji jest także znaczna liczba usług sieciowych i programów umożliwiających zarządzanie informacją bibliograficzną. Od pewnego czasu obserwujemy też wykorzystywanie w celu gromadzenia i dzielenia się adresami ciekawych stron WWW serwisów bazujących na rankingach popularności udostępnianych przez internautów wiadomości (np. digg.com czy wykop.pl). Serwisy tego typu oferują oczywiście podstawową możliwość zarządzania i tematycznego grupowania informacji o dokumentach, jednak funkcjonalność tych usług dotyczy raczej zjawiska *social news* i udostępniania oraz wspólnego oceniania przez internautów różnych wiadomości i dokumentów, mniej natomiast gromadzenia

i współdzielenia społecznej czy prywatnej bazy rozbudowanych o dodatkowe opisy zakładek. Warto zaznaczyć, że w literaturze dotyczącej zjawiska *social bookmarking* znajdziemy informacje o takich narzędziach, które wskazywane są jako przykłady społecznościowego zakładkowania, choć faktycznie nie zawsze nimi są.

Znalezione przez użytkowników informacje są też często gromadzone i udostępniane na łamach różnych portali i serwisów społecznościowych (np. Facebook, Twitter), w ramach których mogą oni tworzyć dodatkowo sieci i grupy osób o podobnych zainteresowaniach, wymieniających się atrakcyjnymi dla nich informacjami. Obecnie znajdziemy wiele usług umożliwiających rozbudowaną organizację informacji o sieciowych dokumentach, np. w postaci profili czy kanałów tematycznych, w ramach których użytkownik gromadzi (sam lub wspólnie z innymi) informacje o multimedialnych zasobach dotyczących konkretnego tematu lub zagadnienia. Przykładem takiej usługi jest serwis Scoop (scoop.it), podobne funkcje oferuje także serwis Pinterest (pinterest.com), będący swoistym agregatorem multimedialnych treści (zdjęć, grafiki, filmów, prezentacji).

Na nieco innej zasadzie działa natomiast serwis Pearltrees (pearltrees.com). W tym przypadku mamy do czynienia z usługą umożliwiającą organizację informacji dotyczących zarówno linków, jak i dodanych przez użytkownika plików (np. zdjęć, dokumentów tekstowych w formacie PDF) czy zapisanych w serwisie notatek. Wszystkie gromadzone informacje użytkownicy mogą dodawać do oddzielnych kategorii (tzw. *pearltrees*), których zawartość jest przedstawiana w formie tematycznych katalogów lub pojedynczych, nieprzyporządkowanych do żadnej kategorii plików. Oczywiście w każdej chwili można zmodyfikować tworzone katalogi. Użytkownicy mogą również importować wcześniej tworzone przez nich zakładki, gromadzone zarówno w przeglądarce internetowej, jak i serwisach społecznościowych (np. Delicious). Zaimportowane w ten sposób dane są następnie automatycznie porządkowane alfabetycznie, ale można zmienić ustawienia przyporządkowania poszczególnych danych do odrębnych kategorii. Można także współtworzyć tematyczne katalogi z innymi internautami zainteresowanymi danym problemem. Podobnie do najpopularniejszych społecznościowych serwisów, Pearltrees umożliwia śledzenie aktywności i katalogów innych użytkowników. Serwis jest także zintegrowany z Facebookiem i Twitterem, dzięki czemu można pobierać z nich informacje, jak również dzielić się zgromadzonymi treściami poprzez te usługi. W przypadku wspólnego tworzenia tematycznych katalogów, każdy z użytkowników może obserwować wprowadzane zmiany i ewentualnie przywracać skasowane przez kogoś innego treści.

Powodem coraz mniejszej popularności serwisów społecznościowych koncentrujących się na ułatwianiu tworzenia i organizacji zakładek jest zapewne także pojawienie się możliwości oferowanych przez aktualne wersje najbardziej popularnych przeglądarek internetowych. Coraz częściej ich funkcjonalność jest bezpośrednio związana z udostępnieniem mechanizmów (rozszerzeń) spersonalizowanego zarządzania zakładkami przechowywanymi w chmurze obliczeniowej. Do gromadzonych w tematycznych folderach zakładek możemy dodawać obrazy i dodatkowe opisy. Oczywiście możliwa jest w tym przypadku także synchronizacja wprowadzanych danych, dzięki czemu są one dostępne na różnych urządzeniach powiązanych z kontem użytkownika. Takie rozwiązania znajdziemy w przeglądarkach Google Chrome, Firefox, Opera czy Internet Explorer (zob. Stelzet-Morawietz, 2016). Co więcej, w przypadku rozszerzenia Bookmark Manager dla Google Chrome mamy także możliwość udostępniania gromadzonych zakładek innym użytkownikom.

Warto zaznaczyć, że z listy ponad 50 serwisów zakładkowych, która jest dostępna w artykule Seana P. Aune'a opublikowanym w 2007 r. na łamach strony internetowej Mashable.com (Aune, 2007), do dzisiaj przetrwało 15, przy czym niektóre z nich, z powodu niewielkiej liczby użytkowników, nie są już na bieżąco aktualizowane tak często, jak to bywało jeszcze kilka lat wcześniej. Również serwis Delicious po kilku latach funkcjonowania jako usługa kupiona od J. Schachtera przez Yahoo!, a później odsprzedawana innym firmom (Delicious, 2016), dzisiaj jest tylko cieniem wcześniejszej funkcjonalności, oferując podstawowe funkcje zapamiętywania i przechowywania zakładek.

Charakteryzując zjawisko *social bookmarking* i omawiając aktualne problemy dotyczące zmian zachodzących w dostępności i funkcjonalności poszczególnych narzędzi, warto również przynajmniej wspomnieć o udostępnionej w 2006 r. usłudze PennTags, przygotowanej i rozwijanej w Bibliotece Uniwersytetu w Pensylwanii. Było to typowe narzędzie służące społecznościowemu zakładkowaniu, o podobnych do Delicious funkcjach, jednak stworzone specjalnie dla czytelników (w tym zwłaszcza studentów, naukowców) i bibliotekarzy. Fakt ten oznacza, że celem PennTags było raczej zarządzanie informacją dotyczącą dokumentów udostępnianych przez macierzystą dla jego twórców bibliotekę, a nie wszystkimi treściami znajdującymi się w sieci. W związku z tym usługa ta umożliwiała lokalizację, organizację i dzielenie się zakładkami do dokumentów i zasobów udostępnianych za pośrednictwem internetowych katalogów bibliotecznych czy pełnotekstowych baz danych wykorzystywanych na Uniwersytecie w Pensylwanii. Oczywiście, podobnie jak w innych narzędziach tego typu, zachowane były tutaj społecznościowe możliwości współtworzenia zakładek (w PennTags mogły być one wykorzystane w ramach tworzenia tematycznych zbiorów informacji czy realizowanych przez studentów i pracowników naukowych projektów), ich przechowywanie w chmurze, jak również wyszukiwanie informacji w zasobach utworzonych przez innych użytkowników i śledzenie ich aktywności. Dodatkowo, serwis oferował możliwość tworzenia swoistych bibliografii adnotowanych i zapisywanie ich w indywidualnych katalogach (w serwisie nazywanych projektami). Bibliografie adnotowane były także tworzone i udostępniane przez samych bibliotekarzy w celu rekomendacji prac dotyczących konkretnych tematów (Farakas, 2007, 142). Warto zaznaczyć, że w przypadku PennTags, zresztą podobnie do innych rozwiązań tego typu, możliwa była automatyzacja czynności dodawania zakładek poprzez zastosowanie tak zwanego bookmarkletu (skryptozakładki), czyli prostego linku w formie przycisku, który w łatwy sposób można umieścić na pasku przeglądarki bez potrzeby instalacji dodatkowych wtyczek czy oprogramowania. Dzięki zastosowaniu bookmarkletu użytkownik, w chwili kiedy chciał dodać nową zakładkę, nie musiał za każdym razem wchodzić na stronę domową PennTags. Wystarczyło, że skorzystał z osadzonego w przeglądarce przycisku i w ten sposób wywołał okno tworzenia i zapisywania dodawanych przez siebie ulubionych stron. Jednak także w przypadku PennTags w ostatnim czasie odnotowywano coraz mniej liczną grupę użytkowników. Pomimo że od dziesięciu lat serwis ten był wymieniany przez badaczy i bibliotekarzy jako przykład modelowego wdrożenia idei *social bookmarking* w ramach informacyjno-bibliotecznych usług, PennTags 15 czerwca 2016 r., po dekadzie funkcjonowania, został zamknięty, a sam projekt przeszedł do historii. Jednym z głównych powodów takiego stanu rzeczy jest wykorzystanie przez użytkowników innych interaktywnych narzędzi umożliwiających organizację metadanych w bardziej przyjazny sposób, lepiej przystosowany do ich indywidualnych potrzeb (PennTags, 2016).

## 9. Zjawisko *social clippingu*

Zjawisko społecznościowego zakładkowania i towarzyszące mu usługi sieciowe od pewnego czasu oferują użytkownikom nieco bardziej rozbudowane możliwości spersonalizowanego zarządzania informacją o dokumentach cyfrowych przy aktywnym wykorzystaniu mechanizmów tagowania. Takie możliwości zapewniają m.in. serwisy społecznościowe typu *social clipping* (zob. Wowra, 2008). Serwisy te nie tylko umożliwiają tworzenie zakładek, ale także oferują dodatkowe usługi, w tym dotyczące swego „wycinania” i opisywania (tagowania, oceniania, adnotowania) fragmentów czy elementów wchodzących w skład stron internetowych lub innych cyfrowych dokumentów. Są to więc serwisy ułatwiające zarządzanie informacją osobistą w jeszcze bardziej spersonalizowany sposób niż miało to miejsce w Delicious lub w przypadku innych narzędzi utożsamianych ze społecznościowym zakładkowaniem. Jednym z najczęściej wymienianych narzędzi tego typu jest serwis Diigo ([diigo.com](http://diigo.com), wcześniej działający pod nazwą Furl), który oprócz opcji „wycinania” poszczególnych elementów stron internetowych i ich opisywania, ułatwia także dołączanie do tak przygotowanych metadanych dodatkowych załączników w postaci dokumentów PDF. Użytkownik ma też możliwość otagowania gromadzonych zasobów jako pojedynczych obiektów cyfrowych lub ich powiązania z innymi dokumentami oraz gromadzenia w postaci tematycznej listy. Co ciekawe, dla każdej tak przygotowanej informacji można wygenerować odrębny adres internetowy (permanenty link), za pomocą którego inni internauci mogą dotrzeć do udostępnionych przez użytkownika treści. Każdym tak przygotowanym dokumentem można dzielić się z użytkownikami sieci za pomocą poczty elektronicznej lub najpopularniejszych serwisów społecznościowych. Można też zintegrować profil w Diigo z profilem w serwisie Twitter, dzięki czemu użytkownik może jednocześnie budować swój zasób zakładek i zamieszczać w tym serwisie mikrobloggingowym informacje o gromadzonych „ulubionych”. Diigo oferuje ponadto tworzenie tzw. *watchlist*, czyli definiowania tagów, za pomocą których w przyszłości użytkownik może śledzić dodawane do serwisu odnośniki. Jest to opcja bazująca na selektywnej dystrybucji informacji i mechanizmie, który umożliwia automatyczne przysyłanie do użytkownika informacji o nowych zasobach odpowiadających jego zainteresowaniom. Funkcja ta przypomina swoją funkcjonalnością technologię RSS (Really Simple Syndication), jednak w tym przypadku dotyczy zasobów dodanych do serwisu Diigo, a nie różnych źródeł informacji. Oczywiście, serwis umożliwia współpracę z innymi jego użytkownikami w ramach tematycznych grup i wspólne opisywanie znalezionych w sieci treści. Dodatkowo, tworząc zakładki dla poszczególnych stron czy ich fragmentów, serwis proponuje słowa kluczowe bezpośrednio powiązane z treścią danego zasobu. Propozycje są przy tym związane z wcześniej wykorzystywanymi przez użytkownika tagami, a także słowami kluczowymi używanymi przez innych użytkowników. Do szybkiej i efektywnej organizacji gromadzonych informacji twórcy serwisu proponują narzędzie Outliner, które pomaga w grupowaniu i przygotowaniu tematycznych list linków wraz z możliwością dodawania do nich swoich notatek. Utworzone w ten sposób dane można łatwo wyeksportować do edytora tekstu Word. Również w przypadku list tematycznych można wygenerować do każdej z nich indywidualny, trwały odnośnik (permalink).

Warto zaznaczyć, że podobny mechanizm „wycinania” i opisywania sieciowych treści oferuje Cyfrowa Biblioteka Narodowa Polona. Tutaj również, oprócz tworzenia prywatnych metadanych dotyczących poszczególnych cyfrowych obiektów, użytkownik może

zaznaczyć wybrany przez niego obszar udostępnionego przez bibliotekę dokumentu, opisać go własnymi tagami oraz opatrzyć dodatkowymi informacjami w postaci adnotacji odnoszących się do dokumentu lub jego fragmentu. Tak przygotowane przez użytkownika zasoby i opisy poszczególnych obiektów są przechowywane przez niego w ramach indywidualnych kolekcji (wirtualnych półek).

## 10. Organizacja i zarządzanie danymi bibliograficznymi

Problem zarządzania informacją o dokumentach od dłuższego czasu dotyczy również możliwości funkcjonalnego wykorzystania sieciowych usług ułatwiających gromadzenie, organizację i dzielenie się danymi bibliograficznymi. W tym kontekście ważne wydaje się współdziałanie internautów w tworzeniu zbiorów opisów, a także wspólne kreowanie bazy danych będącej społecznościowym źródłem informacji o różnego typu dokumentach. Obecnie istnieje znaczna liczba narzędzi online umożliwiających zarządzanie informacją bibliograficzną, jej współdzielenie czy przystosowywanie do indywidualnych potrzeb użytkownika. Co więcej, od ponad dziesięciu lat mamy do czynienia z usługami, których celem jest wspomaganie prac bibliograficznych. Zgodnie z duchem sieci drugiej generacji stanowią one swoiste platformy współpracy online. Nie bez znaczenia jest tutaj też fakt, że użytkownik może z nich korzystać za pośrednictwem przeglądarek internetowych, bez potrzeby instalacji oprogramowania desktopowego. Dotyczy to zwłaszcza usług działających na zasadzie serwisów społecznościowych, skierowanych głównie do studentów, pracowników nauki i bibliotekarzy, dla których korzystanie z gotowych danych bibliograficznych podczas przygotowywania prac dyplomowych czy naukowych może być istotną pomocą w prawidłowym przygotowaniu bibliografii załącznikowej, a także w posługiwaniu się cytowaniami wykorzystanych dokumentów. Funkcjonalność tych usług w dużym stopniu wiąże się z możliwością pobierania gotowych opisów z bibliograficznych baz danych, cyfrowych bibliotek i repozytoriów, a także dzieleniem się tymi opisami z innymi użytkownikami. Przykładem takiego rozwiązania może być serwis CiteULike ([citeulike.org](http://citeulike.org)), jak również naukowy serwis BibSonomy, uruchomiony w 2005 r. na Uniwersytecie Kassel. W obu przypadkach mamy do czynienia z usługami, których fundamentem jest wprawdzie tagowanie i tworzenie zakładek do stron WWW, jednak ich funkcjonalność zdecydowanie różni się od możliwości serwisów typu Delicious czy Diigo. Krótka charakterystyka serwisu BibSonomy pozwoli wyjaśnić na czym polegają te różnice.

## 11. BibSonomy, personomia i semantyczne społecznościowe zakładkowanie

Obok opcji tworzenia zakładek do stron WWW BibSonomy oferuje możliwość organizacji informacji o dokumentach w postaci opisów bibliograficznych przygotowanych w popularnych standardach (w tym BibTex, HTML, EndNote, MS Office XML) zgodnie z zasadami wybranych stylów bibliograficznych (APA, Chicago, Harvard)<sup>13</sup>. Usługa ta

<sup>13</sup> Podobne możliwości oferuje też np. serwis CiteULike.

jest także dobrym rozwiązaniem dla osób, które zamierzają utworzyć własną podręczną bibliotekę dokumentów, wykorzystując możliwość wczytywania plików z komputera i ich powiązania z internetowym kontem użytkownika w BibSonomy. Serwis ten umożliwia też generowanie bibliografii załącznikowej przygotowanej na podstawie gromadzonych przez użytkownika danych. Informacje bibliograficzne można importować i eksportować. Serwis oferuje również tworzenie relacji i powiązań pomiędzy dodawanymi przez użytkownika tagami. Można ręcznie tworzyć związki dotyczące poszczególnych znaczników – zastępować jedne tagi innymi, dodawać tzw. supertagi i subtagi oraz ewentualnie automatycznie kasować je pojedynczo i grupowo (Jäschke et al., 2007). Można także kontrolować relacje słów kluczowych, dotyczące tagów utworzonych przez nas samych i dodawanych przez innych użytkowników.

Warto zaznaczyć, że serwis BibSonomy jest w literaturze przedmiotu podawany jako przykład zarówno folksonomii jak i tzw. personomii (ang. *personomy*) (Anderson, 2016, 48). Określenie to odnosi się do zjawiska tworzenia przez użytkownika swojego własnego systemu organizacji tagów i zależności między nimi. Zasobem takim użytkownik może dzielić się z innymi osobami, korzystającymi z danej usługi sieciowej, które mogą dla własnych potrzeb stosować przygotowane przez niego rozwiązania. Taka indywidualna i spersonalizowana kolekcja znaczników nazywana jest właśnie personomią, natomiast zbiór wszystkich personomii tworzonych w ramach jednego serwisu lub aplikacji tworzy folksonomię (Hotho, 2010, 61)<sup>14</sup>. Serwis BibSonomy w literaturze naukowej zaliczany jest też do narzędzi semantycznego społecznego zakładkowania (ang. *social semantic bookmarking*) (San, 2010, 235) oraz wykorzystania w tym zakresie zamiast tagów adnotacji opartych na słownikach kontrolowanych lub ontologiach. Tzw. semantyczne znaczniki, dzięki odgórnej kontroli słownictwa i możliwości ich prawidłowej interpretacji przez technologie cyfrowe, są w stanie rozwiązać niektóre problemy związane z klasyczną folksonomią, w tym te dotyczące normalizacji spójności znaczników, ich synonimiczności lub wieloznaczności. W publikacjach na temat semantycznego społecznościowego zakładkowania omawiane są także inne projekty i serwisy internetowe, w których autorzy widzą przykłady znaczącej ewolucji narzędzi służących organizacji informacji o dokumentach cyfrowych. Charakterystyczne jest, że po stosunkowo krótkim okresie rozwijania tych przedsięwzięć nie są już one dostępne w sieci. Można tylko domniemywać, że taki stan rzeczy może mieć związek z prezentacją ich prototypowych wersji (pilotażowych lub demonstracyjnych), które nie zgromadziły wokół siebie wystarczającej liczby użytkowników i zostały zamknięte przez swoich twórców<sup>15</sup>.

## 12. Generatory bibliografii

Obecnie dostępne jest wiele sieciowych narzędzi, które wspierają aktywność użytkowników związaną z tworzeniem przez nich spisów bibliograficznych. W pierwszej kolejności należy wymienić tzw. generatory bibliografii i cytowań, które pozwalają na poprawne tworzenie bibliografii załącznikowej z wykorzystaniem opisów bibliograficznych pobieranych

<sup>14</sup> Zob. także Roszkowski, 2015, 147.

<sup>15</sup> Mowa zwłaszcza o takich serwisach jak Int.ere.st, GroupMe, Fuzzzy, SOBOLEO, Faviki i Annotea.

z internetowego źródła, udostępnionego wcześniej szerszej publiczności, np. WorldCat. Usługi te są skierowane przede wszystkim do uczniów i studentów przygotowujących prace pisemne i potrzebujących pomocy przy sporządzaniu zestawień bibliograficznych. Wykorzystane przez nich opisy bibliograficzne mogą dotyczyć różnych typów i rodzajów dokumentów. Ich lista jest zazwyczaj rozbudowana i obejmuje nie tylko tradycyjne dokumenty, ale także np. elektroniczną korespondencję, sekwencje wideo udostępnione na łamach danej strony WWW czy podcasty. Generatory bibliografii nie pozwalają jednak na zaawansowane zarządzanie opisami, ich grupowanie czy współpracę z innymi użytkownikami. Umożliwiają jedynie przygotowanie, generowanie i importowanie danych bibliograficznych dotyczących wykorzystywanych przez użytkownika dokumentów. Zapewniają korzystanie z popularnych stylów bibliograficznych, jednak ich wybór w poszczególnych usługach nie zawsze jest wystarczający (szczególnie jeśli chodzi o darmowe wersje wykorzystywanej przez użytkownika usługi). Generatory bibliografii umożliwiają zazwyczaj także eksport przygotowanego przez użytkownika spisu bibliograficznego do pliku Word lub innego edytora tekstu. Przykładami takich narzędzi są EasyBib ([www.easybib.com](http://www.easybib.com)), Cite This For Me (<http://www.citethisforme.com/>), Citation Machine (<http://www.citationmachine.net/>) czy Harvard Generator (<http://www.harvardgenerator.com/>). W przypadku generatora EasyBib mamy od pewnego czasu do czynienia z usługą bezpośrednio powiązaną z sieciowym pakietem biurowym Dokumenty Google'a. EasyBib jest dodatkiem (wtyczką) do tej usługi. Niektóre z generatorów, np. EasyBib czy Cite This For Me, są także powiązane z mobilnymi aplikacjami, które umożliwiają tworzenie spisów bibliograficznych za pomocą urządzenia przenośnego. Aplikacje te pozwalają skanować kody kreskowe znajdujące się na okładkach tradycyjnych publikacji i automatycznie pobierać do pamięci urządzenia metadane tych dokumentów. Warto jednak zaznaczyć, że w trakcie testów przeprowadzonych z wykorzystaniem wybranych generatorów bibliografii pojawiło się szereg błędów dotyczących pisowni w automatycznie pobieranych opisach. Oczywiście użytkownik ma na każdym etapie tworzenia bibliografii możliwość wprowadzenia korekt do gromadzonych danych (zob. także Kulczycki, 2012; Rozkosz, 2014).

### 13. *Reference Manager* oraz oprogramowanie do organizacji i zarządzania informacją bibliograficzną

Bardziej zaawansowane mechanizmy i funkcje zarządzania informacją bibliograficzną zapewniają narzędzia typu *reference manager*, które w polskim piśmiennictwie są nazywane menedżerami bibliografii. Podstawą działania tych narzędzi jest oferowanie praktycznych narzędzi umożliwiających szybkie i funkcjonalne gromadzenie, zarządzanie i dzielenie się z innymi użytkownikami danymi bibliograficznymi. Pomagają też prawidłowo i w sposób jednolity tworzyć cytowania oraz zapewniają zarówno automatyzację eksportu i importu informacji bibliograficznej ze specjalistycznych pełnotekstowych baz danych czy katalogów OPAC, jak i automatyczne generowanie opisów z innych cyfrowych źródeł (w tym m.in. z serwisów internetowych, podcastów, blogów, prezentacji, kanałów RSS itp.). Nie dziwi więc, że usługi sieciowe lub samodzielne oprogramowanie do zarządzania informacją bibliograficzną (opisami i cytowaniami), jak się wydaje ze względu na różnorodność możliwości ich zastosowania, są w literaturze przedmiotu łączone z wcześniej wspomnianą

koncepcją zarządzania informacją osobistą (Świgoń, 2012, 199–200). Narzędzia te pozwalają na funkcjonalne przechowywanie gromadzonych danych w chmurze, zorganizowanych w zdefiniowany przez użytkownika sposób, oraz ich łatwe sortowanie, edycję i wyszukiwanie według zadanych przez niego kryteriów. Użytkownik może również tworzyć opisy czy cytowania posługując się wybranym stylem bibliograficznym, np. w przypadku usługi Zotero ma dzisiaj do wyboru prawie 8200 stylów. Dla lepszej organizacji gromadzonych treści może też wykorzystywać tagi i tworzyć dodatkowe adnotacje, abstrakty itp. Do dyspozycji użytkownika jest także opcja kreowania powiązań dotyczących informacji o gromadzonych dokumentach, dzięki czemu może on szybko zbudować spersonalizowany system relacji pomiędzy różnymi publikacjami. Narzędzia do zarządzania informacją bibliograficzną zazwyczaj mogą być integrowane z popularnymi (darmowymi i komercyjnymi) edytorami tekstu. Taka integracja ułatwia szybkie przeszukiwanie zgromadzonych danych bibliograficznych oraz np. zamieszczanie odpowiednich cytowań w trakcie przygotowywania przez użytkownika dokumentu tekstowego. Opcje te są dostępne z poziomu edytora tekstu. Kluczowy w przypadku tych narzędzi wydaje się fakt, że często mogą one służyć nie tylko do kreowania list czy wręcz katalogu informacji o interesujących użytkownika dokumentach, a także powiązania tych danych z konkretnymi adresami URL czy plikami przechowywanymi na dysku lokalnym lub zdalnym serwerze. W ten sposób użytkownik ma możliwość budowania spersonalizowanej „biblioteki” dokumentów dostępnej w chmurze za pośrednictwem komputera i urządzenia przenośnego. W większości przypadków możliwa jest współpraca z innymi użytkownikami danego narzędzia przy wspólnym tworzeniu kolekcji opisów bibliograficznych. Przykładem usługi wyposażonej w takie funkcje jest Zotero, które pozwala na kreowanie grup roboczych dotyczących konkretnego tematu, zagadnienia czy danej społeczności (np. użytkowników będących uczniami jednej klasy, szkoły).

Narzędzia do zarządzania opisami bibliograficznymi i cytowaniami mogą mieć formę webowego interfejsu dostępnego za pośrednictwem przeglądarki internetowej lub samodzielnego oprogramowania (względnie aplikacji) instalowanych na dysku stacjonarnego czy mobilnego urządzenia. Użytkownik ma często do dyspozycji dodatek (wtyczkę) do przeglądarki lub specjalnie do tego celu przygotowany bookmarklet. W przypadku Zotero oprogramowanie automatycznie rozpoznaje typ dokumentu, dla którego generowany jest opis, który następnie trafia do zasobu danych gromadzonych przez użytkownika.

#### **14. Zarządzanie informacją o dokumentach za pomocą aplikacji i urządzeń mobilnych**

Coraz ważniejszą rolę w zarządzaniu informacją o dokumentach pełnią urządzenia i aplikacje mobilne. Dostępne dzisiaj aplikacje mobilne charakteryzują się funkcjonalnymi mechanizmami ułatwiającymi gromadzenie danych z różnych źródeł. Również ich prezentacja znacząco odbiega od sposobów udostępniania multimedialnych treści powszechnymi jeszcze kilka lat temu. Użytkownik ma do dyspozycji szereg rozwiązań, których główną zaletą jest selektywna dystrybucja informacji (tzn. generowanie przez aplikację informacji na tematy, które są dopasowane do wcześniej określonych przez użytkownika preferencji), a także niekonwencjonalne i efektowne zaprezentowanie gromadzonych w ten sposób treści (np. prezentowane są one na ekranie urządzenia mobilnego w formie gazety lub

czasopisma). Aplikacje mobilne do zarządzania informacją zapewniają funkcje znane z tradycyjnego oprogramowania i interaktywnych usług udostępnionych za pośrednictwem World Wide Web, w tym sieciowe i społecznościowe gromadzenie informacji o dokumentach. Przykładem takiego rozwiązania jest aplikacja Pocket (<https://getpocket.com/>), pozwalająca na zapisywanie różnego rodzaju cyfrowych treści w celu zapoznania się z nimi w późniejszym czasie, a także umożliwiającą tagowanie i kategoryzowanie tych danych. Użytkownicy mają do dyspozycji także szereg aplikacji mobilnych, służących przede wszystkim do katalogowania dokumentów. Przykładem jest aplikacja Libib (<https://www.libib.com/>), pozwalająca na tworzenie mobilnego katalogu informacji o zgromadzonych przez użytkownika książkach, płytach z muzyką, filmach i grach komputerowych. Za pomocą łatwego w użyciu mechanizmu skanowania kodów paskowych zamieszczonych na okładkach tych dokumentów, aplikacja automatycznie pobiera gotowe dane bibliograficzne wraz z uzupełniającymi informacjami o dokumencie (np. streszczeniem, recenzjami). Dane te można później poszerzać o dodatkowe elementy, takie jak tagi, adnotacje czy własne oceny i recenzje. Również poszczególne programy i usługi sieciowe umożliwiające zarządzanie danymi bibliograficznymi mają odpowiedniki w postaci aplikacji mobilnych. Oferują je np. zarówno usługi darmowe (np. Zotero<sup>16</sup> i Mendeley<sup>17</sup>), jak i komercyjne (np. Papers<sup>18</sup> i EndNote<sup>19</sup>).

## 15. Zakończenie

Sieciowe narzędzia społecznościowe umożliwiające zarządzanie danymi o dokumentach znacznie ułatwiają spersonalizowane porządkowanie informacji pochodzącej z różnych źródeł. Bez wątpienia możliwości współpracy i wspólnego tworzenia treści dotyczących różnorodnych dokumentów to funkcjonalność, która również w przyszłości będzie odgrywała znaczącą rolę w organizowaniu i kompletowaniu informacji (także tej bibliograficznej). Pomimo faktu, że coraz częściej obserwujemy znikanie sieciowych usług ułatwiających zarządzanie informacją o dokumentach (np. niektórych serwisów social bookmarking), mamy do czynienia z powstawaniem szeregu projektów, które korzystając z wcześniejszych doświadczeń podobnych przedsięwzięć, oferują zmodyfikowane, coraz lepsze i coraz bardziej przyjazne dla użytkownika funkcje i sposoby tworzenia oraz dzielenia się danymi o cyfrowych zasobach. Widoczny jest szybki rozwój technologii, oprogramowania i usług sieciowych towarzyszących tym procesom. Dotyczy to także aplikacji na urządzenia mobilne, które zyskują coraz większe znaczenie we współczesnych praktykach zarządzania informacją. Niemniej jednak, mamy także do czynienia z systemami umożliwiającymi użytkownikom samodzielną kategoryzację i tagowanie dokumentów, których funkcjonalność

<sup>16</sup> Przykładem jest aplikacja ZotPad przygotowana z myślą o użytkownikach urządzeń Apple (<https://itunes.apple.com/us/app/zotpad/id497260579?mt=8>).

<sup>17</sup> Zob. stronę internetową dotyczącą aplikacji Mendeley przygotowanej z myślą o użytkownikach urządzeń z systemem Android (<https://play.google.com/store/apps/details?id=com.mendeley&hl=pl>).

<sup>18</sup> Zob. stronę internetową aplikacji Papers 3 przygotowanej z myślą o użytkownikach urządzeń Apple (<http://papersapp.com/ios/>).

<sup>19</sup> Zob. stronę internetową aplikacji EndNote for iPad przygotowanej z myślą o użytkownikach urządzeń Apple (<https://itunes.apple.com/us/app/endnote-for-ipad/id593994211?mt=8>).

nie zawsze jest przez internautów dostrzegana. Np. dzieje się tak w przypadku multiwyszukiwarek towarzyszących katalogom bibliotecznym. Użytkownicy korzystają z tych społecznych klasyfikacji stosunkowo rzadko. Widać dzisiaj wyraźnie, że folksonomie najlepiej sprawdzają się tam, gdzie nie ma narzędzi profesjonalnego indeksowania dokumentów, np. za pomocą kontrolowanego słownictwa<sup>20</sup>. Przykładem są serwisy społecznościowe czy narzędzia służące do organizacji informacji bibliograficznej, w których od kilku lat obserwujemy trend upowszechniania się zjawiska tworzenia znaczników przez internautów. Gromadzą oni dane o dokumentach, dzielą się nimi, ale także korzystają z gotowych, wcześniej przygotowanych przez innych informacji w celu porządkowania i zarządzania posiadanymi przez siebie dokumentami, jak również cyfrowymi treściami, które wcześniej znaleźli w sieci i chcieliby z nich skorzystać także w przyszłości. Warto zaznaczyć, że wśród różnorodnych przedsięwzięć mamy też do czynienia z próbami wdrożenia rozwiązań mających na celu kontrolę jakości wprowadzanych danych, a także z próbami wykorzystania znaczników opartych na słownikach kontrolowanych lub ontologiach.

## Bibliografia

- Anderson, P. (2016). *Web 2.0 and Beyond: Principles and Technologies*. Boca Raton: CRC Press.
- Aune, S.P. (2007). *Social bookmarking: 50+ social bookmarking sites* [online]. Mashable [29.04.2016], <http://mashable.com/2007/08/08/social-bookmarking-2/#kKw3ArUWgGqr>
- Barycki, P. (2015). *Boisz się, że Facebook włączy autooznaczenie? Spokojnie, wyjaśniamy o co chodzi* [online]. Spider'sWeb. Blog blisko technologii [20.04.2016], <http://www.spidersweb.pl/2015/02/facebook-autooznaczenie.html>
- Brzostek-Pawłowska, J. (2011). Zmiany w modelach i technologiach informacyjnych w dobie Web 2.0 i Web 3.0. *Elektronika*, 3, 143–151.
- Clements, L.; Li Liew, Ch. (2016). Talking about tags: An exploratory study of librarians' perception and use of social tagging in a public library. *The Electronic Library*, 34(22), 289–301.
- Delicious (2016). *Wikipedia, The Free Encyclopedia* [online] [29.04.2016], [https://en.wikipedia.org/wiki/Delicious\\_%28website%29](https://en.wikipedia.org/wiki/Delicious_%28website%29)
- dLibra (2010). *Mechanizm tagowania obiektów cyfrowych* [online]. Instrukcja obsługi systemu dLibra 5.0 oraz dMuseion 1.0 [27.04.2016], <https://confluence.man.poznan.pl/community/pages/viewpage.action?pageId=13631676>
- Farkas, M. G. (2007). *Social Software in Libraries: Building Collaboration, Communication, and Community Online*. Medford: Information Today.
- Flickr (b.d.). *The Help Forum. Updates on tags* [online], [28.04.2016], <https://www.flickr.com/help/forum/en-us/72157652019487118/>
- Galanciak, T. (2007). Wywiad: *Joshua Schachter, del.icio.us* [online]. Magazyn T3 [12.05.2016], <http://magazynt3.pl/Wywiad-Joshua-Schachter-delicious/>
- Gilton, D.L. (2016) *Creating and promoting lifelong learning in public libraries : tools and tips for practitioners*. Lanham: Rowman & Littlefield.
- Gmiterek, G. (2012). *Biblioteka w środowisku społecznościowego internetu: biblioteka 2.0*. Warszawa: Wydaw. SBP.
- Gmiterek, G. (2014). Biblioteka 2.0 w 2014 r. – analiza zjawiska. *E-mentor*, 2, 24–33.
- Goodreads (b.d.). *Librarian manual* [online]. Goodreads [24.04.2016], [https://www.goodreads.com/librarian\\_manual](https://www.goodreads.com/librarian_manual)

<sup>20</sup> Por. Sosińska-Kalata, 2008, 43.

- Górska, A. (2012). Tagowanie kontrolowane – oksymoron czy przyszłość języków informacyjno-wyszukiwawczych. *Zagadnienia Informacji Naukowej*, 2, 6–17.
- Hofmokl, J. (2009). *Internet jako nowe dobro wspólne*. Warszawa: Wydaw. Akademickie i Profesjonalne.
- Hotho, A. (2010). Data mining on folksonomies. In: G. Armano et al. (eds.) *Intelligent Information Access*. Berlin: Springer, 57–82.
- Jäschke, R.; Hatho, A.; Schmitz, Ch.; Grah, M.; Krause, B.; Stumme, G. (2007). *Organizing Publications and Bookmarks in BibSonomy* [online]. WWW2007: 16th International World Wide Web Conference, May 8–12, 2007, Banff, Alberta, Canada [27.05.2016], [http://www2007.org/workshops/paper\\_25.pdf](http://www2007.org/workshops/paper_25.pdf)
- Kotuła S. D. (2009). Folksonomia – narodziny i charakterystyka (w aspekcie wyszukiwania informacji). *Folia Bibliologica*, 51, 139–149.
- Kulczycki, E. (2012). *Jak zrobić bibliografię? Internetowe generatory dla studentów i uczniów* [online]. Warsztat badacza – Emanuel Kulczycki [6.06.2016], [http://ekulczycki.pl/warsztat\\_badacza/jak-zrobic-bibliografie-internetowe-generatory-dla-studentow-i-uczniow/](http://ekulczycki.pl/warsztat_badacza/jak-zrobic-bibliografie-internetowe-generatory-dla-studentow-i-uczniow/)
- Lakomy, M. (2014). Tweety na szczycie. Polityka responsywna. *Studia Medioznawcze*, 2, 153–164.
- Lambe, P. (2007). *Organising knowledge : taxonomies, knowledge and organisational effectiveness*. Witney: Chandos Publishing (Oxford) Ltd.
- Lessig L. (2009). *Remiks. Aby sztuka i biznes rozkwitały w hybrydowej gospodarce*. Warszawa: Wydaw. Akademickie i Profesjonalne.
- Lubimyczytać (b.d.). *FAQ dla bibliotekarzy* [online]. Lubimyczytac.pl [27.04.2016], <http://lubimyczytac.pl/faq-dla-bibliotekarzy#temat-1>
- Natanson, B. O. (2012). *A Happy Anniversary: Four Years of Sharing on Flickr* [online]. Picture This. Library of Congress Prints & Photos [26.04.2016], <http://blogs.loc.gov/picturethis/2012/01/a-happy-anniversary-four-years-of-sharing-on-flickr/>
- Northrup, L. (2015). *Flickr Robots Auto-Tag Every Photo, Auto-Annoy Users* [online]. Consumerist, May 21 [26.04.2016], <https://consumerist.com/2015/05/21/flickr-robots-auto-tag-every-photo-auto-annoy-users/>
- Olszański, L. (2012). *Media i dziennikarstwo internetowe*. Warszawa: Wydaw. Akademickie i Profesjonalne.
- Pacek, J. (2010). *Bibliografia w zmieniającym się środowisku informacyjnym*. Warszawa: Wydaw. SBP.
- Palczna, D. (2013). Folksonomia w serwisie bibliosfera.net – liczby, przykłady, charakterystyka [online]. *Podkarpackie Studia Bibliologiczne*, 2 [8.06.2016], <http://psb.ur.edu.pl/psb022013>
- Pash, A.; Trapani, G. (2012). *Lifehacker. Jak żyć i pracować z głową?* Gliwice: Wydaw. Helion.
- PennTags (2016). [online] [20.05.2016], <http://tags.library.upenn.edu/>
- Pirmann, C. M. (2012). Tags in the Catalogue: Insights From a Usability. Study of LibraryThing for Libraries. *Library Trends*, 61, 1, 234–247.
- Popielarczyk, T. (2015). *Facebook samodzielnie oznaczy osoby na każdym opublikowanym zdjęciu. Czy tego chcą czy nie* [online]. Antyweb [20.04.2016], <http://antyweb.pl/facebook-samodzielnie-oznaczy-osoby-na-kazdym-opublikowanym-zdjeciu-czy-tego-chca-czy-nie/>
- Roszkowski, M. (2008). Czym jest tagowanie? [online] *Warsztaty Bibliotekarskie*, 3–4, [19.04.2016], <http://www.pedagogiczna.edu.pl/warsztat/2008/3-4/080306.htm>
- Roszkowski, M. (2007). Folksonomia jako narzędzie społecznego tagowania. *Warsztaty Bibliotekarskie* [online], nr 4, [19.04.2016], <http://www.pedagogiczna.edu.pl/warsztat/2007/4/070404.htm>
- Roszkowski, M. (2015). Wykorzystanie technologii informacyjnych do gromadzenia i analizy danych ze społecznościowego serwisu Lubimy Czytać. W: G. Gmiterek, M. Ochmański, M. Roszkowski (red.), *LaTel – z badań nad wykorzystaniem technologii informacyjnych w bibliologii i informatologii*. Warszawa: Wydaw. SBP.
- Roth, M. (2015). *The Flickr Commons welcomes our 100th institution* [online], Flickr blog [29.04.2016], <http://blog.flickr.net/en/2015/03/25/the-flickr-commons-welcomes-our-100th-institution/>

- Rozkosz, E. (2014). *Generator bibliografii w Dokumentach Google* [online]. Edukacja informacyjna [6.06.2016], <http://www.edukacjainformacyjna.pl/2014/03/generator-bibliografii-w-dokumen-tach-google/>
- San, M. (2010). *Handbook of Research on Web 2.0, 3.0, and X.0: Technologies, Business, and Social Applications: Technologies, Business, and Social Applications, vol. 1*. Hershey: Information Science Reference.
- Sosińska-Kalata, B. (2008). Kryteria ocen stosowanych wobec systemów porządkowania dokumentów W: *Dokument, książka i biblioteka w badaniach naukowych i nauczaniu uniwersyteckim*. M. Skalska-Zlat & A. Żbikowska-Migoń. Wrocław: Wydaw. Uniwersytetu Wrocławskiego, 18–46.
- Springer, M.; Dulabahn, B.; Michel, P.; Natanson, B.; Reser, D.; Woodward, D.; Zinkham, H. (2008). *For the Common Good: The Library of Congress Flickr Pilot Project* [online]. Library of Congress. Prints & Photographs Reading Room [26.04.2016], [www.loc.gov/rr/print/flickr\\_report\\_final.pdf](http://www.loc.gov/rr/print/flickr_report_final.pdf)
- Sriharee, G. (2015). An ontology-based approach to auto-tagging articles [online]. *Vietnam Journal of Computer Science*, 2, 85 – 94 [26.04.2016], <http://link.springer.com/article/10.1007%2Fs40595-014-0033-6>.
- Stelzel-Morawietz, P. (2016). Test przeglądarek internetowych. *PC World*, 5, 84–91.
- Stępień K. (2010). *Folksonomie czyli społecznościowe opisywanie treści: poradnik*. Warszawa: Wydaw. SBP.
- Świgoń, M. (2012). *Zarządzanie wiedzą i informacją. Podstawy teoretyczne. Badania w wymiarze indywidualnym*. Olsztyn: Wydaw. UWM.
- Tomaszczyk, J. (2007). Taksonomia jako narzędzie organizacji informacji. *Zagadnienia Informacji Naukowej*, 1, 40–49.
- Vander Wal, T. (2005). Explaining and Showing Broad and Narrow Folksonomies [online]. *Vanderwal.net* [27.04.2016], <http://www.vanderwal.net/random/entrysel.php?blog=1635>
- Werla, M. (2012). Web 2.0 i (polskie) biblioteki cyfrowe [online]. *Biuletyn EBIB*, [27.04.2016], [http://www.ebib.pl/images/stories/numery/129/129\\_werla\\_.pdf](http://www.ebib.pl/images/stories/numery/129/129_werla_.pdf)
- Wowra, W. (2008). Ulubione on-line. *Magazyn Internet*, 8, 50.
- Woźniak-Kasperek, J. (2011). *Wiedza i język informacyjny w paradygmacie sieciowym*. Warszawa: Wydaw. SBP.
- Woźniak-Kasperek, J. (2015). Bibliografia a katalog biblioteczny – dyskusja o pojęciach i terminach. *Przegląd Biblioteczny*, 4, 517–532.
- Yang, L.; Sun, T.; Zhang, M.; Mei, Q. (2012). *We Know What @You #Tag: Does the Dual Role Affect Hashtag Adoption?* [online]. University of Michigan. Personnel WWW Server [22.04.2016], <http://www-personal.umich.edu/~qmei/pub/www2012-yang.pdf>
- Zalewski, P. (2007). *Osiem cech Web 2.0* [online]. Internet Standard [12.05.2016], <http://www.internetstandard.pl/news/107199/Osiem.cech.Web.2.0.html>

## Networked Approaches to Social Information Management on Documents

### Abstract

**Purpose/Thesis:** The aim of this paper is to describe the phenomena of social organization of information in networked environment and related tools and applications. This includes the concept of social cataloguing, user generated metadata and library related projects where crowdsourced metadata are the subject of interest.

**Approach/Research methods:** The paper is based on the literature review and critical analysis methodology. Where applicable, the functional analysis of software and web services was used to frame the process of information organization.

**Results and conclusions:** In the recent years the dynamic change has been witnessed in social networking services for web document management. One may observe the phenomenon of the disappearance of sites, which lost their popularity. They are often replaced with new, more interactive and user-friendly tools. Moreover, the information users more and more often change the reality of information with mobile applications which foster the process of recording and sharing information on the Web. These applications and tools, such as desktop software and web services, help users to collect, manage, and share content. The phenomenon involves the possibility of using online tools the functionalities of which are based on the idea of openness, online web communities and the idea of Semantic Web.

**Originality/Value:** The paper is aimed to present a synthetic view on the evolution of software and web services dedicated to web document management. It also pinpoints the issue of ephemerality of this type of tools and applications. This approach has not been thoroughly investigated in the literature yet. The paper may be used as a background for further analyses and comparative studies in the field.

### Keywords

Folksonomies. Tags. Hashtags. Bookmarking services. Bibliography management tools. Bibliography generators.

---

*Dr GRZEGORZ GMITEREK jest adiunktem w Katedrze Bibliografii i Dokumentacji na Wydziale Dziennikarstwa, Informacji i Bibliologii Uniwersytetu Warszawskiego. Specjalizuje się w problematyce stosowania nowych technologii w bibliotekach (ze szczególnym uwzględnieniem wykorzystania przez bibliotekarzy elementów i narzędzi wchodzących w skład szeroko pojmowanego zjawiska Web 2.0). Najważniejsze publikacje: Biblioteka w środowisku społecznościowego internet (biblioteka 2.0), Warszawa 2012.; The book in era of new technologies, integration and interactivity of media [W:] History 2.0. Panta Rhei. Symposium Proceedings 19th General Congress of Polish Historians, September 17 Szczecin, Anna Sobczak, Marta Cichocka, Piotr Frąckowiak, ed. Lublin 2014, 75–81; eBook 2.0. Książka w rzeczywistości sieci drugiej generacji i mobilnych technologii [W:] W kręgu kultury czytelnicznej dzieci i młodzieży, pod red. M. Antczak, A. Walczak – Niewiadomskiej, Łódź 2015, 51 – 68; Rzeczywistość rozszerzona a książka i prasa [w:] La-Tel – z badań nad wykorzystaniem technologii informacyjnych w bibliologii i informatologii, pod. red. G. Gmiterka, M. Ochmańskiego i M. Roszkowskiego. Warszawa 2015, 41–66.*

*Kontakt z autorem:*

*ggmiterek@uw.edu.pl*

*Katedra Bibliografii i Dokumentacji*

*Wydział Dziennikarstwa, Informacji i Bibliologii*

*Uniwersytet Warszawski*

*ul. Nowy Świat 69*

*00–927 Warszawa*

**Dorota Siwecka:** *Światowy model informacji bibliograficznej : programy i projekty (1950–2010)*. Wrocław; Warszawa: Wydaw. SBP, 2015, ss. 355, [1] (Nauka, Dydaktyka, Praktyka; 162). ISBN 978-83-64203-48-0

Polska literatura naukowa dotycząca teorii bibliografii przechodzi obecnie kryzys: w ciągu ostatnich 16 lat (2000–2015) ukazało się zaledwie 14 pozycji zwartych z tego zakresu, z czego: cztery są wznowieniami dzieł dawniejszych<sup>1</sup>, trzy są poświęcone wąskim zagadnieniom (monografia bibliograficzna, bibliografia osobowa historyka)<sup>2</sup>, dwie wydano w formie podręcznika i mają charakter przeglądowy<sup>3</sup>, jedna traktuje bibliografię jako naukę pomocniczą literaturoznawstwa<sup>4</sup>, a ostatnia jest zbiorem refleksji i anegdot w popularnym ujęciu<sup>5</sup>. Po odjęciu tej sumy zostajemy z liczbą trzech oryginalnych publikacji naukowych zajmujących się teorią lub metodyką bibliografii<sup>6</sup>. Jest to dorobek dość skromny, wzięwszy pod uwagę burzliwy okres, jaki pod wpływem nowoczesnych technologii i sposobów rejestracji piśmiennictwa przechodziła ostatnio bibliografia. Stąd niezwykle ważne jest pojawienie się publikacji Doroty Siweckiej; bynajmniej nie z tego powodu, że nastąpiło to w dotkliwym okresie stagnacji polskiej myśli bibliograficznej, ale przede wszystkim dlatego, że jest pracą rzetelną, wszechstronną, przekrojową i dogłębną, niemającą odpowiedników w polskiej i zagranicznej literaturze.

Dorota Siwecka jest doktorem nauk humanistycznych w zakresie bibliologii i informatologii o specjalności bibliografia i stopień ten uzyskała – pod panieńskim nazwiskiem Chłopkowska – w 2013 r. dzięki rozprawie *Kształtowanie się światowego modelu informacji bibliograficznej. Programy i projekty (1950–2010)*. Dysertacja ta, w poszerzonym wymiarze, została udostępniona szerszej publiczności jako *Światowy model informacji bibliograficznej*. Książka składa się z siedmiu rozdziałów szczegółowych, omawiających kolejne aspekty współpracy międzynarodowej (organa odpowiedzialne za współpracę i stworzone przez nie programy; unifikacja sposobów identyfikacji i opracowania dokumentów oraz wymiany tych danych; organizacja pracy bibliograficznej na poziomie narodowym) w zakresie tworzenia bibliografii narodowych pomyślanych jako kolejne ogniwa, które zsumowane przybliżą w miarę pełny obraz produkcji wydawniczej świata. Części te zostały poprzedzone rozdziałem o istocie bibliografii (w którym niestety zabrakło odniesienia do definicji słownikowych i normalizacyjnych, a tym samym oceny autorki, czy wytrzymują one próbę czasu w kontekście nowoczesności; w ogóle egzemplifikację w tej kwestii trzeba uznać za niewystarczającą), a zakończone analizą zawartości poszczególnych bibliografii narodowych. Pracę dopełniają spis treści w języku polskim i angielskim,

<sup>1</sup> M. Dembowska Maria: *Metoda Bibliografii Polskiej Karola Estreichera*. Warszawa 2001 [pierwotruk: 1954]. K. Estreicher: *O bibliografii*. Oprac. E. Stefańczyk. Warszawa 2008 [zbiór artykułów, pierwotruk: 1865]. H. Hleb-Koszańska: *Spór o początki bibliografii w Polsce*. Oprac. J. Sadowska, E. Szalecka-Mróż. Warszawa 2003 [pierwotruk artykułu: 1973]. A. Sitarska: *Z warsztatu bibliografa*. Białystok 2006 [zbiór artykułów, pierwotruki: 1970–2006].

<sup>2</sup> A. Bajor: *Monografia bibliograficzna. Analiza modelowa*. Częstochowa 2003. A. Bajor: *Z zagadnień form bibliografii. Monografia bibliograficzna*. Warszawa 2005. A. Znajomski: *Bibliografie osobowe historyków polskich*. Lublin 2004.

<sup>3</sup> *Bibliografia. Metodyka i organizacja*. Pod red. Z. Żmigrodzkiego. Warszawa 2000. *Informacja naukowa. Rozwój, metody, organizacja*. Pod red. Z. Żmigrodzkiego, Warszawa 2006.

<sup>4</sup> T. Winek: *Nauki pomocnicze literaturoznawstwa*. Warszawa 2007.

<sup>5</sup> E. Dombek: *Potyczki bibliografa czyli bibliografia od kuchni*. Warszawa 2015.

<sup>6</sup> A. Matczuk: *Polskie bibliografie nauk humanistycznych i społecznych do roku 1989*. Historia i metodyka, Lublin 2014. J. Pacek: *Bibliografia w zmieniającym się środowisku informacyjnym*. Warszawa 2010. M. Skalska-Zlat: *Bibliografia w Polsce 1945–1996. Naukoznawcza analiza dyscypliny*. Wrocław 2002.

wstęp, zakończenie, obszerna i wyczerpująca bibliografia (276 źródeł, 32 normy, 271 opracowań, 15 informatorów – łącznie 594 pozycje), wykaz skrótów, skrótów tytułów czasopism, indeks przedmiotowo-osobowy. Zabrakło niestety wykazu 11 tabel i 19 rysunków. Także bibliografia ma pewne konstrukcyjne braki, którym można było poświęcić nieco więcej namysłu, bowiem choć kolejność i grupowanie wykorzystanych dokumentów jest słuszne, to jednak spis jest mało przejrzysty. W części poświęconej źródłom znalazły się bardzo zróżnicowane dokumenty (programy różnych sekcji IFLA, bibliograficzne bazy danych, instrukcje obsługi programów, akty prawne czy teksty tablic klasyfikacyjnych), co może zdezorientować czytelnika. Można było wprowadzić bardziej szczegółowy podział w obrębie przyjętych działów. Podobna sytuacja ma miejsce w części dotyczącej opracowań.

Należy podkreślić przemyślaną kompozycję tekstu we wspomnianych rozdziałach szczegółowych oraz konsekwencję w kolejności omawianych elementów – korzystanie z publikacji ułatwia rozpoczynanie omówienia od przedsięwzięć o najszerszym zasięgu (Program Powszechnej Rejestracji Bibliograficznej), by przez standardy międzynarodowe poprowadzić czytelnika do formatów opisów bibliograficznych i wymiany danych, stosowanych w poszczególnych państwach. To także doskonale ilustruje stopień skomplikowania omawianego zagadnienia. Bardzo interesująco prezentuje się rozdział szósty, który ukazuje rozwój formatów grupy MARC, uzupełniony o opis formatów powstałych w państwach byłego bloku socjalistycznego, a także bardzo cenne wyjście poza omówienie potrzeb bibliotek i uwzględnienie także rozwiązań przeznaczonych dla rynku wydawniczego i księgarskiego. Złożoność nakreślonych w tym fragmencie problemów uwypukla rolę prób unifikacji opracowania rzeczowego dokumentów, które stały się przedmiotem rozdziału siódmego. Jego podstawowym atutem jest ukazanie powiązań między różnymi językami informacyjno-wyszukiwawczymi i przedstawienie procesu ich doskonalenia. Ciekawie zaprezentowano również ewolucję opracowania rzeczowego w bibliografiach narodowych, od lat 50. do 90. XX w. Zabrakło jednak przedstawienia stanu rzeczy z początku XXI w., co jest o tyle zastanawiające, że w wielu miejscach pracy autorka powołuje się na najnowsze dane. Zastanawia także brak zaznaczenia kontekstu historyczno-politycznego w niektórych miejscach (np. w rozdziale 6.1.2.), co utrudnia pełne zrozumienie treści.

Autorka w swej pracy dążyła „do przedstawienia i podsumowania zjawisk, jakie wystąpiły w bibliografii na forum międzynarodowym w okresie powojennym, a praktycznie poczynawszy od lat 50. XX w. do pierwszej dekady XXI w. włącznie” (s. 19) i przyjęła założenie, że zmiany w bibliografii i jej metodyce, podyktowane w dużej mierze wprowadzeniem nowych rozwiązań technologicznych, przebiegły od etapu „narodowego” do międzynarodowej unifikacji i globalizacji. Analizowane sześćdziesięciolecie podzieliła na trzy okresy: pierwszy (lata 50. XX w.), był czasem formułowania w poszczególnych krajach własnych zasad opisu bibliograficznego; w drugim (druga połowa XX w.), międzynarodowe wytyczne opracowywania bibliografii przystosowywano do ustalonych rozwiązań krajowych; trzeci (od końca XX w.) charakteryzuje się bezpośrednim przyjmowaniem norm międzynarodowych. Ta periodyzacja nie była jednak punktem wyjścia dla konstrukcji opracowania.

Przedstawienie i podsumowanie omówionych przez autorkę zjawisk oparte zostało głównie na analizie istniejącej literatury, a liczba pozycji w bibliografii daje ogłęd stopnia jej wykorzystania. Autorka sięgnęła do każdej możliwej pozycji<sup>7</sup>, która mogła ją wesprzeć w tworzeniu tej rozległej syntezy, ale przede wszystkim do źródeł (zwłaszcza cenny jest wgląd w oficjalną dokumentację, zawierającą zalecenia i wytyczne z zakresu metodyki i organizacji bibliografii na szczeblu międzynarodowym). Było to zadanie niełatwe, bo jak sama przyznaje „wymienione opracowania albo skupiają się na wybranych jej [bibliografii – JMŁ] aspektach (opracowaniu formalnym, rzeczowym, formatach wymiany danych itd.), albo stanowią ogólne opracowania, których aktualność ograniczona jest do lat 80. XX w. [...] lub

<sup>7</sup> Choć być może nie każda powinna być wzięta pod uwagę: trudno za właściwe uznać powołanie się na dokument pt. *ZSIW – zagadnienie egzaminacyjne* (nadesłała: dr D. Pietruch-Reizes), zamieszczony w witrynie internetowej *Bibliomaniak* – <http://bibliomaniak.w.interiowo.pl/zsiw-egzamin.html> [ostatni dostęp: 24.01.2016], a o takim powołaniu informuje nas przypis 72.

do ostatniego dziesięciolecia" (s. 24). Dlatego też analizę piśmiennictwa autorka postanowiła uzupełnić przeglądem witryn internetowych bibliotek i bibliografii narodowych, a także wynikami autorskich badań ankietowych, przeprowadzonych w 2011 r. Niestety, ankieta spotkała się ze stosunkowo małym odzewem (18 odpowiedzi na 34 rozesłane formularze; w siedmiu przypadkach odesłano autorkę do witryn internetowych serwisów bibliograficznych). Co więcej, choć te szcztkowe wyniki posłużyły do opracowania niektórych zestawień (m.in. Tab. 4–5), do pracy nie dołączono formularza ankiety, stąd nie wiadomo, na jakie pytania autorka poszukiwała odpowiedzi. Autorka przyznaje także, że samodzielnie uzupełniała braki w odpowiedziach na drodze własnych poszukiwań w dostępnych źródłach i opracowaniach (s. 146, przypis 486). Tu pojawia się pytanie o zasadność przeprowadzenia takich badań i jakość ich wyników, kiedy nie wiadomo, jakie kwestie nurtowały autorkę oraz w jakim stopniu odpowiedzi pochodziły od ankietowanych, a w jakim – od samego ankietera.

Z powyższego wynika, że celem pracy było wypełnienie dotkliwej luki badawczej, jaką był brak analizy wykorzystywania międzynarodowych zaleceń i wytycznych w organizacji pracy bibliograficznej w ostatnim półwieczu; oprócz tego autorka chciała wskazać „kierunek, w jakim podąży współczesna informacja bibliograficzna" (s. 20), choć trzeba przyznać, że ten aspekt nie został zbyt wyeksponowany. Ponadto, o ile proces kształtowania się i ewolucji modelu autorka przedstawiła wzorowo i bardzo skrupulatnie, to jednak praca, ukazując się pod przyjętym tytułem, powinna być wzbogacona także o to, co z tego kształtowania się wyłoniło – o opis owego współczesnego światowego modelu informacji bibliograficznej; tego w pracy zabrakło. Autorka ostatecznie jednak dowodzi, że współpraca bibliograficzna jest coraz ściślejsza i mimo pierwotnych założeń pewnej swobody i różnicowania, ostatecznie wypracowano jednolite systemy współpracy w zakresie opracowania i informowania o dokumentach.

Niestety, choć merytorycznie, mimo wyżej opisanych niedociągnięć, praca jest bardzo dobra, to nie została należycie potraktowana na etapie prac redakcyjnych. Daje się w niej zauważyć szereg usterek, które choć nie uniemożliwiają właściwego korzystania z dzieła, to jednak odbijają się negatywnie na obcowaniu z nim. Wynika to, jak się zdaje, z dwóch przyczyn: brak czujności przy przeredagowywaniu dysertacji do druku oraz niedostateczna kontrola redakcyjna wydawnictwa. Praca ma rozbudowaną strukturę, stąd przy nawigacji pomocna byłaby dobra żywa pagina. Niestety, zamiast informować o bieżącym rozdziale i podrozdziale (a trzeba jeszcze wiedzieć, że niektóre z nich zawierają podrozdziały drugiego, a nawet trzeciego stopnia), informuje o rozdziale i... tytule książki, co jest absolutnie zbędne. Skracanie powszechnych w bibliologii terminów wielowyrazowych również przebiegło niejednolicie, przykładowo bieżąca bibliografia narodowa została skrócona w *Wykazie skrótów* minuskułą (bbn), zaś narodowa centrala bibliograficzna – majuskułą (NCB). Wykaz ten jest również niepełny – zabrakło choćby skrótu LISTA czy UNESCO – oraz niekonsekwentny: podaje rozwinięcia skrótów komitetów Międzynarodowej Rady Archiwów, ale podobnych rangą komitetów Międzynarodowej Federacji Dokumentacji już nie. Z kolei na s. 41 powołany jest artykuł z *Informatora Bibliotekarza i Księgarza*, którego tytuł nie został skrócony i stąd brakuje go w *Wykazie skrótów tytułów czasopism*. We *Wstępie*, z nieznanых przyczyn, obok przypisów dolnych występują przypisy harwardzkie, a niekiedy do przypisu harwardzkiego dołączony jest przypis dolny (na szczęście w dalszych rozdziałach pojawiają się konsekwentnie przypisy dolne). W rozdziale 2 tłumaczenia nazw kolejnych komitetów w ramach Międzynarodowej Federacji Dokumentacji są pisane raz z wielkiej, raz z małej litery. W wątpliwość trzeba także podać sens fonetycznego zapisu instytucji i programów w oryginale zapisywanych cyrylicą (43–44), kiedy w *Wykazie skrótów* są one transliterowane oraz współtowarzyszą im tłumaczenia polskie (np. MSNTI vs. MSINT). Takich niekonsekwencji jest mnóstwo, pojawiają się także – i to w liczbie przekraczającej poziom przyzwoitości – literówki, a nawet błędne dane bibliograficzne (w tekście ciągłym na s. 24 praca Jarosława Packa została datowana na rok 2011 zamiast na 2010). Konstrukcja *Indeksu osobowego i przedmiotowego* też jest uciążliwa, albowiem nie zaznaczono wcięciami jednostek stanowiących część większych organizacji. Sprawia to, że na pierwszy rzut oka indeks nie zachowuje porządku alfabetycznego, gdyż np. po „Biblioteka

Brytyjska” następuje „Lending Division”, a po nim „Biblioteka Cambridge University”. Ostatnie słowo krytyki należy się konstrukcji tabel nr 2–4 oraz 7–10, które nie konfrontują dwóch cech i są w istocie alfabetycznymi wyliczeniami. Powoduje to bezskuteczne próby poszukiwania ukrytego sensu konstrukcji tych tabel, którego w istocie nie ma.

Lista uchybień redakcyjnych jest długa, ale to przede wszystkim nowatorski charakter książki daje nadzieję na rychłe drugie wydanie *Światowego modelu informacji bibliograficznej*. Bo nie można wątpić, że będzie ono potrzebne – nakład tej pozycji, z uwagi na świetny poziom merytoryczny oraz zaspokojenie potrzeby nowej refleksji w nauce o bibliografii, powinien zostać szybko wyczerpany, gdyż można ją polecić wszystkim, którzy interesują się nowoczesną bibliografią.

Jakub Maciej Łubocki

*Instytut Informacji Naukowej i Bibliotekoznawstwa, Uniwersytet Wrocławski*

*Nadesłano: 25 lutego 2016.*

**Mariusz Jarocki:** *Open Source w bibliotekach w świetle badań publicznych uczelni akademickich w Polsce*. Toruń: Wydaw. Naukowe Uniwersytetu Mikołaja Kopernika, 2015, ss. 300. ISBN: 978-83-231-3479-4

Książka Mariusza Jarockiego pt. *Open Source w bibliotekach w świetle badań publicznych uczelni akademickich w Polsce* jest jedną z niewielu polskich publikacji traktujących o zjawisku Otwartego Źródła w kontekście działalności instytucji bibliotecznych. Książka stanowi skróconą wersję dysertacji doktorskiej noszącej tytuł *Zastosowania oprogramowania Open Source w bibliotekach publicznych uczelni akademickich w Polsce* obronionej w 2014 r. na Wydziale Nauk Historycznych UMK w Toruniu. Warto zaznaczyć, że wydana drukiem praca to także pierwsza publikacja z zakresu bibliologii i informacji naukowej, która koncentruje się nie tylko na genezie i historii oprogramowania Open Source, jego charakterystyce, ale stanowi przede wszystkim przejrzyste kompendium wiedzy na temat możliwych do zastosowania – zwłaszcza w uczelnianych bibliotekach – tworzonych w ramach tego zjawiska aplikacji. Dodajmy, że w 2014 r. nakładem Wydawnictwa SBP ukazała się praca Sebastiana D. Kotuły pt. *Wstęp do Open Source*, będąca syntetycznym wprowadzeniem do zagadnienia Open Source, także z perspektywy problemów będących częścią informatologii i bibliologii. Książka M. Jarockiego jest doskonałym rozwinięciem tez i charakterystyk przygotowanych przez S. D. Kotułę, ujętym z punktu widzenia systemów i aplikacji przeznaczonych dla działalności bibliotek. Nie będzie chyba nadużyciem, jeśli stwierdzę, że te dwie prace uzupełniają się wzajemnie. Książka S. D. Kotuły przybliża genezę, historię ruchu Open Source, zawiera szczegółową charakterystykę oprogramowania (także z punktu widzenia porównań z komercyjnymi rozwiązaniami programistycznymi). Jest źródłem informacji dla osób chcących nabyć wstępną wiedzę na temat wybranych informatycznych rozwiązań. Stanowi, jak zaznacza zresztą sam autor, wprowadzenie do zagadnienia Open Source w ogóle. W przypadku pracy M. Jarockiego, mamy natomiast do czynienia z bardziej szczegółową charakterystyką różnorodności oprogramowania, przedstawioną z perspektywy jego wykorzystania w poszczególnych procesach bibliotecznych, zarówno podstawowych, jak i pomocniczych. Co więcej, intencją autora recenzowanej pracy było przygotowanie wskazówek dla bibliotek, które swoje działania oparły właśnie na systemach otwartych. W związku z tym M. Jarocki, na podstawie przeprowadzonych badań i analiz, nie tylko ukazał schemat aktualnego wykorzystania otwartego oprogramowania w badanych jednostkach, ale zaprezentował

także własną propozycję modelu automatyzacji działania biblioteki publicznej uczelni akademickiej w oparciu o rozwiązania bezpośrednio odnoszące się do omawianego w pracy zjawiska. Model ten stanowi przy tym wskazówkę dotyczącą możliwości wykorzystania różnego typu oprogramowania z perspektywy wykonywanych w bibliotece działań. Celem rozprawy doktorskiej była więc analiza rynku i wdrożenia poszczególnych programów oraz systemów z punktu widzenia działalności polskich księżnic. Na podstawie własnych obserwacji i przeglądu poszczególnych narzędzi autor pokazał także, w których obszarach tej działalności otwarte programy i systemy mają najczęstsze zastosowanie. M. Jarocki skupił się na funkcjonowaniu polskich bibliotek publicznych uczelni akademickich. Powody takiego wyboru są jak najbardziej zrozumiałe. W bibliotekach tego typu, na tle wszystkich instytucji książki, komputeryzacja została przeprowadzona stosunkowo najwcześniej. Co więcej, zazwyczaj posiadają one w swojej strukturze organizacyjnej dział ds. komputeryzacji lub informatyzacji, w którym zatrudnione są osoby posiadające odpowiednie kompetencje w zakresie instalacji, konfiguracji i administracji poszczególnymi otwartymi systemami czy programami. Jest to bardzo ważny argument. Instytucje te, pomimo, że korzystają z oprogramowania z otwartym kodem źródłowym, muszą brać pod uwagę trudności związane z samą implementacją i występującymi w późniejszym czasie problemami w użytkowaniu poszczególnych programistycznych rozwiązań. W tym przypadku mamy do czynienia z sytuacją, kiedy dostęp do aplikacji czy systemów Open Source jest najczęściej bezpłatny, ale prace mające na celu ich wdrożenie i wykorzystanie zawsze generują pewne koszty. M. Jarocki zastanawia się przy tym nad intensywnością wykorzystania tego typu rozwiązań przez biblioteki publicznych uczelni akademickich. Stawia też tezę, że większość tych instytucji implementuje i wykorzystuje aplikacje Open Source głównie jako wsparcie pomocniczych procesów bibliotecznych. Podstawowe procesy (gromadzenie zbiorów, budowanie katalogów, opracowanie dokumentów, udostępnianie zasobów, wyszukiwanie i przekazywanie informacji) są wspierane przez oprogramowanie Open Source tylko w największych tego typu placówkach. Należy zaznaczyć, że autor wziął pod uwagę oprogramowanie działające w oparciu o systemy Windows i Linux, a pominął system Mac OS. Powodem takiej decyzji była mała – według autora – popularność tego systemu w Polsce. Jednak wydaje się, że urządzenia oferowane przez firmę Apple również w naszym kraju są wykorzystywane przez coraz większą liczbę użytkowników. Wiele na to wskazuje, że w najbliższych latach firma ta może także stać się jednym z głównych graczy na polskim rynku informatycznym i komputerowym. Oferowane przez nią tablety już dzisiaj są wykorzystywane przez pracowników bibliotek i czytelników na przykład w ramach realizowanego przez Fundację Rozwoju Społeczeństwa Informacyjnego projektu „Tablety w Twojej Bibliotece” (w 2015 r. przekazano ponad 1250 tabletów iPad 318 bibliotekom).

Recenzowana książka składa się z czterech głównych części, które zostały przygotowane z punktu widzenia analizowanych i charakteryzowanych w pracy problemów. Pierwsza, najkrótsza część, dotyczy samego zjawiska Open Source – jego genezy, definicji najważniejszych terminów oraz historii ewolucji społeczności tego ruchu. Autor słusznie wskazał na dwa, z pozoru bardzo podobne terminy – wolne oprogramowanie oraz oprogramowanie o otwartym kodzie źródłowym. W przypadku tego drugiego, nie do końca prawidłowo został przetłumaczony tytuł pierwszego warunku, który musi być spełniony, aby program mógł być uważany za Open Source. Chodzi mianowicie o zasadę, która w definicji Otwartego Źródła w języku angielskim brzmi *free redistribution*. Nie chodzi tutaj jednakże o darmową, ale wolną redystrybucję, tzn. licencja nie może wprowadzać zakazów dotyczących sprzedaży lub darmowego przekazywania oprogramowania jako części pewnej całości. Dobrze ten mankament wyjaśnia S. D. Kotuła we wcześniej wspomnianej książce. M. Jarocki słusznie natomiast wskazał na różnice pomiędzy oprogramowaniem należącym do tzw. *free software* a rozwiązaniami typu Open Source. Niuanse wynikające z tego porównania wydają się być źródłem częstych błędów popełnianych przez osoby, które niedokładnie rozumieją istotę zjawisk dotyczących otwartości w dzisiejszym udostępnianiu, a może nawet bardziej dzieleniu się różnego typu cyfrowymi wytworami, w tym aplikacjami stworzonymi czy współtworzonymi w ramach współpracy użytkowników.

Słuszne wydaje się także wskazanie wad oprogramowania Open Source, szczególnie z punktu widzenia kosztów wdrożenia, wsparcia i serwisu dla użytkownika końcowego. Przydatny dla zainteresowanego tym zjawiskiem czytelnika okaże się na pewno wykaz źródeł informacji dotyczących społeczności skupionych wokół poszczególnych projektów, w tym zwłaszcza krótkie charakterystyki serwisów i portali związanych z ruchem Open Source.

Druga i trzecia część książki dotyczy systemów przeznaczonych do automatyzacji procesów bibliotecznych. Znajdziemy tutaj szczegółowy wykaz i syntetyczną charakterystykę przywołanych przez autora rozwiązań. Drugi rozdział koncentruje się na podstawowych procesach bibliotecznych, w tym, oczywiście, tych dotyczących gromadzenia, opracowania i udostępniania zbiorów. M. Jaroński przedstawia najczęściej wykorzystywane i na potrzeby pracy sprawdzone rozwiązania (z pominięciem programów nieaktualizowanych lub przygotowanych dla systemu Mac OS). Źródłem wiedzy służącej znajdowaniu informacji o oprogramowaniu były dla autora: fachowa literatura przedmiotu oraz specjalistyczne serwisy internetowe bezpośrednio powiązane z problematyką Open Source i rozwiązaniami programistycznymi wdrażanymi w bibliotekach. W opisie zastosowanej metodologii brakuje jednak informacji na temat kwerend odnoszących się do prowadzonych poszukiwań wykorzystanej literatury. Autor korzystał zapewne z różnego rodzaju baz pełnotekstowych czy cyfrowych repozytoriów i bibliotek. Przydałoby się je tutaj przynajmniej wymienić. Tym bardziej, że w tekście książki pojawia się informacja o doborze oprogramowania m.in. na podstawie prostych zapytań zadanych za pośrednictwem wyszukiwarki Google.

Rozdział drugi, dotyczący systemów przeznaczonych dla podstawowych procesów bibliotecznych został podzielony na cztery podrozdziały, w których autor przedstawił informacje na temat oprogramowania, które może być wykorzystywane przez biblioteki w ich codziennych pracach. Znajdziemy więc tutaj charakterystykę zintegrowanych systemów bibliotecznych, oprogramowania do zarządzania zasobami cyfrowymi czy wspierającego procesy wyszukiwania dokumentów. Na szczególną uwagę zasługuje zwłaszcza charakterystyka platform i programów służących do budowy cyfrowych bibliotek i repozytoriów. Nie ukazała się do tej pory w języku polskim żadna inna publikacja, która zawierałaby bardziej kompletną bazę informacji dotyczących oprogramowania Open Source przeznaczonego do prezentacji i zarządzania kolekcjami cyfrowymi. W przypadku zautomatyzowanych systemów bibliotecznych, na uwagę zasługują zwłaszcza charakterystyki popularnych dzisiaj systemów Evergreen i Koha. Przydałoby się jednak nieco więcej informacji odnoszących się do wdrożeń Koha w polskich instytucjach. Jest ich dzisiaj całkiem sporo, nie tylko zresztą w przypadku bibliotek akademickich. W tym rozdziale książki M. Jaroński posłużył się jedynie kilkoma przykładami wdrożeń z bibliotek zagranicznych. Ciekawie za to prezentuje się podrozdział dotyczący narzędzi wspierających proces wyszukiwania dokumentów udostępnianych w bibliotekach. Autor przedstawił poszczególne rozwiązania słusznie zresztą przyporządkowując je do trzech rodzajów programów (narzędzi indeksujących, katalogów OPAC oraz interfejsów informatyczno-wyszukiwawczych). Dobrym uzupełnieniem rozważań na temat narzędzi wspierających podstawowe procesy biblioteczne jest podrozdział dotyczący zarządzania bibliografią i cytowaniami. M. Jaroński skupił się przy tym na dwóch typach rozwiązań: bibliograficznych bazach danych oraz programach umożliwiających zarządzanie informacją bibliograficzną i cytowaniami. Zwłaszcza to drugie rozwiązanie jest dzisiaj coraz częściej postrzegane jako istotne wsparcie prac mających na celu gromadzenie i łatwe zarządzanie informacjami o dokumentach, ale także – samymi dokumentami. Aspekt ten może też dotyczyć możliwości tworzenia przez użytkownika swojej osobistej elektronicznej biblioteki złożonej z wyselekcjonowanych przez niego cyfrowych zasobów. Co więcej, oprogramowanie to służy z jednej strony, do zarządzania zgromadzonymi informacjami bibliograficznymi oraz zapewnienia jednolitego formatu cytowania w przygotowywanych dokumentach tekstowych (np. z wykorzystaniem edytora Microsoft Word), jak również do dzielenia się gromadzonymi danymi z innymi użytkownikami i łatwej współpracy przy tworzeniu tematycznych zbiorów informacji bibliograficznych. Takie możliwości oferuje na przykład charakteryzowane przez M. Jarońskiego Zotero, program będący jednocześnie

usługą sieciową wykorzystywaną przez użytkowników z różnych stron świata w celu zarządzania gromadzoną przez nich informacją bibliograficzną.

Trzeci rozdział książki dotyczy szerokiego spektrum oprogramowania Open Source, które może być wykorzystane do realizacji bibliotecznych procesów pomocniczych. Poszczególne systemy i aplikacje zostały przyporządkowane do pięciu kategorii rozwiązań. M. Jarocki zwrócił więc tutaj szczególną uwagę na systemy operacyjne, bazy danych, języki programowania, systemy zarządzania treścią oraz wybrane aplikacje użytkowe (w tym pakiety biurowe, aplikacje wspierające pracę z plikami graficznymi, muzycznymi, wideo, komunikatory internetowe, przeglądarki internetowe itp.). Ta część pracy jest najbardziej ogólna, tzn. w swojej istocie omawiane rozwiązania najmniej koncentrują się na działalności bibliotek. Jednakże, ze względu na szeroki wachlarz możliwości wykorzystania omawianego oprogramowania, także z punktu widzenia przetwarzania multimedialnych dokumentów, komunikacji czy zarządzania i administracji pracy instytucji bibliotecznych, rozdział ten stanowi najobszerniejszą część recenzowanej książki. Popularność i dostępność różnorodnych rozwiązań informatycznych dotyczących zjawiska Open Source utrudniały zapewne wybór oprogramowania poddanego przez autora dalszym analizom. M. Jarocki oprócz wyszukiwania danych za pośrednictwem najpopularniejszej wyszukiwarki, kwerend w fachowej literaturze czy specjalistycznych serwisach internetowych gromadzących informacje o programach Open Source, posłużył się w tym celu autorską ankietą, której wyniki szczegółowo zostały przedstawione w ostatnim rozdziale książki (*Oprogramowanie Open Source w bibliotekach publicznych uczelni akademickich w Polsce – wyniki autorskich badań ankietowych*). Przedmiotem tych badań, jak wskazuje autor, „były wybrane aspekty technologii informatycznej wspierającej usługi i procesy zachodzące w publicznych bibliotekach akademickich”. Wśród tez badawczych postawionych przez M. Jarockiego znajdują się założenia dotyczące częstotliwości i popularności wykorzystania poszczególnych rozwiązań, wpływu na korzystanie z nich czynnika finansowego, jak również tezy odnoszące się do kwestii edycji, współtworzenia i przystosowania przez bibliotekarzy do swoich potrzeb kodu źródłowego oprogramowania. Autor słusznie założył, że większość planowanych wdrożeń programów Open Source w bibliotekach jest powiązana z udostępnieniem wprowadzanych usług w trybie online. Zrozumiałe wydaje się też założenie dotyczące wdrażania projektów Open Source w kontekście wykorzystywania systemów informacyjnych, narzędziowych czy systemów operacyjnych (serwerowych).

Jak wcześniej wspomniano, aby zweryfikować postawione tezy, autor posłużył się metodą sondażu diagnostycznego, a kwestionariusz wykorzystanej w nim ankiety jest zamieszczony na końcu książki. Warto zaznaczyć, że M. Jarocki oprócz rozsyłania ankiet drogą elektroniczną, wykorzystał także technikę wywiadu i bezpośredni kontakt z wybranymi pracownikami bibliotek. Przy czym prowadzący badanie, z różnych względów, także tych dotyczących bezpieczeństwa bibliotecznych systemów, słusznie skorzystał z anonimowej formy udostępniania ankiet. Na treść ankiety składały się pytania otwarte, półotwarte oraz zamknięte, co umożliwiło autorowi przeprowadzanie analiz ilościowych i jakościowych. Bardzo istotny był także podział ankiety na trzy części, z których jedna była skierowana do osób w ogóle nie wykorzystujących rozwiązań Open Source. Obiektem badań były największe biblioteki akademickie w Polsce (odpowiedzi udzieliło 55 instytucji), w których, jak się okazało w trakcie badania, rozwiązania Open Source są rzeczywiście popularne. Nie jest też zaskoczeniem, że większość bibliotek wybiera je ze względu na niższe koszty wdrażania oraz brak konieczności przeprowadzania procedury przetargowej. Potwierdziła się też teza autora, że najczęściej wykorzystywane jest oprogramowanie użytkowe i narzędziowe, choć znajdziemy też przypadki implementacji zintegrowanych systemów bibliotecznych, tworzenia na bazie tych rozwiązań bibliotek cyfrowych, repozytoriów czy platform e-learningowych. Większość bibliotekarzy uważa również, że idea ruchu Open Source znajduje zrozumienie w środowisku bibliotekarskim, a wdrażanie w bibliotekach towarzyszących jej informatycznym rozwiązaniom wydaje się dzisiaj niezbędne.

Książka M. Jarockiego to bardzo ważna pozycja na rynku wydawniczym. Bez wątpienia ma duże znaczenie zarówno dla wyjaśnienia czym jest samo zjawisko Open Source, jak i dla poznania

możliwości oferowanych dzisiaj przez oprogramowanie tego typu instytucjom bibliotecznym. Polskie środowisko bibliotekarskie i naukowe czeka na tego typu opracowania aktualnych i kluczowych dla bibliotek zagadnień. Nie mam co do tego żadnych wątpliwości. Szybka ewolucja nowych technologii i towarzyszących im zjawisk powoduje konieczność ciągłego przygotowywania przez badaczy zarówno teoretycznych, jak i także praktycznych kompendiów wiedzy.

Praca nie posiada niestety indeksu. Jestem przekonany, że byłby on doskonałym uzupełnieniem treści, bardzo przydatnym dla czytelników. Książka, co ważne, dostępna jest w wersji elektronicznej, m.in. w wirtualnej czytelnicy iBuk Libra, z której korzysta dzisiaj duża liczba polskich bibliotek.

Reasumując, wartość pracy M. Jarockiego oceniam bardzo wysoko. Książka jest doskonałym źródłem wiedzy dla osób poszukujących informacji na temat zjawiska Open Source, jak i przykładów funkcjonalnego i użytecznego dla bibliotek oprogramowania. Może też w przyszłości stanowić materiał porównawczy dla podobnych analiz.

Grzegorz Gmiterek

Katedra Bibliografii i Dokumentacji  
Wydział Dziennikarstwa, Informacji i Bibliologii  
Uniwersytetu Warszawskiego

Nadesłano: 21 lipca 2016.

**Małgorzata Kowalska: Crowdsourcing internetowy – pozytywny wymiar partycypacji społecznej. Konteksty – istota – uwarunkowania.** Warszawa: Wydaw. SBP, 2016, ss. 366. Seria „Nauka, Dydaktyka, Praktyka”; 165. ISBN 978-83-64203-60-2

Znacząco rośnie liczba naukowych tekstów z różnych dziedzin wiedzy, w których pojawiają się takie określenia jak: zbiorowa inteligencja, mądrość tłumu, inteligencja współpracy, wikinomia, otwarta innowacja, prosumpcja, crowdsourcing, crowdfunding, produkcja partnerska i in. Podkreśla się w nich pewien wymiar nowej ekonomii, przede wszystkim aspekt biznesowy i marketingowy.

Książka Małgorzaty Kowalskiej – *Crowdsourcing internetowy – pozytywny wymiar partycypacji społecznej. Konteksty – istota – uwarunkowania*, wydana przez Stowarzyszenie Bibliotekarzy Polskich, jest włączeniem się w żywo toczący się naukowy dyskurs dotyczący kierunków rozwoju nowoczesnych społeczeństw, w których nie same technologie, lecz ich inteligentne użycie zdecyduje o ich kondycji. Na tle innych publikacji, dotyczących partycypacyjnego modelu współpracy użytkowników sieci, monografię M. Kowalskiej odróżnia staranne uporządkowanie i doprecyzowanie zakresów znaczeniowych wielu terminów powiązanych z prezentowanym fenomenem, a także ułożenie szeroko zakrojonej tematyki crowdsourcingu w obszarze bibliologii i nauki o informacji.

Wybór tematu badań należy uznać za bardzo trafny, szczególnie w okresie konwergencji mediów i kształtowania się wzorów zachowań społecznych nowych technicyzowanych społeczeństw. Znacząco rośnie bowiem zainteresowanie inicjatywami angażującymi ochotników z sieci. Potwierdza to wzrastająca liczba ukazujących się na ten temat publikacji, jak i analiza zapytań kierowanych z hasłem „crowdsourcing” do wyszukiwarki Google, na co wskazuje także autorka.

Książka została starannie przemyślana, co odzwierciedla kompozycja i ujęcie rozprawy. M. Kowalska podzieliła tekst na dwie zasadnicze części: część I. *Otoczenie crowdsourcingu* jest wprowadzeniem i bazą do części II. skupiającej się na wyjaśnieniu samego fenomenu crowdsourcingu.

Zarówno technologiczne, jak i ekonomiczne otoczenie crowdsourcingu oraz jego społeczne i kulturowe konteksty zanalizowane przez autorkę w części pierwszej prowadzą do wyjaśnienia warunków sprzyjających powstawaniu nowych modeli partycypacji społecznej. Śledząc rozwijanie

się idei wspólnotowości – poprzez pokazanie i interpretację licznych koncepcji filozoficznych, socjologicznych, z zakresu psychologii społecznej, psychologii tłumu – autorka dochodzi do czasów społeczności wirtualnych, których warunkiem istnienia jest infrastruktura technologiczna (cyberprzestrzeń) oraz potrzeba relacji międzyludzkich. W dobie konwergencji mediów społecznościowa współpraca, stanowiąca istotny paradygmat usieciowionej gospodarki i nowy model biznesowy oraz charakterystyczny dla całej rzeczywistości społeczno-kulturalnej model kultury partycypacji zaznaczają się szczególnie wyraziście. Bez względu na sposób podejścia do determinantów kultury partycypacji „jej istota leży w zacieraniu się granicy między producentem a konsumentem dóbr kultury” (s.111), a wytwory kultury uczestnictwa „współtworzą kapitał społeczny, stając się źródłem wiedzy, pomysłów, okazji i wsparcia” (s.112). Chociaż badacze dokonują rozróżnienia pomiędzy inteligencją współpracy a inteligencją zbiorową (różnica wynika z innego poziomu zaangażowania użytkowników sieci), to niektórzy skłonni są traktować taką kolektywną inteligencję jako alternatywę dla dominującej do tej pory wiedzy eksperckiej.

Część druga książki ukazuje fenomen crowdsourcingu, a *de facto* fenomen tłumu (ang. *crowd-sourcing*) i jego siłę sprawczą. Autorka odnosi się tu do najważniejszych koncepcji teoretycznych i terminologicznych, a w szczególności prezentuje poglądy Jeffa Howe’a – autora terminu *crowdsourcing* – oraz Darrena C.Brabhama, który wprowadził rozważania na temat *crowdsourcingu* do literatury naukowej; dokonuje także zestawienia wszystkich dostępnych w polskiej literaturze definicji badanego zjawiska. W rozdziale drugim podejmuje próbę konceptualizacji *crowdsourcingu* internetowego z odniesieniem do takich determinantów badanego zjawiska, jak: *crowdsourcer* (inicjator), tłum oraz elektroniczna platforma, a także prezentuje różnorodne propozycje taksonomiczne *crowdsourcingu* internetowego. Rozdział trzeci, lokujący *crowdsourcing* w kręgu zainteresowań bibliologii i informatologii, może być najtrudniejszy w odbiorze dla czytelnika spoza kręgów naukowych – prezentacja wybranej metodologii do zbadania obecności problematyki *crowdsourcingu* w literaturze przedmiotu nasycona jest tytułami publikacji, czasopism, liczbami i procentami. Z całą pewnością jednak zaprezentowane postępowanie badawcze i uzyskiwane na poszczególnych etapach rezultaty są bardzo ważne dla pokazania kryteriów doboru światowej i polskiej bazy źródłowej poszukiwań prowadzonych przez autorkę. Warto w tym miejscu zasygnalizować, że w typowaniu czasopism do badań, M. Kowalska wykorzystała metodologię mieszaną: ilościową (bibliometryczną) i jakościową (analizy i krytyki piśmiennictwa). Podstawowym narzędziem wyboru wiodącego czasopiśmiennictwa zagranicznego była baza Journal Citation Report (JCR), a w zasadzie wariant JCR SSE (JCR Social Science Edition). Zagraniczny materiał badawczy został zebrany na platformie Web of Science Core Collection w bazach: Social Science Index oraz Arts and Humanities Index, a uzupełniając na platformie EBSCOhost w bazie LISTA (Library, Information Science and Technology Abstracts). Do wytypowania wiodących czasopism polskich z zakresu bibliologii i informatologii wykorzystana została część B „Listy czasopism punktowanych” przygotowanej przez Ministerstwo Nauki i Szkolnictwa Wyższego, a zebrany materiał uzupełniono danymi z Polskiej Bibliografii Bibliologicznej oraz Bibliografii zawartości Czasopism.

Zgromadzony zbiór badawczy unaoczniał stosunkowo późne zainteresowanie tematyką *crowdsourcingu* światowych, a tym bardziej polskich, przedstawicieli bibliologii i informatologii. Jest to zasób zróżnicowany pod względem tematycznym – najliczniej w literaturze zagranicznej odnosi się do kategorii określonej przez autorkę jako „*crowdsourcing* jako mechanizm generowania, kontekstualizacji, optymalizacji i oceny jakości treści, danych, zasobów i systemów informacyjnych” (s. 194).

*Crowdsourcing* jako sposób rozwiązywania konkretnych problemów zdominował ostatni, czwarty rozdział rozprawy, w którym zostało przedstawione 21 wybranych przykładów wdrożeń realizowanych przez biblioteki zagraniczne w następujących obszarach zastosowań: transkrypcja; katalogowanie, kategoryzacja, kontekstualizacja; kolaboratywne tagowanie i linkowanie; generowanie treści; nadawanie georeferencji; komentowanie, opiniowanie, rekomendowanie oraz społecznościowe pozyskiwanie funduszy. Przejawy *crowdsourcingu* – jako szeroko rozumianej partycypacji użytkowników bibliotek

polskich – M. Kowalska starała się pokazać na przykładzie współtworzenia zasobów cyfrowych, kółkowskiego tagowania, komentowania i oceniania zasobów, a także społecznościowego pozyskiwania funduszy (crowdfunding), chociaż – jak stwierdza autorka – trudno jest wskazać inicjatywy realizowane „z takim rozmachem i w takim wymiarze jak ma to miejsce w innych krajach świata” (s. 251).

Warto powtórzyć za autorką, że „(...) doświadczenia bibliotek i innych instytucji sektora kultury w zakresie wdrażania crowdsourcingu dowodzą, że niesie on przede wszystkim pozytywne implikacje – zarówno natury społecznej, jak i ekonomicznej. Z jednej strony bowiem pozwala na budowę wirtualnych społeczności i grup użytkowników skupionych wokół biblioteki, a co za tym idzie – lepsze dostosowywanie oferty usługowej do potrzeb użytkowników i zwiększenie stopnia wykorzystywania zbiorów, z drugiej – na obniżenie kosztów działalności oraz szybsze osiągnięcie celów wymagających niekiedy zaangażowania dużych nakładów finansowych i kadrowych...” (s. 224). Problemem jest jednak niewystarczająca obecność tego typu inicjatyw w świadomości społecznej – ogranicza ją jednojęzyczność oraz brak promocji. Autorkę książki zainteresowała kwestia obecności crowdsourcingu internetowego w świadomości bibliotekarzy i pracowników innych instytucji ochrony dziedzictwa kulturowego. Dotarła więc do badań ankietowych amerykańskiej organizacji CrowdConsortium z 2014 i 2015 r. prowadzonych wśród respondentów (w ponad 75% bibliotekarzy) z całego świata (głównie Stanów Zjednoczonych). Badania te dowodzą istnienia szeregu problemów organizacyjnych, prawnych, finansowych oraz mentalnych. Niezadowolająco wypadły także autorskie próbkowe badania realizowane w środowisku polskim w trakcie webinarium poświęconych crowdfundingowi oraz crowdsourcingowi. Za jedyną możliwą do realizacji w bibliotekach formę crowdsourcingu bibliotekarze uznają finansowanie społecznościowe.

Podsumowując rozdział *Crowdsourcing internetowy w działalności bibliotecznej*, można powiedzieć, że przeniesienie badanego fenomenu w ramy działalności bibliotecznej jest prawie całkowitym *novum* na polskim gruncie biblio- i informatologicznym, poza kilkoma wcześniejszymi artykułami wskazanymi przez autorkę. Egzemplifikacja możliwych zastosowań crowdsourcingu w działaniach bibliotecznych oraz innych instytucji kultury jest istotnym wkładem autorki w rozwój teorii i praktyki bibliotecznej oraz innych instytucji kultury XXI wieku, co należy dodatkowo podnieść na tle bardzo słabo dotychczas reprezentowanej tematyki w polskich periodykach. Chociaż rzetelność naukowa każe autorce wskazywać zarówno pozytywne, jak i negatywne implikacje wynikające z praktycznej implementacji crowdsourcingu, M. Kowalska zdecydowanie staje po stronie jego zwolenników i propagatorów, co zresztą uwidacznia w tytule pracy.

Autorka prezentuje świetnie opanowany warsztat badawczy. Swobodnie porusza się po literaturze anglojęzycznej – z dużą dojrzałością badacza wykorzystuje metodę krytycznej analizy piśmiennictwa – poddaje analizie i krytycznej refleksji blisko 300 zidentyfikowanych przez nią artykułów zagranicznych i kilkanaście polskich, sięga do metod bibliometrii. Z uznaniem należy odnieść się do naprawdę sporej liczby przypisów – nie tylko bibliograficznych, ale również komentujących, wyjaśniających, faktograficznych. Rozprawa oparta jest więc na solidnej wiedzy i postępowaniu badawczym. Rozważania teoretyczne, z wykorzystaniem bogatego zbioru starannie dobranych publikacji, poparte są znaczącą liczbą przykładów konkretnych implementacji crowdsourcingu, co dodatkowo podnosi ocenę pracy. Starannie przygotowaną *Bibliografię*, liczącą w sumie 35 stron, badaczka podzieliła na dwie części: *Publikacje drukowane* oraz *Dokumenty elektroniczne*, co z uwagi na wykorzystany materiał i sposób opisu nie budzi zastrzeżeń.

Do zasadniczego tekstu liczącego 285 stron oraz *Bibliografii* autorka dołączyła także dwa aneksy oraz indeks rzeczowy. Aneks 1. *Wykaz publikacji na temat crowdsourcingu poddanych analizie bibliometrycznej* badaczka podzieliła na obszary tematyczne wyróżnione podczas analizy. Zabieg ten należy ocenić pozytywnie, gdyż ułatwić może samodzielne studiowanie literatury i pogłębianie wiedzy w wybranych kontekstach. Aneks 2. *Lista bibliotecznych projektów crowdsourcingowych* – z podziałem na projekty zagraniczne i polskie – spotka się zapewne z uznaniem i ze znacznym zainteresowaniem bibliotekarzy, pracowników informacji oraz społeczności instytucji kultury.

Należy dodać, że praca została także we właściwych proporcjach ubogacona graficznie, a przejrzysty *Spis tabel, rysunków i wykresów* czytelnicy znajdą przed indeksem rzeczowym na końcu książki.

Książkę M. Kowalskiej należy uznać za bardzo wartościową pozycję w repertuarze Wydawnictwa SBP. Jest to znakomite kompendium wiedzy na temat tego, jak współpraca wirtualnych społeczności może zmieniać świat, w tym świat bibliotek i innych instytucji kultury. *Crowdsourcing internetowy – pozytywny wymiar partycypacji społecznej. Konteksty – istota – uwarunkowania* to książka dla tych, których interesują społeczności online (kreatywny tłum) i drzemący w nich kapitał intelektualny i wielka sprawcza siła; to publikacja dla tych, którzy chcą znaleźć przykłady pozytywnego wykorzystania działania tłumu, wirtualnych społeczności, partnerskiej produkcji, zbiorowej inteligencji, mądrości tłumu, otwartej innowacji, inteligencji współpracy; dla tych, którzy chcą lepiej poznać wymiary współczesnej wikipedii poprzez crowdsourcing właśnie oraz crowdfunding i outsourcing; dla tych wszystkich, którzy w sieci Web. 2.0 chcą być prosumentami informacji lub pracować z nimi na rzecz rozwoju społeczeństw informacji i wiedzy.

Katarzyna Materska

*Biblioteka Główna*

*Uniwersytet Kardynała Stefana Wyszyńskiego*

*Nadesłano: 25 lipca 2016.*

## Wśród zagranicznych książek

Jacek Wojciechowski

Sytuacja w komunikacji publicznej, a tym samym również w transmisji oraz w obiegu informacji, ulega bieżącym zmianom. Wprawdzie nie z dnia na dzień, ale jednak sukcesywnie. W obserwacji trudno nieraz dotrzymać kroku.

Przemiany doraźne są fragmentarycznie relacjonowane w sieci oraz w czasopiśmiech pozasiecowych. Natomiast doniesienia pogłębione i perspektywne koncepcje, a także wyniki poważnych badań, oraz najistotniejsze konkluzje – jak w całej humanistyce i częściowo w obszarze nauk społecznych – nadal ujawniają się głównie w drukowanych publikacjach książkowych. Kto je zatem ignoruje, ten ponosi ryzyko, że wypadnie z naukowej gry. I na tym opiera się sens skrótowego referowania treści książek naukowych, w czym niekiedy biorę udział. Tym razem sygnalizuję publikację, z których dwie powinien (jak sądzę) przeczytać każdy, komu nauka o informacji jest rzeczywiście bliska.

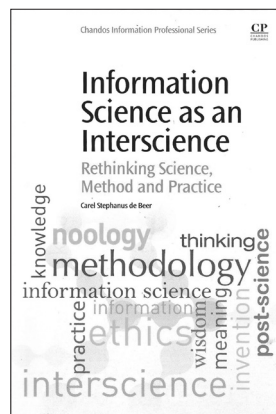
### **Nowa infozofia – raz jeszcze [\*\*\*\*]**

Carel Stephanus de Beer (2015): *Information science as an interscience. Rethinking science, method and practice*. Kidlington: Chandos Publishing, ss. 139. ISBN 978-0-08-100140-0

Rok temu przedstawiłem tu niezwykle ciekawą i ważną książkę<sup>1</sup> na temat nowej filozofii oraz zmodernizowanej koncepcji nauki o informacji – nie w pełni zresztą podzielając autorskie opinie, moim zdaniem nadmiernie skażone ideami postmodernizmu. Zwłaszcza apoteoza chaosu budzi wątpliwości, trudno ją bowiem z założeń teoretycznych przeszczepić na grunt praktycznych badań naukowych. No bo – poza samą ramą ogólną – niby jak? Aczkolwiek z tekstem prof. L. Floridiego niełatwo polemizować, bo jest napisany rewelacyjnie.

Tej opinii nie mogę jednak powtórzyć literalnie wobec sygnalizowanej tym razem publikacji – tematycznie oraz ideowo wysoce zbieżnej, również w odniesieniach postmodernistycznych. Jej autor, Carel Stephanus de Beer, to profesor informatologii z uniwersytetu w Pretorii (RPA), od lat zajmujący się filozofią informacji oraz tak jak L. Floridi postulujący nowe dla niej ramy i zmieniony charakter – co trafnie sygnalizuje w tytule określeniem internauki. Autorska znajomość filozofii w ogóle, a tym bardziej infozofii, oraz filozofii nauki, jest zresztą imponująca.

Natomiast sporo do życzenia pozostawia forma eksplicacji. Przekaz jest ciężki, mało klarowny, chwilami nadmiernie rozwlekły, ale przede wszystkim męczy niebywałą ilość powtórzeń tych samych stwierdzeń i opinii. Wygląda to tak, jakby całość powstała z autonomicznych segmentów, które nie do końca przefiltrowano i zsynchronizowano. Traci na tym siła sugestii, mimo że wiele stwierdzeń ma kreatywny charakter i odpowiednie uzasadnienie.



<sup>1</sup> L. Floridi: *The 4-th revolution. How the infosphere is reshaping human reality*. Oxford: Oxford University Press 2014 – rec. w ZIN 2015, vol. 53, nr 1(105), s.132–134.

Już na samym wstępie pojawia się niejasność, kiedy przesłanki swoich wywodów de Beer określa jako postawę niekrytyczną, tłumacząc mętnie, że to jest objaśnianie bez krytykowania. Przypuszczam, że sygnalizuje w ten sposób podejście racjonalistyczne, chociaż głowy nie dam; mało to klarowne. Odwołuje się zaś również do noologii (termin także niejednoznaczny), czyli – w zasadzie – do wiedzy o umyśle oraz do gry *ideami*.

Tak czy inaczej, z tego właśnie punktu widzenia nobilituje naukę o informacji do rangi interdyscyplinarnej, a nawet internaukowej, bo mieści się na styku wielu dyscyplin oraz wykracza poza naukę. To miłe, ale skąd taki wniosek? Otóż stąd, że informacja jest bazą myślenia (mam wątpliwości, czy tylko, a poza tym to kolejny przejaw paninformacjonizmu). W związku z tym jednak, sama wymaga nowego sposobu myślenia o sobie. Ale konkretnie jakiego – to już nie wyjaśnia się w pełni.

Dalszy wywód nawiązuje do sugestii Edmunda Husserla, że oto nauka tradycyjna przeżywa kryzys, w następstwie ograniczającej formalizacji jej zakresów, sztywnej metodologii, nadmiernej ideologizacji, oraz pretendowania do uniwersalnej wszechmocy. To jest czynnik stagnacyjny, knebel (nie da się powiedzieć, że to całkowita nieprawda), który trzeba zerwać. Jak?

Opierając się na myśleniu, czyli na **rozumie**, ale inaczej niż w praktyce dotychczasowej, bo w oparciu przede wszystkim o wartości, a o prawdę – już mniej. Z tym akurat trudno się pogodzić, ale autor tak właśnie twierdzi. Natomiast ma rację, że to musi być myślenie konstruktywne a nie destrukcyjne, mocno inwencyjne i skorelowane z dużą dawką entuzjazmu.

W sumie de Beer dość demagogicznie każe wybierać tryb postępowania: determinizm czy indeterminizm, kreacja czy destrukcja, swoboda czy schematyzacja. Trochę to przypomina pytanie, czy chce się być bogatym i pięknym. Na wszelki wypadek, w odniesieniu do nauki o informacji, autor jednak anuluje dychotomię i zaleca jednoznacznie: indeterminizm, swobodę, oraz kreatywność. Czyli: myśleć trzeba kreatywnie, ale – bez przesady?

Jest też zdania, że ta dyscyplina w obecnej postaci uległa petryfikacji: skostniała i doszła do ściany. Być może, częściowo, nie jest to diagnoza mylna, zwłaszcza w dopowiedzi, że oderwała się od życia. No bo rzeczywiście – co konkretnie ma z niej praktyka informacyjna oraz biblioteczna?

Natomiast autorskie zalecenia konstruktywne trochę ociekają sloganami. De Beer twierdzi bowiem, że (nie tylko) w tej dyscyplinie trzeba unikać pewników i myśleć w sposób otwarty, z inwencją, całkowicie *po nowemu*. No i właśnie – to znaczy: jak? Dopowiada więc, że to ma być nauka **ruchliwa**, nomadyczna, wykraczająca poza naukę dotychczasową, czyli postnauka. Owszem, impuls innowacyjny jest nieodzowny, ale to akurat są same ogólniki albo metafory.

Autor wielokrotnie nawołuje, żeby wyjść poza naukę – ku rzeczywistości. Absolutnie tak, tylko że powtarza to w całej książce nieustannie, toteż trudno nie zadać pytania, co to właściwie znaczy. Oraz dopowiada, też w serii powtórzeń, że trzeba w tej dyscyplinie myśleć kompleksowo. Według mnie, to oczywistość. Którą de Beer dopełnia zaskakującą opinią, że nie potrzeba już więcej wiedzy, lecz więcej mądrości. Mądrość bez wiedzy? Otóż to jest ewidentna żonglerka werbalna. Ale oczywiście pomyśleć nad tym warto, przełamując mglistą metaforę.

W tej książce zresztą z reguły jest tak, że im bliżej do konkretów, tym mocniej śwędzą wątpliwości. Oto powiada autor, że internetyzacja informacji wprowadziła do informatologii formalne ograniczenia oraz technologiczny fundamentalizm: króluje inżynieria informacyjna. Owszem, ale w infotonice. W odniesieniu do nauki o informacji powiedziałbym raczej, że dokonała się biurokratyzacja opinii, a górę wzięło sloganiarstwo i myślowe lenistwo, natomiast nie widzę wypowiedzi rzeczywiście inżynierskich. Jeśli już, to pseudoinżynierskie pozy, a to nie to samo. Ale być może autor widzi to szerzej.

Ratunkiem wobec zastoju ma być dociekanie prawideł nadawania, transmisji, oraz odbioru informacji – i to jest bezdyskusyjnie racja. Ale co w tym nowego i odkrywczego? Wszak przecież do tego właśnie sprowadza się wykorzystywanie w informatologii metod i narzędzi, przejętych już dawno z teorii komunikacji. Ale powtórzę: zapewne warto to wszystko zastanowienia oraz rekapitulacji.

Autorska koncepcja teorii informacji zakłada oddalenie się w analizach od linearności, tabelaryzacji, a nawet od sieciowości. Na to miejsce de Beer wprowadza inną konfigurację, a na pewno

nazwę: to atlas wiedzy. Wielokierunkowy, otwarty, zanurzony w nieskończonej konwersacji, która po tym atlasie rozlewa się nieustannie, współkreując nowe treści. W takiej referencji to nawet ładnie wygląda, ale jakie są konkrety? A miało być bliżej życia...

Jego apologia nauki o informacji osiąga wyżyny w charakterystyce pracownika informacji. Z woli autora, to ma być mianowicie reprezentant całej nauki oraz całej wiedzy! Hmm... Czasem chciałoby się teoretyków informacji oddelegować na miesiąc do pracy w informatorium – pomiędzy konkrety. Zwłaszcza, że miejscem dosiężnym autorskiego wywodu jest nieoczekiwane określenie trubadur wiedzy. Jakkolwiek sformułowanie wygląda imponująco, to jednak powtórzę: hmm.

Klarowność eksplikacji nie jest dla tej książki charakterystyczna. Przeciwnie – wywody bywają mętne, przymglone dodatkowo przez wielokrotną powtarzalność. Natomiast fundamentalną zaletą jest próba wyrwania informatologicznych rozważań ze stagnacji, z zakłętego kręgu banałów i z mało ważnych problemów zastępczych. A chociaż autorskie sformułowania nie rzucają na kolana, bo są zbyt ogólnikowe, to jednak skłaniają do niestandardowych refleksji. Dlatego jest to publikacja, którą koniecznie trzeba przeczytać.

Z całą pewnością oryginalne i nobilitujące są autorskie odniesienia do różnych przesłanek filozoficznych. W bibliologii oraz w informatologii to jest absolutna rzadkość, tymczasem każdej nauce zaplecze filozoficzne – jeżeli jest – pomaga wyznaczyć rangę. De Beer odwołuje się do rozmaitych kierunków oraz szkół filozoficznych, biele poruszając się wśród nich slalomem, czasem zresztą w zaskakujących zestawieniach. W większości wywodów ujawniają się, wprost lub pośrednio, nawiązania do koncepcji postmodernistycznych oraz postponowoczesnych – co chwilami budzi mój sceptycyzm, ale w końcu to nie ja rozstrzygam, co jest trafne i ważne, a co nie.

Wśród wielorakich relacji oczywiste są autorskie nawiązania do krytyki czystego rozumu E. Kanta, oraz do rozważań G. Hegla o sprzecznościach pomiędzy logiką i faktami. To zaś daje asumpt do sugestii, że świat jest zbyt skomplikowany, aby do zrozumienia go wystarczyły proste reguły poznawcze. Nauka – twierdzi de Beer – wymaga rozległych i nieskrępowanych poszukiwań, oraz specjalnego myślenia kompleksowego (Edward Morin nazywał je spiralnym), możliwie pluralistycznego, jak zalecał ojciec filozofii postmodernistycznej, Jean-Francois Lyotard. Trudno z tym polemizować, ale wobec tego obecne u nas aktualnie bezkrytyczne bałwochwalstwo dla sformalizowanych ocen nauki, w postaci parametryzacji, należałoby z obrzydzeniem odrzucić. Tymczasem nikt tego nie robi...

Z drugiej strony: pełne przeniesienie tak zróżnicowanych założeń na grunt nauki o informacji, w autorskim opisie bywa nieraz nie tylko mało klarowne, lecz wręcz niemożliwe do zaakceptowania. Choćby w sugestii, że podstawy informacji oraz całej o niej nauki nie może stanowić prawda, bo to jest konstrukt abstrakcyjny – lecz zamiast niej: sensy i wartości. Te jednak, na moje wyczucie, też są abstraktami, toteż cały ten wywód wygląda na kolejną zonglerkę terminologiczną.

W bezpośrednim nawiązaniu do nowych pomysłów na wiedzę o informacji, de Beer przywołuje opinie kolejnych filozofów. Oto więc Kanadyjczyk, Bernd Frohman uważa, że najwyższy już czas, aby oderwać się od abstrakcyjnej koncepcji komunikacji (również informacyjnej), a postrzegać oraz badać przebieg rzeczywistych procesów transmisji, a zwłaszcza odbioru. Natomiast Amerykanin David Blair dopowiada, że konkretnie trzeba rozpoznawać sygnalizacje oraz odbiorcze kreacje znaczeń, dalekie wszak od wzajemnej jednoznaczności, bo kontekst przyjmowania (dopowiem: i odbiorcza świadomość) w istotny sposób je różnicuje. To są sugestie szczególnie ważne, ale (jak widać w praktyce) rzadko brane pod uwagę.

Stosunkowo niecodzienne refleksje wynikają z poglądów francuskiego filozofa Emmanuela Levina-sa – że nauka i wiedza generują się głównie z etyki – oraz z rozważań Rainera Kuhlena (Niemcy) nad etyką informacji. Pada mianowicie pytanie o własność i o prawo własności informacji, prywatnych oraz publicznych w dużym przeciwieństwie pomieszanu, a także o odpowiedzialność za ich treść i dystrybucję.

Jednak głównym filozoficznym guru pozostaje dla autora francuski uczony Michel Serres, cytowany w książce obficie i często. Który jest twórcą idei inteligencji zbiorowej, kreowanej wspólnie w oparciu o łączność między ludźmi, a więc o wymianę informacji, czemu sprzyja *nieograniczona*

możliwość relacji w przestrzeni wirtualnej. Otóż w rzeczywistości niekoniecznie tak jest, a teza o nieograniczoności to czysta abstrakcja.

Kończąc w tym segmencie opinia jest taka, że dotychczasowy system informowania oraz kształcenia sprzyja raczej zapamiętywaniu istniejących treści, aniżeli kreowaniu nowych, więc w niemałym stopniu ogranicza myślenie, petryfikuje, sprowadzając wychowanie do opanowania sprawności, użytecznych obecnie, ale w przyszłości – już niekoniecznie. Niby racja, jednak to zaledwie cienki dodatek w szerokim dyskursie na temat kryzysu edukacji szkolnej oraz pozaszkolnej, w którym inni autorzy<sup>2</sup> mają do powiedzenia znacznie więcej.

Mimo zgłaszanych tu obiekcji wobec postponowoczesnej apoteozy chaosu i niepewności – w nauce to może być czynnik sprawczy, lecz nie jedyna wykładnia – główne autorskie opinie warto wziąć pod uwagę, co najmniej dla przemyślenia. Zwłaszcza jego apel o rozumowanie otwarte, pełne inwencji kreatywnej, a więc wykraczające poza samą tylko innowacyjność, która jest z góry ograniczona.

Do tego zaś konieczna jest nowoczesna koncepcja teorii informacji, którą autor lokalizuje w samym centrum nauki w ogóle. Jednak z ukierunkowaniem nie tyle na informacyjną technologię, co głównie na przemyślenie miejsca i roli człowieka w świecie informacji.

Z tej książki, a jeszcze bardziej z przywołanej tu już publikacji L. Floridiego, wyłania się interesujący obraz refleksji naukowej nad informacją i panorama jej filozoficznej podbudowy. Poziom jest imponujący. Zdecydowanie przewyższa jakością filozoficzne refleksje nad bibliotekarstwem, jakie stosunkowo niedawno trafiły mi się do czytania<sup>3</sup>. Co stwierdzam z niejaką zazdrością, bo jednak klasyczne bibliotekoznawstwo jest mi trochę bliższe.

### Informacja polityczna [\*\*]

Tetiana Granczak (2014): *Biblioteka w informacyjnym suprowodzie uprawlinnia suspilnimi procesami: politiko-komunikacijnij aspekt*. Kiiw: Nacjonalna Akademia Nauk Ukraini, Nacjonalna Biblioteka Ukraini im. W.I.Wernadśkogo, ss. 182. ISBN 978-966-02-7136-4

Zasadność opinii de Beera rzuca się w oczy najbardziej tam, gdzie sformalizowane ograniczenia pseudonaukowe sprowadziły rozmaite dyscypliny naukowe do parteru, skąd teraz trudno im się wygrzebać. Tak jest w Rosji i na Białorusi i na Ukrainie (zapewne nie tylko), gdzie post-sowieckie kanony nadal w tamtejszym bibliotekoznawstwie i w nauce o informacji stanowią zauważalne obciążenie, mimo ewidentnych wysiłków, żeby się z nimi rozstać.

Dotyczy to także drobiazgów (ale uciążliwych), jak mania sygnalizowania autorskich imion tylko przez pierwsze litery, więc wielu autorów nie da się zidentyfikować. Także – zagranicznych, bo z kolei notowanych tylko po rosyjsku (białorusku, ukraińsku), bez łacińskiej transkrypcji. No i jest też praktyka operowania skrótami nazw, dla obcokrajowca nie do rozszyfrowania, a żadnych indeksów nie ma. Na to zaś nakłada się jeszcze dotkliwie odcięcie od źródeł zachodnich, trochę za sprawą odmiennych alfabetów i odrębnego internetu, a także niedostatku środków na zachodnie publikacje, ale przede wszystkim to jest efekt zasiedziałości tam izolacjonizmu. Szkoda, bo potencjał wydaje się znaczny.

Місце і роль бібліотек  
у формуванні  
національного  
інформаційного  
простору

Матеріали міжнародної наукової конференції  
(Київ, 21-23 жовтня 2014 р.)

<sup>2</sup> Np. G. Thomas: *Education*. Oxford: Oxford University Press, 2013; *Neuroscience in education*: Oxford: Oxford University Press, 2012.

<sup>3</sup> Bill Crowley: *Renewing professional librarianship. A fundamental rethinking*. Westport: Libraries Unlimited, 2008; Charles B. Osburn: *The social transcript. Uncovering library philosophy*. Westport: Libraries Unlimited, 2009.

Jednak najdotkliwsza tamtejsza petryfikacja odnosi się do sposobu uprawiania obu dyscyplin i nauk humanistycznych w ogóle – w dokładnej opozycji do sugestii de Beera i Floridiego – gdzie górę biorą roztrząsania terminologiczne. Otóż sporo wypowiedzi, poza tym nieraz ciekawych, sprowadza się do tego, co znaczy jakieś sformułowanie, oraz kto co powiedział. Mam nadzieję, że my mamy to za sobą, ale za pewność nie poręczę. W sygnalizowanej tu książce ukraińskiej (oczywiście nietrudno zrozumieć, że uprawianie nauki tam teraz to zadanie szczególnie skomplikowane), przywołuje się bagatelne opinie autorów głównie ukraińskich i rosyjskich, równie *odkrywcze*, jak stwierdzenie, że po środzie następuje czwartek. Chwilami trudno to czytać.

Nie lekceważę terminologii, ale to jest tylko narzędzie (a raczej narzędko), a nie przedmiot i cel naukowego poznania. To tak, jak krakowsko-warszawska kontrowersja, czy wychodzi się *na dwór*, czy też *na pole*, podczas gdy w rzeczywistości wychodzi się na ulicę. Z tego nie wynika absolutnie nic.

W sygnalizowanej tu książce autorka z kijowskiej Akademii Nauk podjęła rozważania na temat istoty oraz funkcjonowania informacji politycznej i roli bibliotek w jej transmisji. Zdawałoby się, że w kraju o postsowieckiej proveniencji oraz z tak niebywałymi zawirowaniami politycznymi w historii i w teraźniejszości, trudno o temat bardziej fascynujący. Tymczasem tekst utonął w żonglerce terminami oraz w jałowej kazuistyce, miejscami nudząc ponad wytrzymałość. Nie twierdzę, że jest w ogóle bez wartości (bo nie byłoby tego omówienia), ale sporą część książki skażił właśnie formalizm, w ZSRR uznawany za ważny wskaźnik nauki, ale dzisiaj prezentujący się smętnie.

Właściwie tylko raz, na dwóch stronach, ujawnił się cały dramatyzm wydarzeń politycznych, oraz skomplikowania politycznej informacji, w powiązaniu z bibliotekarstwem, na Ukrainie zilustrowany akurat... zmianami nazwy obecnej Biblioteki Narodowej. Otóż w 1918 r., decyzją hetmana Skoropadskiego, została powołana Biblioteka Narodowa Ukrainy. W rok później władze komunistyczne zmieniły jej nazwę na Bibliotekę Ogólnonarodową. W 1934 r. nowa nazwa brzmiała: Ogólnoukraińska Biblioteka Akademii Nauk, a po dwóch kolejnych latach – Biblioteka Akademii Nauk USRR. W 1965 r. dokonano następnego przemianowania na Centralną Bibliotekę Naukową Akademii Nauk USRR. W 1988 r. dodano: im. W.I. Wernadskiego. No i wreszcie od 1996 r. to jest Narodowa Biblioteka Ukrainy im. I.W. Wernadskiego.

Już samo to jest symptomatyczną charakterystyką chwiejności politycznej informacji oraz politycznego uwikłania bibliotek. Zarazem też żonglerka nazwami i celebrowanie rejestracyjnych detali ilustruje pośrednio absurdu pseudonaukowego formalizmu, od którego tak zdecydowanie odciął się de Beer, nie mając zresztą pojęcia, jak to wygląda w okolicznościach ekstremalnych. A ich ślady, wciąż jeszcze obecne za naszą wschodnią granicą, redukują produktywność tamtejszych dociekań informatologicznych i bibliologicznych znacznie poniżej możliwości.

W gąszczu rozmaitych mielisz formalnych zaplątała się też Tetiana Granczak, w swojej mocno zresztą chaotycznej relacji – przywołując przy tym (bez żadnego pożytku) mnóstwo tekstów ukraińskich oraz rosyjskich, które po prostu napompowały bibliografię. Natomiast zabrakło potrzebnego, ale dobrego uściślenia, co to właściwie jest ta polityczna informacja. Znalazł się tam wprawdzie zapis, że to zawartość dokumentów państwowych (rządowych i samorządowych) oraz międzynarodowych, ale trudno wszak na tym poprzestać. Za to można się zgodzić, że – jak napisano – w komunikacji politycznej biblioteki pełnią role archiwizatorskie oraz transmisyjne, rozpowszechniając stosowne materiały i w ten sposób przyczyniają się do demokratyzacji życia politycznego, które dzięki temu jest mniej utajone. Ale istnieje także, niestety, jeszcze jedna funkcja, nierzadko *standardowo* narzucana bibliotekom – co Granczak słusznie sygnalizuje – mianowicie rola agent politycznej propagandy.

W dążeniu do pełnego ustalenia relacji pomiędzy informacją polityczną i bibliotekami, autorka liczy (nie wiem czy słusznie) na nowe opinie naukowe, wynikające z innego niż dawniej sposobu postępowania w naukach o bibliotece i o informacji. W ZSRR mianowicie traktowano biblioteki jako ośrodki propagandy ideologicznej – z dramatycznymi tego skutkami – a podstawy ku temu sformułował O. Czubarjan, główny ideolog (a moim zdaniem: dewastator) tamtejszej bibliologii. A co

zmieniło się konkretnie? Otóż w następstwie uznania bibliotek za instytucje społeczne (ale nie ma wyjaśnienia: przez kogo), obecnie biorą górę społeczne oraz komunikacyjne analizy bibliotekarstwa i funkcjonowania informacji, zdecydowanie produktywniejsze, niż ideologizacyjne. Może to i racja, jednak zaniepokoiło mnie wplątanie w ten wywód wzmianki o metodzie dialektycznej, o której tacy jak ja zdążyli zapomnieć, a młodzi nigdy nie słyszeli. Czyżby miało wrócić nowe?

Jest w tej książce bardzo drobiazgowy, chociaż wybiórczy, przegląd historyczny bibliotecznych i okołobibliotecznych wydarzeń na świecie, potwierdzający konieczne, ale niebezpieczne uwikłanie bibliotek – czasami wbrew woli, a czasem nie – w procesy politycznej informacji oraz propagandy. Co wprawdzie w tekście zakłóca ogólny tok eksplikacji, ale samo w sobie jest interesujące.

Granczak uwypukla manipulatorską rolę bibliotek w tym zakresie, wymuszoną przez władze, przede wszystkim w hitlerowskich Niemczech, w carskiej Rosji, oraz w ZSRR. O współczesności wypowiada się mniej i ostrożnie, deklarując ufność w demokratyczne systemy władzy, jakkolwiek te nie są przecież powszechne. Tonuje natomiast entuzjazm wobec demokratyzacyjnych cech Internetu, ponieważ w komunikacji sieciowej występują licznie również zjawiska przeciwstawne. Dlatego, jej zdaniem, właśnie biblioteki mogą i powinny stać się neutralnymi agendami rzeczywiście informacyjnymi, likwidując ewentualne służebności propagandowe.

Co łatwo napisać, ale w praktyce wszystko jest znacznie bardziej skomplikowane. A swoją drogą, to mógłby być obszar do dyskusji z tezami C. S. de Beera, jakoby w informacji prawda nie musiała stanowić kategorii fundamentalnej. Lecz takiej konfrontacji poglądów oczywiście nie ma.

Bieżące powiązania bibliotek z informacją polityczną – głównie w kontekście ukraińskim i to akurat jest najciekawsze – autorka sprowadza do czterech wariantów. Charakteryzuje mianowicie biblioteki parlamentów, upatrując ich główne powinności w informacyjnej obsłudze organów władzy wszystkich szczebli, natomiast ich ofertę dla szerokiej publiczności traktuje marginalnie.

Odrębnie opisuje biblioteki prezydenckie – w USA, w Rosji, oraz na Białorusi – jakkolwiek ich rola klaruje się mętnie i to bynajmniej nie z winy autorki: tak chyba rzeczywiście jest. Na Ukrainie zresztą takiej biblioteki nie ma, jednak w strukturze tamtejszej Biblioteki Narodowej istnieją kolekcje prezydenckie, złożone z zagranicznych podarunków dla czterech (teraz pięciu?) kolejnych prezydentów Ukrainy. W opisie wygląda to na składnicę rozmaitych rekwizytów, więc nie tylko książek, a jaki miałby z tego być pożytek, odgadnąć niełatwo. Tak czy inaczej, takie biblioteki istnieją i – przynajmniej formalnie – biorą *jakiś* udział w procesach politycznego informowania.

W ogóle natomiast kijowska Biblioteka Narodowa (tak jak inne biblioteki narodowe), zdaniem autorki, wypełnia w komunikacji politycznej rolę doniosłą, opracowując i rozpowszechniając informacje polityczne oraz komentarze – w rozmaitych formach. Granczak zreferowała to bardzo szczegółowo, dodając jeszcze wysoką ocenę tej działalności przez władze polityczne, ale nie wspomniała: które. Wygląda na to, że na Ukrainie oraz w innych krajach jest mniej więcej tak, jak napisała; i tu pojawia się refleksja, że być może nie byłoby źle, gdyby akurat biblioteki narodowe w informacyjne politykowanie angażowały się jak najmniej. To jest gra naprawdę ryzykowna. W końcu zaś to właśnie biblioteki parlamentarne (oraz ewentualnie prezydenckie) są powoływane do takich zadań.

Inne biblioteki naukowe oraz publiczne na Ukrainie (łącznie 865), wspólnie z narodową, tworzą i rozpowszechniają polityczne bazy informacyjne, bibliografie, oraz inne materiały, w tym także prawne – z retrospektywą aż do 1938 roku, ale nie wyjaśnia się po co i do czego miałyby to komukolwiek służyć – na szeroki użytek publiczny. W znacznej części istnieje do nich dostęp przez internet, czasami w *popularnej* postaci pytań oraz odpowiedzi. I to jest czwarty wariant obiegu informacji politycznych.

Granczak nie ma wątpliwości, że biblioteki powinny informować społeczeństwo o polityce władz, a dla nich z kolei konsultować stan i jakość politycznej informacji, oraz ponadto uczestniczyć w wychowaniu obywatelskim – jakkolwiek pod istotnym warunkiem: unikania jakiegokolwiek konkretnej indoktrynacji ideologicznej. Oczywiście nie sposób wyłączyć biblioteki z obiegu politycznej informacji, ale grunt jest śliski, a cytowany warunek, chociaż niby słuszny, jest rozbrajająco nierealny. Rzeczywistość bowiem nie zna nigdzie sytuacji całkowitej neutralności bibliotek (i bibliotekarzy)

wobec swoich organizatorów, tak publicznych, jak i pozarządowych oraz pozasamorządowych. Każdy kto płaci, zawsze jest pierwszy w kolejce po laury, sympatię i po aprobatę.

Dlatego na racjonalną i neutralną informację polityczną można liczyć tylko wówczas, kiedy w jej rozpowszechnianie (z przejmowaniem od rozmaitych nadawców) zaangażują się biblioteki różnej proveniencji, ale wyłącznie dla transmisji, kategorycznie natomiast bez komentowania. Bo tym akurat zajmują się inni.

Tego właśnie uczy – wydawałoby się, że na Ukrainie szczególnie – obfita historiografia najnowsza. Okazuje się bowiem, że te same informacje polityczne, oraz zdarzenia i fakty, prowadzą do całkowicie odmiennych konkluzji, a nieraz nawet do konfrontacji fizycznych. Trzeba więc postępować ostrożniej, aniżeli z kwasem solnym. Oraz na przekór opinii de Beera, przywrócić informacji konieczny atrybut prawdziwości. Na ile jest to możliwe.

W książce Tetiany Granczak przeważa koncepcja wykreowania na Ukrainie (i zapewne wszędzie) względnie jednolitego, spójnego systemu informacji politycznej, przede wszystkim za pośrednictwem bibliotek. Ta idea, ewentualnie słuszna w obszarze informacji potocznej, albo naukowej, lub innej przedmiotowej, w odniesieniu akurat do politycznego informowania, wydaje się błędna. Praktyka podpowiada bowiem, że to pluralizm nadawczy i transmisyjny jest w tym wypadku rozwiązaniem najlepszym.

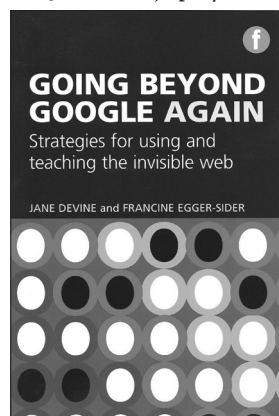
### Internet głęboki [\*\*\*\*\*]

Jane Devine, Francine Egger-Sider (2014): *Going beyond Google again. Strategies for using and teaching the invisible web*. London: Facet Publishing, ss. 180. ISBN 978–1-85604–838-5

Znalazłem dowód na galopadę czasu. Kiedy mianowicie przed sześcioma laty zauważyłem książkę<sup>4</sup> tych samych (jak wyżej) autorek, zaniechałem czytania, bo temat wydał mi się marginalny. Dotyczył mianowicie tzw. internetu głębokiego albo ukrytego. Określenia są zresztą różne: *invisible Web*, *deep Web*, *hidden Web*, *darknet* i dotyczą tej zawartości sieci (w tym mnóstwa bibliotecznych baz danych), której nie wyluskują popularne wyszukiwarki, w tym również Google. Czego nie należy mylić z niezamazywalnymi śladami każdego zapisu w sieci, sygnalizowanymi jako *deep log*, bo to coś zupełnie innego<sup>5</sup>. Otóż wystarczyło kilka lat, żeby rozległość zjawiska przerosła wszelkie wyobrażenia.

Tak więc to był błąd. Nie miałem racji – również ze względu na jakość referencji. Oto bowiem druga książka na ten temat tych samych autorek jest pod każdym względem znakomita. Świetnie napisana, kompetentna i głęboko mądra. Okazja do takiej lektury zdarza się bardzo rzadko. Obie panie pracują w bibliotece La Guardia Community College (Devine to szefowa) – ogromnej, wielodzielinowej szkoły wyższej, jednej z największych w USA, w Long Island City, blisko Nowego Jorku – i doświadczenia praktyczne dodatkowo cały wywód racjonalizują.

Zdaniem autorek, granice ukrytego Internetu (jest jeszcze nowa nazwa segmentująca *cognitive invisible Web*) są płynne. Aktualnie to podobno 85% informacji, w ogóle zgromadzonych na stronach WWW, z czego standardowe wyszukiwarki wylawiają teraz coraz więcej, ale to raptem 1/3 ogółu treści<sup>6</sup>. Jeżeli to prawda, to opowieści o tzw. społeczeństwie informacyjnym są zbiorem mitów.



<sup>4</sup> J. Devine, F. Egger-Sider: *Going beyond Google. The invisible web in learning and teaching* New York: Neal-Schuman 2009.

<sup>5</sup> D. Nicholas, D. Clark: *Evidence of user behaviour: deep log analysis. W: User studies for digital library development*. London: Facet Publishing 2012, 85-94.

<sup>6</sup> Podobna opinia: T. Mann: *The Oxford guide to library research*. Ed. 4. New York: Oxford University Press 2015, 1.

Jakkolwiek niedostępność dla wyszukiwarek bywa często zamierzona, z góry jest także efektem różnych filtracji, personalizacji, oraz passwordów. Z drugiej strony: są też inne formuły wyszukiwania, ogarniające częściowo ukryty Web. Poważnym rywalem dla Google jest teraz Facebook, rośnie też korzystanie z Wikipedii, zaś specjalnie dla głębokiego Internetu powstała wspierająca baza *Silence Direct*.

W Internecie ukrytym przeważają informacje wyspecjalizowane, więc szczególnie ważne dla specjalistów właśnie. Tymczasem według rozeznania autorek, najczęściej docierają do tych wiadomości *komputerowcy* i w ogóle informatycy, natomiast inni fachowcy już znacznie rzadziej i daleko nie wszyscy. Również bibliotekarze oraz pracownicy służb informacyjnych – nieczęsto (!?), jeżeli nie liczyć korzystania z własnych baz danych. To sygnalizuje, że są w tym zakresie ogromne rezerwy. Tym bardziej, że obecnie widać radykalny wzrost aplikacji na smartfony, co potencjalny zasięg takiej oferty może rozszerzyć jeszcze bardziej.

Na razie publiczność zdecydowanie preferuje wyszukiwarki Google, ze względu na szybkość, wygodę, oraz (już) przyzwyczajenie, mimo że posiłkowanie się słowami kluczowymi (to obecnie norma) nieraz sprawia trudności<sup>7</sup>. Mało kto natomiast zdaje sobie sprawę, że ukryty internet zawiera inne treści, toteż nawet poszerzenie obszaru wyszukiwawczego przez Google – co ostatnio ma miejsce – nie zapewnia dostępu do wszystkich informacji potrzebnych i ważnych.

Tak czy inaczej (badania na dużej próbie) poszukiwanie informacji zaczyna od wyszukiwarek 80% internautów, a ponad 83% studentów<sup>8</sup>, podczas gdy od bibliotek nie zaczyna tego prawie nikt. Potem, w dalszym korzystaniu, proporcje są już inne. Mianowicie sieciowe wyszukiwarki wykorzystuje w ogóle 93% publiczności, natomiast biblioteczne bazy informacyjne 33% publiczności ogólnej, a 57% studentów. I to niezależnie od tego, że ostatnio zaufanie do standardowych wyszukiwarek podobno spada<sup>9</sup>.

Uytuowanie autorek w bibliotece uczelnianej ułatwiło analizy zbiorowości studenckiej – w dużej (ponad 8000 respondentów) grupie, co czyni wyniki w miarę wiarygodnymi. W najogólniejszym ujęciu, Devine i Egger-Sider są zdania, że studenci rozpaczliwie walczą z nadmiarem informacji, toteż traktują automatyczne wyszukiwarki jak zbawienie, nie uświadamiając sobie różnicy pomiędzy informacją szybką a rzetelną, zwłaszcza naukową<sup>10</sup>. Najchętniej zadowolają się informacją dostateczną, bez intencji pogłębienia lub wzbogacenia. Potwierdza to prof. Naomi Baron w przeglądzie doniesień badawczych; okazuje się mianowicie, że 77% studenckich cytowań w różnych rozprawach pochodzi z pierwszych trzech stron cytowanych źródeł<sup>11</sup>. A z moich własnych obserwacji wynika, że tak postępują nie tylko studenci.

Interesujący jest też, przedstawiony przez autorki, hierarchiczny rejestr sygnalizowanych trudności przy studenckim łowieniu informacji. Otóż dla 84% trudny jest sam start, potem zaś – ustalenie konkretnego zakresu poszukiwań (66%), oraz odrzucenie treści zbytecznych (61%). Odsetki są niepokojąco wysokie.

O internetowych wyszukiwarkach – tak jak wielu innych znawców – autorki nie mają najlepszego zdania, twierdzą bowiem, że proces wyszukiwania przebiega w nich trybem *jak leci*. Ewentualna personalizacja opiera się na odnotowaniu wyszukiwań poprzednich (cookies) oraz dotychczasowych kontaktów, a czasem także miejsc zamieszkania. To może wystarczyć dla reklamy butów, bądź najbliższego fryzjera, ale nie dla informacji poważnej. Tę oferują (w tym celu wszak weryfikowane) bazy biblioteczne, lecz ich wykorzystanie nie jest wystarczająco szerokie.

<sup>7</sup> Mann, 2015, 116.

<sup>8</sup> *Perceptions of libraries 2010. Context and community*. Dublin (USA): Online Computer Library Center 2011, 54.

<sup>9</sup> *Perceptions*, 2015, 43, 56.

<sup>10</sup> Mann, 2015, 339.

<sup>11</sup> N. S. Baron: *Words onscreen. The fate of reading in a digital world*. New York: Oxford University Press 2015, 39.

Z tego zaś bierze się autorska sugestia, że do używania internetu głębokiego trzeba przysposabiać – już uczniów w szkole, a potem studentów. W praktyce: dwuetapowo. Najpierw każdy powinien opanować ogólne reguły szukania oraz korzystania z informacji w sieci, a potem, w drugiej fazie – docierania do internetu ukrytego. Ze specjalnych sondaży wynika, że w USA i w UK praktykuje się to bardzo rzadko. Przypuszczam, że u nas też.

Referując to, chcę zwrócić uwagę na jeszcze jedną opinię auterek, zasygnalizowaną na koniec rozważań, niby marginalnie, ale dobrze udokumentowaną. Otóż (nie tylko) ich zdaniem, rozwój technologii komunikacyjnej zmierza w zauważalnym stopniu do częściowego zastąpienia dotychczasowych – w zasadzie pisemnych (z dodatkami) – relacji między nadawcami, urządzeniami i publicznością, przez systemy głównie dźwiękowe i głosowe. Podobno nie tylko w obszarze komend i porozumień zwrotnych (co wyeliminowałoby klawiatury i wskaźniki manualne), ale też w zakresie serwowania treści. Stąd biorą się różne inne głosy, że oto być może zanosí się na fragmentaryczny powrót do komunikacji oralnej – jak ongiś – tyle, że teraz w wersji stechnicyzowanej, pośredniej oraz masowej.

Myślę, że ma w tym przypadku miejsce nieuprawniona nadinterpretacja. Dla refleksji, dla nauki i dla edukacji, a zresztą nawet dla komunikacji potocznej, wyłącznie audialny lub audiowizualny system porozumiewania się absolutnie nie wystarczy. Gdyby było inaczej, po co byłoby rezygnować w swoim czasie z magnetofonów i audiokaset, oraz modernizować transmisje telewizyjne, a i współczesny obieg audiobooków nie byłby tak marginalny, jak jest. Osobiście dostrzegam w tych głosach kolejne próby zdyskredytowania pisma jako ważnego narzędzia w publicznym porozumiewaniu się. Myślę, że trzeba do tego podchodzić z najwyższą nieufnością, ponieważ twardych dowodów nie ma.

Natomiast jest jakaś racja w sugestiach obu auterek, że uznanie jakoby obecny kształt semiotyczny komunikacji publicznej, więc także informacyjnej, był już zamknięty i ukształtowany ostatecznie, to co najmniej nieostrożność, albo nawet błąd. Po prostu nie wiemy, co w tym zakresie można jeszcze wykreować nowego. Trzeba zatem obserwować i w miarę potrzeby reagować na to, co ulegnie zmianie. Również, jeśli kiedyś Internet głęboki przestanie być w pełni ukryty.

*Nadesłano: 26 września 2015.*

## Przegląd polskich nowości wydawniczych

Anna Stanis

*Biblioteka Uniwersytecka w Warszawie*

***Biblioteka, książka, informacja, Internet 2014.*** Praca zbior. pod red. Zbigniewa Osińskiego, Renaty Malesy, Sebastiana D. Kotuły. Lublin: Wydaw. Uniwersytetu Marii Curie-Skłodowskiej, 2015, ss. 286. ISBN 978–83-7784–685-8

Ukazał się kolejny (już trzeci) tom z cyklu „Biblioteka, książka, informacja, Internet”, który przedstawia przegląd najnowszych badań naukowych z dziedziny informatologii i bibliologii. Składa się z dwóch części. Część pierwsza zatytułowana *Informacja i wiedza w dobie Internetu* zawiera dziewięć artykułów poświęconych zjawiskom i procesom informacyjnym zachodzącym w sieci. Rozważano w jaki sposób Internet wpływa na rozwój humanistyki (Zbigniew Osiński), a także wykorzystanie idei Open Source w działalności m.in. informacyjnej (Sebastian D. Kotuła). Poruszono problemy wyszukiwania materiałów wizualnych w Internecie (Kamil Stępień), serwisów dziennikarstwa obywatelskiego (Wanda A. Ciszewska), digitalizacji (Monika Sobczak-Walis; Rafał Wójcik). Zanalizowano, stosując indeks cytowań bibliograficznych CYTBIN, recepcję tekstów z dwudziestolecia międzywojennego we współczesnym piśmiennictwie bibliologicznym (Agnieszka Bajor). Część pierwszą zamyka artykuł poświęcony charakterystyce wybranych portali wymiany poglądów i wyników badań pracowników naukowych (Lidia M. Jarska) oraz artykuł na temat repozytorium instytucjonalnego, jako elementu promocji nauki (Lidia Mikołajuk).

Część druga zatytułowana *Kształcenie, książka i biblioteka a możliwości stwarzane przez Internet* zawiera 8 artykułów dotyczących m.in.: zmian w kształceniu w zakresie informatologii i bibliologii (Małgorzata Kowalska), bibliotek cyfrowych i bibliotek internetowych wykorzystywanych w procesie kształcenia (Żaneta Szerksznis; Anita Has-Tokarz), zastosowania technologii mobilnych w dostępie do wybranych zasobów cyfrowych (ośmiu polskich i 17 zagranicznych) (Mariusz Jarocki). Obszerny tekst poświęcono analizie zastosowania grywalizacji w działalności różnego typu bibliotek (Natalia Pamuła-Cieślak). Renata Malesa ukazała działanie lubelskich bibliotek publicznych na rzecz ograniczenia e-wykluczenia. Część drugą zamykają dwa teksty poświęcone bibliotekom uczelnianym, pierwszy o integracji książek elektronicznych z systemem biblioteczno-informacyjnym szkoły wyższej (Teresa Górniak), drugi opisuje wyniki badań dotyczących wykorzystywania informacji zawartych w rekordach MARC 21 przez użytkowników w procesie wyszukiwania w OPAC (Izabela Swoboda).

**Siwecka Dorota:** *Światowy model informacji bibliograficznej. Programy i projekty (1950–2010)*. Wrocław-Warszawa: Wydaw. SBP 2015, ss. 355. Seria „Nauka, Dydaktyka, Praktyka”; 162. ISBN 978-83-64203-48-0<sup>12</sup>

W pracy przeanalizowano i przedstawiono działalność międzynarodowych instytucji zajmujących się bibliografią na przestrzeni sześćdziesięciu lat, tj. od początku lat pięćdziesiątych XX w. do pierwszej dekady XXI w. włącznie. Termin bibliografia odnosi się tu do „zbioru opisów/rekordów bibliograficznych udostępnianych w postaci publikowanego spisu tradycyjnego, elektronicznego lub w postaci bazy danych”. Analiza zaleceń i wytycznych międzynarodowych instytucji w ostatnich sześćdziesięciu latach pozwoliła na określenie jak kształtował się model informacji bibliograficznej poczynając od bibliografii narodowej, poprzez ujednolicenie i normalizację w skali międzynarodowej do etapu globalizacji. Na ostatni etap rozwoju modelu informacji bibliograficznej istotny wpływ miało także zastosowanie nowych technologii. Kierunek ostatnich zmian zmierza, ze względu na potrzeby użytkowników wynikające z samodzielnego korzystania z bibliografii, do uproszczeń opisu bibliograficznego, opisu rzeczowego i formatu opisu danych z jednoczesnym zastosowaniem prostych dla użytkownika możliwości wyszukiwawczych.

Praca składa się z dziewięciu rozdziałów. W pierwszym przedstawiono bibliografię jako przedmiot badań i źródło informacji historycznej, podano znaczenie słowa „bibliografia”. W kolejnym – dokonano analizy działalności poszczególnych organizacji międzynarodowych i wybranych narodowych (Biblioteka Kongresu, Biblioteka Brytyjska). Szeroki kontekst dla opisu kształtowania się modelu informacji bibliograficznej stanowią również teksty dotyczące rejestracji bibliograficznej, opisu bibliograficznego, formatu wymiany danych.

Osobny rozdział poświęcono opracowaniu rzeczowemu dokumentów w bibliografii uwzględniając międzynarodowe zalecenia IFLA i UNESCO. Opisano różne języki informacyjno-wyszukiwawcze i ich zastosowanie w bibliografiach narodowych. Interesującym projektem jest system MACS (Multilingual Access to Subjects), który pozwala na równoległe przeszukiwanie czterech baz danych bibliotek narodowych uczestniczących w projekcie (Biblioteka Niemiecka, Biblioteka Narodowa Francji, Biblioteka Narodowa Szwajcarii, Biblioteka Brytyjska) niezależnie od języka w jakim sformułowane jest zapytanie.

Dwa ostatnie rozdziały poświęcono organizacji pracy nad bibliografiami, formie i zawartości bibliografii narodowych. Pracę uzupełniono obszerną bibliografią, indeksami osobowym i przedmiotowym oraz wykazami skrótów.

**Tomaszczyk Jacek:** *Model systemu informacji terminologicznej*. Katowice: Wydaw. Uniwersytetu Śląskiego, 2014, ss. 218. „Prace naukowe Uniwersytetu Śląskiego w Katowicach”; nr 3257. ISBN 978-83-8012-339-7

Przedmiotem książki jest prezentacja badań prowadzonych nad organizacją i udostępnianiem informacji terminologicznej. Z przeprowadzonej analizy polskiej i zagranicznej literatury wynika, że powyższa tematyka nie ma jeszcze kompleksowego opracowania teoretycznego. W pierwszym rozdziale pracy zdefiniowano nowe pojęcie – informacja terminologiczna – poprzedzając je definicjami terminów: *informacja*, *termin* i *terminologia*. Opisano historię badań, szkół terminologii, ich twórców i teorie w nauce o terminologii. Określono cele działalności terminologicznej, różne formy jej organizacji, zarówno w skali krajowej, jak i ogólnoświatowej. Osobny podrozdział poświęcono

<sup>12</sup> Szczegółową recenzję książki Doroty Siweckiej przygotowaną przez Jacka Łubockiego prezentujemy w obecnym numerze ZIN na s. 134–137.

terminologii występującej w językach informacyjno-wyszukiwawczych najczęściej opisywanych w literaturze i stosowanych w praktyce, takich jak: klasyfikacje, języki haseł przedmiotowych, języki deskryptorowe oraz języki słów kluczowych.

Dla informacji terminologicznej zaproponowano definicję „...to informacja o terminach i wyrażanych przez nie pojęciach oraz źródłach, produktach, usługach i narzędziach terminologicznych, a także o specjalistach, stowarzyszeniach, organizacjach, instytucjach i firmach prowadzących działalność w zakresie terminologii”. Scharakteryzowano cechy informacji terminologicznej, jej rolę, znaczenie, źródła i odbiorców.

Rozdział drugi poświęcony jest problemom i systemom informatycznym związanym z informacją terminologiczną. Zarządzanie informacją terminologiczną różni się od zarządzania informacją ogólną. Jej celem jest umożliwienie użytkownikowi łatwego dostępu do zasobów terminologicznych, tak aby mógł on w różny sposób wyszukiwać, porównywać, analizować i graficznie przedstawiać uzyskane w ten sposób aktualne i historyczne informacje. Wśród problemów związanych z zarządzaniem informacją terminologiczną omówiono organizację dostępu do tego typu informacji, prawa autorskie, informatyczne systemy zarządzania informacją terminologiczną, format wymiany danych terminologicznych, wykorzystanie sieci semantycznej.

W rozdziale trzecim zaproponowano oryginalny ogólnopolski model wielodyscyplinowego systemu informacji terminologicznej, którego najważniejszymi cechami są: multidyscyplinarność, kompleksowość informacji opisującej terminy, funkcjonalność, elastyczność i otwartość na potrzeby użytkownika. Opisany model infrastruktury systemu opiera się na współpracy sieci instytucji gromadzących zasoby terminologiczne koordynowanej przez Centrum Terminologiczne.

W pracy zastosowano jakościowe metody badawcze: analizę piśmiennictwa z zakresu informatologii, terminologii rozumianej jako dyscyplina naukowa oraz terminografii, a także – metodę analizy zawartości (do badania serwisów internetowych), zastosowano również technikę wywiadu swobodnego. Książkę zaopatrzone w bogaty aparat pomocniczy.

***Teoretyczne zagadnienia bibliologii i informatologii. Studia i szkice.*** Pod red. Elżbiety Gondek. Katowice: Wydaw. Uniwersytetu Śląskiego, 2015, ss. 191. „Prace naukowe Uniwersytetu Śląskiego w Katowicach”; nr 3237. ISBN 978-83-8012-365-6

Książka zawiera zbiór rozpraw poświęconych wybranym zagadnieniom teoretycznym z zakresu bibliologii i informatologii napisanych przez grono specjalistów z tej dziedziny. Poruszono w niej problemy związane m. in. z drukarstwem, sztuką książki, bibliografią i czytelnictwem. Różnorodność problemów stworzyła okazję do szerszego spojrzenia na przedmiot badań dziedziny wiedzy, jaką jest nauka o książce i informacji. Kilka rozpraw omawia przedmiot badań bibliologii. Problemom badań czytelnictwa i jego opisu w teorii bibliologii poświęcona jest praca Ireny Sochy *Zachowania lekturowe Polaków – problemy i kategorie opisu czytelnictwa*. Na sztukę książki w badaniach bibliologicznych zwraca uwagę Małgorzata Komza. Do aktualnych zadań badawczych bibliologii Maria Juda zalicza dawne drukarstwo, proponując wyodrębnienie dziejów edytorstwa, formy książki oraz jej wyposażenia zdobniczego jako subdyscypliny. Miejsu drukarstwa w bibliologii poświęcona jest też praca Elżbiety Gondek *Zawód drukarza i technologia druku w bibliologii i poza jej zakresem*.

Pod wpływem rozwoju nowych technologii i technik komunikacji zmienia się także przedmiot badań informatologii. Barbara Sosińska-Kalata w rozprawie *Ewolucja koncepcji informatologii (nauki o informacji)* wnikliwie opisuje jak zmieniał się przedmiot badań tej dyscypliny i jak pod wpływem tych zmian kształtowała się nazwa, a tym samym jej tożsamość. Problemem przedmiotu badań informatologii zajmuje się też Jadwiga Woźniak-Kasperek w pracy *Z głównych problemów samoświadomości*

*informatologii*, zadając pytania: Czy w chwili obecnej informatologia jest samodzielną dyscypliną naukową, czy też stanowi dopełnienie bibliologii? Jakie są relacje informatologii z bibliotekoznawstwem i bibliografią? Co jest głównym przedmiotem badań informatologicznych? Z podsumowania autorki wynika, że informatologia nie jest jednolitą dyscypliną, a raczej zbiorem w różny sposób ze sobą powiązanych subdyscyplin o różnym stopniu szczegółowości.

Rozważania nad bibliometrią używaną do badania komunikacji piśmienniczej stanowią treść pracy Marty Skalskiej-Zlat. Wiesław Babik opisuje cele, zadania, strukturę i program organizacji ISKO (International Society for Knowledge Organization). Remigiusz Sapa w swojej rozprawie *Biblioteka naukowa w integrującym się środowisku informacyjnym człowieka* opisuje rolę bibliotek naukowych, wskazując na kolejny etap zadań przewidzianych dla tego typu bibliotek – budowanie spójnej infrastruktury integrującej tradycyjne zasoby z zasobami informacji cyfrowej i technologiami dostępnymi w środowisku cyfrowym.

Temat środowiska cyfrowego wpływającego na kształt współczesnej bibliografii podejmuje Jadwiga Sadowska w artykule *Dokąd zmierza bibliografia w „erze cyfrowej”?* Wychodząc od podstawowych zmian w organizacji bibliografii w ostatnim półwieczu, opisując nowe rozwiązania FRBR, RDA, ICP, autorka nakreśla perspektywy informacji bibliograficznej. W tekście J. Sadowskiej znalazło się również omówienie dwóch dokumentów istotnych dla przyszłości informacji bibliograficznej: 1) *On the Record. Report of the Library of Congress Working Group on the Future of Bibliographic Control* opracowany przez Grupę Roboczą w Bibliotece Kongresu w 2008 r., 2) *Guidelines for National Bibliographies in the Electronic Age* opracowany w tym samym roku przez Zespół IFLA. Prace nad bibliografią w „erze cyfrowej” mają przede wszystkim uczynić ją wygodną i przydatną użytkownikowi.

Książkę wyposażono w indeks osobowy.

**Włodarczyk Bartłomiej:** *Mapy tematów jako podstawa semityzacji języka haseł przedmiotowych*. Warszawa: Wydaw. SBP 2016, ss. 323. Seria „Nauka, Dydaktyka, Praktyka”; 167. ISBN 978-83-64203-62-6

Upowszechnienie Internetu przyczyniło się do zwiększenia samodzielności użytkowników w wyszukiwaniu informacji. Dostosowanie języków haseł przedmiotowych, jednego z najpowszechniejszych narzędzi wyszukiwania, do współczesnych potrzeb użytkowników, stało się zadaniem priorytetowym dla bibliotek oferujących czytelnikom ten rodzaj wyszukiwania.

Celem prezentowanej publikacji jest opis koncepcji, jednej z możliwych transformacji języków haseł przedmiotowych, która usprawniłaby wyszukiwanie rzeczowe w katalogach bibliotecznych. Autor określił cel rozprawy jako „analiza i zaprezentowanie jej rezultatów w zakresie możliwości optymalizacji wykorzystania danych semantycznych zorganizowanych jako JHP w celu stworzenia systemu zapewniającego skuteczniejsze, bliższe oczekiwaniom użytkowników wyszukiwanie informacji”. W literaturze przedmiotu powstało już wiele prac opisujących mapy tematów. Praca Bartłomieja Włodarczyka jest jedną z najobszerniejszych na ten temat i jedyną opisującą możliwość przekształcenia JHP w mapę tematów na konkretnym przykładzie JHP BN.

Książka składa się z trzech części. W pierwszej, potraktowanej jako wstęp do całej rozprawy, przedstawiono przegląd dotychczasowych wyników polskich i zagranicznych badań przeprowadzonych nad JHP, zarówno z punktu widzenia metodyki ich tworzenia jak i stosowania. Omówiono propozycje przekształcenia LCSH, włoskiego JHP Soggettario, prace nad automatyzacją indeksowania z zastosowaniem języków informacyjno-wyszukiwawczych. Prace badawcze nad przekształceniem JHP w model map tematów prowadził do tej pory jedynie japoński zespół z Center for Integrated Area Studies (Uniwersytet w Kioto) pod kierunkiem Motomu Naito m.in. dla języka japońskiego NDLSh, LCSH, AGROVOC.

W drugiej części (rozdziale pierwszym) przedstawiono mapy tematów jako model reprezentacji wiedzy, opisano historię ich powstania, scharakteryzowano elementy wchodzące w skład modelu i sposób ich zapisu. Opisano oprogramowanie pozwalające tworzyć i zarządzać mapami tematów (Ontopia, Kamala, Wandora), pozabiblioteczne obszary zastosowań map tematów tj. cyfrową humanistykę, e-learning (wykorzystanie map tematów w tworzeniu sieciowych materiałów edukacyjnych). Przedstawiono podstawowe rozwiązania stosowane w wizualizacji map tematów w różnych projektach. Porównano mapy tematów i model SKOS w przedstawieniu języków informacyjnych o notacji paranaturalnej. Wśród wymagań stawianych sposobom prezentacji map tematów jako najważniejsze wymieniono: dynamiczną adaptację do oczekiwań użytkownika oraz możliwość szybkiej i intuicyjnej nawigacji.

W części trzeciej (rozdział drugi) przedstawiono propozycję przekształcenia JHP BN w mapę tematów (w dniu 30.04.2014 słownictwo JHP BN liczyło 171 592 tematów ogólnych, formalnych i jednostkowych). Propozycję przekształcenia poprzedzono podsumowaniem dotychczasowych prac prowadzonych nad transformacją JHP w mapę tematów. Zaprezentowano model indeksowania i wyszukiwania oparty na ontologii autorstwa Winfrieda Goederta, Jessici Hubrich i Matthiasa Nagelschmidta. Celem przekształcenia JHP BN w mapę tematów jest udoskonalenie możliwości wyszukiwawczych systemu bibliotecznego, tak aby odpowiedzi na zapytania użytkowników były kompletne i precyzyjne. Jako dwa podstawowe założenia wymieniono: 1) wykorzystanie informacji zapisanych w rekordach wzorcowych jednostek leksykalnych, słowniku i hasłach rozwiniętych JHP BN, 2) zgodne z modelem danych, maksymalne wykorzystanie elementów i konstrukcji wchodzących w skład standardu map tematów. Wśród wymagań stawianych JHP BN przekształcanemu w mapę tematów wymieniono: 1) poszerzenie zapytania informacyjnego, 2) uwzględnienie w jednakowym stopniu relacji paradygmatycznych i syntagmatycznych. Przedstawiono przykładowe techniczne rozwiązania problemu konwersji danych z formatu MARC 21 do jednego ze standardowych formatów wykorzystywanych do przedstawiania map tematów. W przekształceniach rekordów wzorcowych tematów wskazane zostały rodzaje tematów JHP BN denotujących więcej niż jeden przedmiot w rzeczywistości pozadokumentacyjnej oraz sposób ich rozdzielenia, tak aby jeden temat reprezentował jeden przedmiot, co można wykonać jeszcze przed przekształceniem JHP w mapę tematów. Szczegółowo przedstawiono problemy występujące przy różnych rodzajach tematów (warianty nazw imprez, nazwy korporatywne, nazwy geograficzne). Omówiono reprezentacje różnych rodzajów relacji w mapie tematów. Rezultatem analizy JHP BN w przekształceniu w mapę tematów jest opracowanie wstępnej ontologii tej mapy. Schemat mapy zaprezentowano w oparciu o diagramy skonstruowane z wykorzystaniem programu Onotoa. Do wizualizacji wykorzystano wstępny projekt standardu GTM. Wstępna ontologia mapy tematów stworzonej w oparciu o JHP BN składa się z 19 typów tematów, dwóch typów nazw, 29 typów powiązań i 50 typów ról. Schemat może być rozbudowywany o dodatkowe typy tematów, powiązań i wystąpień.

Praca Bartłomieja Włodarczyka może stanowić źródło inspiracji dla twórców języków haseł przedmiotowych a także dla badaczy zajmujących się językami informacyjno-wyszukiwawczymi. Książkę zaopatrzone w obszerną bibliografię i aneks *Wskaźniki przedmiotu ontologii mapy tematów opartej na JHP BN*.

*Nadesłano: 18 maja 2016.*

## „Europejska konferencja edukacji informacyjnej” (Tallinn, Estonia, 19–22 października 2015 r.)

W dniach 19–22 października 2015 r. w Tallinie odbyła się 3. edycja „European Conference on Information Literacy” (ECIL), zorganizowana przez Katedrę Nauk o Informacji Szkoły Technik Cyfrowych Uniwersytetu w Tallinie, Wydział Nauk o Informacji i Komunikacji Uniwersytetu w Zagrzebiu oraz Wydział Zarządzania Informacją Uniwersytetu Hacettepe w Ankarze.

W konferencji uczestniczyło 361 osób z 50 krajów. ECIL, to, obok angielskiego LILAC, największe i najważniejsze coroczne spotkanie bibliotekarzy i badaczy zajmujących się kompetencjami informacyjnymi i edukacją informacyjną, odbywające się w Europie. Podczas czterodniowych obrad odbyło się 10 warsztatów, wygłoszono 129 referatów (w sesjach równoległych), 18 krótkich prezentacji (tzw. *pecha-kucha*), zaprezentowano 19 posterów, sześć prezentacji z badań doktoranckich, odbyła się jedna dyskusja panelowa.

Wybrane referaty zostały opublikowane w recenzowanym tomie pokonferencyjnym, nakładem wydawnictwa Springer w ramach serii *Communications in Computer and Information Science* (Kurbanoglu et al., eds. 2015).

Referaty w sesjach zostały podzielone na 23 bloki tematyczne:

- Approaches to information literacy,
- Copyright literacy,
- Development and evaluation of information literacy,
- Digital literacy,
- E-learning and m-learning,
- Everyday life media and information literacy,
- Health information literacy,
- Information and digital literacy instruction,
- Information literacy – government and society,
- Information literacy and academic libraries,
- Information literacy and information culture in higher education,
- Information literacy and learning,
- Information literacy and sustainability,
- Information literacy and technology,
- Information literacy instruction,
- Innovative information literacy approaches,
- Media and visual literacy,
- Policy and strategy,
- Professional development and information literacy,
- Professional information literacy,
- Reading format preferences of students – print vs electronic,
- Social networking and gaming,
- Standards and professional practice.

W sprawozdaniu omówione zostały jedynie wybrane referaty. Szczegółowy program konferencji oraz abstrakty dostępne są na stronie ECIL 2015, <http://ecil2015.ilconf.org/>.

Organizatorzy zaprosili jako główne referentki: Sonię Livingstone, Susan Danby oraz Carol Collier Kuhlthau.

Sonia Livingstone z Uniwersytetu Londyńskiego wygłosiła referat podczas sesji otwierającej. Nawiązała w nim do badań realizowanych w ramach sieci EU Kids Online, w latach 2006–2014. Szczególnie interesujące były przytoczone przez Livingstone wyniki badań jakościowych (obserwacji rodzin), których celem było poznanie sposobu postrzegania przez rodziców zachowań i kompetencji medialnych ich dzieci. Badania pokazały, że rodzice koncentrują się przede wszystkim na kompetencjach technicznych, np. sposobie obsługi urządzeń. Mniej uwagi poświęcają kompetencjom krytycznym, które wiążą się z rozumieniem przez dzieci, czym jest Internet, kto udostępnia w nim treści i w jakim celu. Komponent jakościowy w EU Kids Online rzucał nowe światło na wyniki badań ilościowych, np. niskich kompetencji krytycznych w odniesieniu do korzystania z mediów. Livingstone akcentowała również znaczenie wielowymiarowej analizy danych, która pozwalała interpretować zgromadzone dane na różnych poziomach: (1) jednostkowym, (2) społecznym (np. rodziny) i (3) kraju. Jest to szczególnie istotne, jeżeli badania mają być podstawą do budowania polityki informacyjnej, w tym polityki dla edukacji medialnej i informacyjnej opartej na dowodach, i uwzględniającej lokalne potrzeby.

Wystąpienie Susan Danby z Politechniki Queensland również odnosiło się do dzieci. Danby badała zachowania medialne małych dzieci (5–7 lat) w przedszkolu i w ich środowisku domowym. Sprawdzała również jaki jest stosunek nauczycieli przedszkolnych do wykorzystania mediów w pracy z dziećmi. Stosowała metodę etnograficzną, która umożliwiła interpretację i rozumienie sytuacji uczenia się różnych narzędzi cyfrowych przez bardzo młodych użytkowników, np. korzystania z map Google przez dzieci niepotrafiące jeszcze czytać. Zwracała uwagę na interakcje dzieci-komputer i dzieci-dorośli. Badania były realizowane w Australii. Kultura tego kraju – w szczególności czynniki demograficzne (np. gęstość zaludnienia) czy podejście do technologii mieszkańców obszarów rolniczych – stanowi istotny kontekst do rozumienia wyników badań.

Carol Collier Kuhlthau, emerytowana profesor Uniwersytetu Rutgersa, zaprezentowała podejście do uczenia się określane jako dociekanie ukierunkowane lub ukierunkowane dociekanie naukowe (ang. *guided inquiry*). Charakterystyczną cechą tego podejścia jest wyszukiwanie i wykorzystywanie różnych źródeł informacji, w celu rozwijania wiedzy o danym problemie, zagadnieniu. Wymaga to od uczącego zaangażowania. Proces uczenia się następuje poprzez budowanie pytań, eksplorowanie źródeł, badanie i konstruowanie odpowiedzi (wiedzy) w oparciu o znalezione informacje. Uczący się otrzymuje wsparcie od nauczycieli i bibliotekarzy, jednocześnie musi wykazać się dużą samodzielnością. To podejście przygotowuje osobę do rozwiązywania problemów (związanych z życiem codziennym, pracą) z wykorzystaniem nowoczesnych technologii. Kuhlthau poświęciła wiele uwagi w wystąpieniu modelowi Procesu Wyszukiwania Informacji (ang. *Information Search Process* [ISP]). Jest to model teoretyczny, przedstawiający proces wyszukiwania na trzech poziomach: fizycznym (działania), kognitywnym (myśli) i afektywnym (uczucia). Pokazuje, że wyszukiwanie to nie tylko czynności, które wykonuje użytkownik, ale także różnego rodzaju odczucia i przemyślenia, np. poczucie zagubienia, frustracja, optymizm, satysfakcja.

Sesje plenarne i równoległe były podporządkowane określonym obszarom badań/praktyki *information literacy*. Autorki, z uwagi na własne zainteresowania, uczestniczyły głównie w sesjach prezentujących obszar nauki i szkolnictwa wyższego. W dalszej części tekstu przedstawione zostały refleksje nad wybranymi wystąpieniami. Opisane zostały również referaty autorek zaprezentowane podczas ECIL.

Wiele wystąpień dotyczyło wsparcia przez biblioteki pracowników naukowych uczelni. Szeroko rozumiana edukacja informacyjna, skierowana na doskonalenie procesu przygotowywania do publikacji wyników badań ma pomóc w zwiększeniu liczby publikacji danej uczelni w prestiżowych czasopismach (tj. czasopismach indeksowanych w uznanych bazach). W tym kontekście Ivanka Yankova, Rumelina Vasileva, Tzvetelina Dimitrova, Kamelia Nusheva z Uniwersytetu Bibliotekoznawstwa i Technologii Informacyjnych w Sofii przedstawiły założenia programu szkoleń skierowanych zarówno do autorów jak i wydawców. Natomiast Hana Janečková i Petra Dědičová z Politechniki w Brnie zaprezentowały kurs e-learningowy skierowany do doktorantów i pracowników naukowych *Publikowanie naukowe od A do Z*.

Susanne Mikki i Marta Zygmuntowska z Uniwersytetu w Bergen przedstawiły wyniki badań dotyczących obecności naukowców z ich uniwersytetu w serwisach przeznaczonych do współpracy naukowej, takich jak: ResearchGate, Akademia.edu, czy Gogle Scholar. Te badania pomogły m.in. zdefiniować najpopularniejsze serwisy, które można by polecać doktorantom (młodym naukowcom) do podnoszenia swojej rozpoznawalności na forum międzynarodowym.

Zuza Wiorogórska wygłosiła referat podsumowujący wyniki pilotażowych badań porównawczych przeprowadzonych wśród doktorantów psychologii Uniwersytetu Warszawskiego i Uniwersytetu Lille 3, dotyczących potrzeb i zachowań informacyjnych tej grupy młodych naukowców. Wyniki badań nie wykazały szczególnych różnic między użytkownikami informacji z Polski i Francji. W obu krajach respondenci zwracali uwagę na potrzebę specjalistycznej edukacji informacyjnej w ich dziedzinie nauki, takiej, która pozwoliłaby na poszerzenie umiejętności informacyjnych i – co podkreślano w obu krajach – zaoszczędzenie czasu podczas wyszukiwania literatury przedmiotu. Problemami wskazywanymi przez respondentów były: wysokie ceny najnowszych wydań psychologicznych e-czasopism, ograniczony dostęp (np. swobodny dostęp dopiero po upływie okresu embarga) do wielu tytułów lub jego brak, wreszcie wysokie ceny, jakie trzeba zapłacić w przypadku indywidualnego zakupu poszczególnych artykułów. Podkreślano wartości płynące z dzielenia się wiedzą i wspólnego działania (ang. *Community of Practice*) – zarówno w świecie realnym (konferencje, współpraca krajowa i zagraniczna) jak i wirtualnym (media społecznościowe). Respondenci w obu krajach dostrzegli wartości, jakie niesie za sobą publikowanie w otwartym dostępie (ang. *Open Access*), ale albo uważali to za zbyt drogą (i dla młodych naukowców nieosiągalną) drogę publikacji wyników swoich badań, albo pragmatycznie orzekali, że najważniejsze czasopisma (zatem te posiadające obliczony wskaźnik Impact Factor) nie są wydawane w otwartym dostępie, a tylko publikacje w takich czasopismach dają gwarancję szybkiego rozwoju kariery naukowej.

Ewa A. Rozkosz przedstawiła wyniki badań nad usytuowanym uczeniem się kompetencji informacyjnych przez małą społeczność badaczy pedagogiki krytycznej. Celem było poznanie, zrozumienie i opisanie sposobów uczenia się przez jej członków kompetencji informacyjnych. Jako ramę teoretyczną wykorzystała teorię usytuowanego uczenia się Leave i Wengera, z którą wiąże się pojęcie wspólnoty praktyk (ang. *Community of Practice*). Uczenie się kompetencji rozumiane było jako angażowanie się we wspólnotę praktyk informacyjnych, poprzez: (1) odtwarzanie zastanych praktyk, polegające na przesuwaniu się uczestników wspólnoty z jej obrzeży do centrum, w wyniku poznawania przez nich znaczeń, reguł i działań właściwych dla wspólnoty, (2) ich transformację, będącą następstwem napięć i negocjacji nowych znaczeń dla działań podejmowanych przez wspólnotę. Społeczność wypracowała podzielane przez członków znaczenia dla różnych sposobów korzystania z informacji. Osoby włączające się do społeczności nabywały zastane wzory i reguły (zachowań informacyjnych). Doświadczenia biograficzne członków społeczności, w tym naukowe wyjazdy zagraniczne, jak również współpraca z organizacjami pozarządowymi przy realizacji projektów naukowych, wpływały na zmianę znaczeń nadawanych określonym działaniom (zachowaniom informacyjnym) i były punktem wyjścia do ich zmiany. Jednocześnie społeczność mogła być hamulcem zmian (transformacji zachowań informacyjnych), jeżeli propozycje zmian nie uzyskiwały aprobaty członków zajmujących wyższą pozycję w hierarchii tej społeczności.

Bez względu na obszar, w którym uczestnik konferencji sytuuje siebie (jako badacza kompetencji informacyjnych lub ich edukatora), z pewnością jego uwagę przyciągnąć powinny toczące się podczas ECIL dyskusje nad pojęciem *information literacy*. I tak wspomniana wcześniej Carol Collie Kuhlthau zdefiniowała *information literacy* jako umiejętność zastosowania informacji do tworzenia wiedzy potrzebnej do mądrych działań (oryg. *ability to use information to construct knowledge for wise action*). Z kolei Heidi Julien określiła je jako zestaw umiejętności, wiedzy i postaw wymaganych, aby efektywnie, sprawnie i w etyczny sposób uzyskiwać dostęp do informacji (oryg. *set of skills, knowledge and attitudes required to access information effectively, efficiently and ethically*). W różnych definicjach akcentowano odmienne wymiary kompetencji informacyjnych: krytyczny (np. zdolność do

refleksji nad ochroną swojego wizerunku w mediach), technologiczny (np. umiejętność korzystania z narzędzi do zarządzania informacją), ekologiczny (np. rozumienie zjawiska przeciążenia informacją) i inne. Zwracano uwagę na związek definicji pojęcia z kontekstem, w jakim ono się pojawia. Inaczej rozumiane są kompetencje informacyjne w badaniach efektów edukacji informacyjnej akademickiej czy szkolnej oraz w badaniach nad uczeniem się kompetencji informacyjnych w miejscu pracy. Warto odnotować, że w referatach poświęconych działaniom podejmowanym przez biblioteki w obszarze edukacji informacyjnej, kompetencje informacyjne były często utożsamiane z umiejętnością korzystania z różnych źródeł informacji oferowanych przez te instytucje. Dyskusje nad pojęciem *information literacy* toczące się podczas ECIL nie służyły więc wypracowaniu ujednoliconej definicji, ale podkreśleniu różnorodności podejść badawczych, kontekstów, stanowisk oraz praktyki. To podejście może inspirować do konstruowania nowych pytań badawczych i wykorzystywania nowych perspektyw w badaniach fenomenu *information literacy*.

Zagadnienia poruszane na konferencjach ECIL odzwierciedlają najnowsze tendencje w badaniach kompetencji informacyjnych oraz edukacji informacyjnej i ich zastosowaniu w praktyce. Na gruncie nauki i szkolnictwa wyższego warto zwrócić uwagę na to, że coraz więcej mówi się o edukacji informacyjnej dla doktorantów i pracowników uczelni, najczęściej w kontekście pomocy w przygotowywaniu publikacji naukowych, głównie tych przeznaczonych do obiegu międzynarodowego. Coraz częściej przedmiotem pogłębionej refleksji są zagadnienia węższe niż ogół kompetencji informacyjnych, np. podstawowa umiejętność związane z prawem autorskim (ang. *copyright literacy*) lub wykorzystywaniem surowych danych badawczych (ang. *data literacy*). Ponadto – edukacja informacyjna prezentowana jest jako swojego rodzaju promocja biblioteki i jej zbiorów. Im lepiej wyszkoleni użytkownicy, tym pełniej i chętniej będą korzystać z zasobów informacji oferowanych przez uczelnię.

Polskę na ECIL 2015 reprezentowały Ewa Rozkosz (DSW) i Zuza Wiorogórska (UW).

ECIL 2016 odbędzie się w dniach 10–13 października 2016 r. w Pradze.

## Bibliografia

Kurbanoglu, S.; Boustany, J.; Spiranec, S.; Grassian, E.; Mizrachi, D.; Roy, L., eds. (2015). *Information Literacy: Moving Toward Sustainability*. Cham: Springer International Publishing.

Zuza Wiorogórska

Biblioteka Uniwersytecka w Warszawie

Ewa A. Rozkosz

Dolnośląska Szkoła Wyższa

Nadesłano: 2 marca 2016.

## „Digital Humanities 2016”

(Kraków, 11–15 lipca 2016 r.)

W dniach 11–15 lipca 2016 r. odbyła się międzynarodowa konferencja naukowa pt. „Digital Humanities 2016”, która należy do najważniejszych konferencji z zakresu cyfrowej humanistyki. Była to już 28 edycja tej corocznej konferencji, którą organizuje Alliance of Digital Humanities Organizations (ADHO). Na miejsce obrad wybrano Kraków – po raz drugi w historii konferencji miasto z Europy Środkowo-Wschodniej. Przewodniczącym Rady Programowej konferencji został wybrany Manfred Thaller (Universität zu Köln, Niemcy), członek komitetu programowego European Association for Digital Humanities. Wśród lokalnych partnerów wspierających ADHO w organizacji konferencji znalazły się Uniwersytet Jagielloński (UJ) oraz Uniwersytet Pedagogiczny im. Komisji Edukacji Narodowej (UP), reprezentowanych przez dr. hab. Macieja Edera z Instytutu Filologii Polskiej UP oraz dr. Jana Rybickiego z Instytutu Filologii Angielskiej UJ. W zespole polskich organizatorów znaleźli się również przedstawiciele bibliologii i informatologii: dr hab. Grażyna Wrona (UP), dr hab. Władysław Marek Kolasa oraz dr Iwona Pietrzkiewicz z Instytutu Nauk o Informacji UP.

Podczas pierwszego i drugiego dnia konferencji jej uczestnicy mieli możliwość wzięcia udziału w warsztatach. Spośród 31 proponowanych warsztatów zdecydowana większość dotyczyła wykorzystania konkretnych narzędzi lub technik badawczych przydatnych w pracy cyfrowym humanistom. Wśród prezentowanych narzędzi znalazły się aplikacje do tworzenia kolekcji cyfrowych, komputerowej analizy tekstów, wizualnej analizy danych, semantycznego przetwarzania dokumentów czy do rozpoznawania treści cyfrowych wersji zniszczonych manuskryptów. Uczestnicy warsztatów mogli dowiedzieć się więcej o metodach analizy nie tylko tekstów, ale także materiałów audio-wizualnych czy metadanych gromadzonych w bibliotekach cyfrowych. Oprócz tego, mieli okazję poznać bliżej możliwości efektywnego wykorzystania indeksów nazw geograficznych w badaniach naukowych oraz dobre praktyki angażowania studentów w prace ośrodków uprawiających cyfrową humanistykę, takich jak laboratoria cyfrowej humanistyki czy biblioteki naukowe.

Oficjalne rozpoczęcie „Digital Humanities 2016” odbyło się drugiego dnia konferencji. Ceremonię otwarcia uświetnił wykład prof. Agnieszki Zalewskiej (Instytut Fizyki Jądrowej PAN, przewodnicząca Europejskiej Organizacji Badań Jądrowych CERN), podczas którego autorka pokazała, czego organizacje związane z cyfrową humanistyką mogą uczyć się od CERN.

Podczas kolejnych trzech dni konferencji prelegenci przedstawiali wyniki swoich badań w formie krótkich lub dłuższych referatów w ramach 78 sesji, w których znalazło się łącznie 269 wystąpień. Ponadto, trzeciego dnia konferencji miała miejsce sesja posterowa poprzedzona tzw. poster slams – kilkuminutowymi wystąpieniami autorów posterów, podczas których opowiadali o tematyce swoich prac i zachęcali do ich obejrzenia. W sumie podczas sesji posterowej zaprezentowano 118 plakatów. Czwartego dnia w ramach *Busa lecture* Helen Agüera (National Endowment for the Humanities – NEH, USA) przedstawiła swoje obserwacje dotyczące cyfryzacji badań i projektów naukowych wspieranych przez NEH. Konferencję zakończył referat Claire Warwick (Durham University, Wielka Brytania) wygłoszony podczas ceremonii zamknięcia „Digital Humanities 2016”. W swoim wystąpieniu autorka zastanawiała się z czego wynika niesłabnąca popularność papierowych źródeł informacji pomimo ich digitalizacji prowadzonej na szeroką skalę.

Nie sposób w jednym sprawozdaniu chociaż pokrótce streścić wszystkie wystąpienia, które zostały wygłoszone podczas „Digital Humanities 2016”, dlatego poniżej zostaną scharakteryzowane główne tematy referatów, paneli czy posterów zilustrowane wybranymi przykładami.

Opis tematyki wystąpień warto zacząć od stwierdzenia, że zdecydowana większość referatów czy posterów dotyczyła konkretnych projektów cyfrowej humanistyki. Niewiele miejsca poświęcono teoretycznym rozważaniom nad istotą cyfrowej humanistyki, kierunkom jej rozwoju czy metodom badawczym wykorzystywanym w pracy naukowej cyfrowych humanistów.

Wiele z prezentowanych projektów dotyczyło infrastruktury badawczej i struktur organizacyjnych cyfrowej humanistyki (przedstawiono np. HathiTrust Research Center oraz National Data Center for the Humanities) oraz analizy i wykorzystania nowych mediów (takich jak blogi, Wikipedia, Twitter czy inne serwisy społecznościowe). Często przedstawiano także projekty z zakresu stylometrii komputerowej, eksploracji danych i tekstu (ang. *topic modeling*, *text mining*, *data mining*) oraz analizy zasobów sieciowych. Dzięki wspomnianym technikom naukowcy badali, np. autorstwo tekstów antycznych (*Authorship Verification with the Ruzicka Metric*, Mike Kestemont, Justin Stover, Moshe Koppel, Folger Karsdorp i Walter Daelemans), czy też poufność w nowożytnej korespondencji transkonfesyjnej (*Exploring Networks Of Confidentiality And Secrecy In Early Modern Transconfessional Correspondences*, Ingeborg van Vugt). Nie zabrakło także projektów opartych na wizualizacji wyników badań, np. zaprezentowano mapę literatury i kultury okresu romantyzmu (*Digitally Mapping Romantic Literature and Culture*, Christopher Donaldson, Matthew Sangster, Joanna Taylor).

Wiele uwagi poświęcono także standardom i technologiom Semantic Web, m.in. XML, Linked Open Data czy RDF oraz przykładom ich implementacji. Do tej grupy prac można zaliczyć m.in. referat pt. *Converting the Liddell Scott Greek-English Lexicon into Linked Open Data using lemon*. Autorzy – Fahad Khan (University of Ca' Foscari Venezia, Włochy), Francesca Frontini (Italian National Research Council, Włochy), Monica Monachini (Istituto di Linguistica Computazionale "A. Zampolli", Włochy) oraz Federico Boschetti (Italian National Research Council, Włochy), przedstawili metodologię oraz cząstkowe wyniki projektu, którego celem jest konwersja słownika *Liddell Scott Greek-English Lexicon* do postaci semantycznej bazy wiedzy. Grupa badaczy z Uniwersytetu Bolońskiego (Włochy) – Marilena Daquino, Silvio Peroni, Francesca Tomasi i Fabio Vitali w referacie pt. *The Project Zeri Photo Archive: Towards a Model for Defining Authoritative Authorship Attributions* podjęła problem wykorzystania technologii Sieci Semantycznej do reprezentacji informacji na temat autorstwa dokumentów ikonograficznych w archiwum cyfrowym Zeri Photo Archive.

W wystąpieniach poruszono też kwestię crowdsourcingu oraz kształcenia w zakresie humanistyki cyfrowej. Np. Francesco Beretta (CNRS – Université de Lyon, Francja) omówił wyniki badań nad problemami studentów historii dotyczącymi użytkowania cyfrowych systemów informacyjnych oraz próbami ich rozwiązania (*From Index Cards to a Digital Information System: Teaching Data Modeling to Master's Students in History*). Nie zabrakło również projektów dotyczących budowy i analizy korpusów językowych. Jeden z takich projektów przedstawili Berenike Herrmann i Gerhard Lauer (oboje z Göttingen University, Niemcy) w referacie *KARREK: Building and Annotating a Kafka/Reference Corpus* dotyczącym cyfrowego korpusu tekstów Franza Kafki.

Przedmiotem referatów były również zagadnienia z obszaru bibliologii i informatologii. W wielu wystąpieniach poruszano kwestię wykorzystania ontologii do opisu kolekcji cyfrowych zgromadzonych w bibliotekach czy archiwach cyfrowych. Do tej grupy prac zaliczyć można referat Terhi Nurmikko-Fuller (University of Oxford, Wielka Brytania), Jacob Jett (University of Illinois at Urbana-Champaign, USA), Timothy Cole (University of Illinois at Urbana-Champaign, USA), Chris Maden (University of Illinois at Urbana-Champaign, USA), Kevin R. Page (University of Oxford, Wielka Brytania) i J. Stephen Downie (University of Illinois at Urbana-Champaign, USA) pt. *A Comparative Analysis of Bibliographic Ontologies: Implications for Digital Humanities*. W wystąpieniach często podejmowano temat tworzenia cyfrowych repozytoriów, bibliotek oraz archiwów i wykorzystywania ich kolekcji w badaniach naukowych. W tym miejscu warto wspomnieć o warsztatach *Data Mining Digital Libraries* dotyczących komputerowej analizy metadanych oraz materiałów publikowanych w bibliotekach cyfrowych, które przygotowali i przeprowadzili bibliotekarze z Biblioteki Narodowej Norwegii – Lars Gunnarsonn Johnsen, Magnus Breder Birkenes i Arne Martinus Lindstad).

Oprócz metadanych w kolekcjach cyfrowych analizom poddawano również bibliografie publikacji naukowych (referat pt. *A Study of Knowledge Integration in Digital Humanities Based on Bibliographic Analysis* autorstwa Muh-Chyun Tang, Yun Jen Cheng, Kuang Hua Chen i Jieh Hsiang z National Taiwan University na Tajwanie), rekordy z katalogów bibliotecznych (referat pt. *Printing in a Periphery: a Quantitative Study of Finnish Knowledge Production, 1640–1828* autorstwa Mikko Tolonen, Niko Iilomäki, Hege Roivainen i Leo Lahti z Uniwersytetu Helsińskiego w Finlandii) oraz cytowania kolekcji cyfrowych poświęconych I wojnie światowej (referat pt. *The Great War on the Web: the Making of Citing and Referencing by Amateurs* autorstwa Valérie Beaudouin i Zeynep Pehlivan z Politechniki Telecom ParisTech w Paryżu). Zaprezentowano także narzędzie do komputerowego przetwarzania przypisów (poster pt. *CTRaCE: Canonical Text Reader and Citation Exporter* autorstwa Martina Reckziegela, Stefana Jänicke i Gerika Scheuermanna z Uniwersytetu w Lipsku w Niemczech).

W wystąpieniach podejmowano również temat elektronicznych publikacji naukowych oraz otwartości w dzieleniu się wynikami prac badawczych (Open Access i Open Data). W tym kontekście Nora McGregor, Mia Ridge, Stella Wisdom i Aquiles Alencar-Brayner (British Library, Wielka Brytania) w referacie pt. *The Digital Scholarship Training Programme at British Library: Concluding Report & Future Developments* zwrócili uwagę na rolę bibliotek w kształceniu kompetencji cyfrowych przydatnych cyfrowym humanistom. W ten sposób biblioteki włączają się w realizację koncepcji cyfrowej nauki (ang. *digital scholarship*).

Analizując treść wygłoszonych referatów i zaprezentowanych posterów pod kątem reprezentowanych przez nie dyscyplin naukowych i kierunków badawczych, można zauważyć olbrzymią różnorodność podejmowanych zagadnień. Przedstawiono prace z zakresu historii literatury, stylistyki, lingwistyki, historii i historiografii, archiwistyki, bibliologii, informatologii, paleografii, folklorystyki, muzykologii, historii sztuki, pedagogiki, dydaktyki czy medioznawstwa.

Warto również wymienić przedstawicieli polskich ośrodków badawczych, którzy wystąpili na konferencji w roli prelegentów. Maciej Maryl (IBL PAN, Warszawa), Maciej Piasecki i Ksenia Młynarczyk (oboje z Politechniki Wrocławskiej) zaprezentowali analizę gatunków blogów opartą na metodach komputerowych (*Where Close and Distant Readings Meet: Text Clustering Methods in Literary Analysis of Weblog Genres*). M. Maryl wraz z Piotrem Wciślikiem (obydwaj z IBL PAN, Warszawa), przedstawili zagadnienie remediacji Polskiej Bibliografii Literackiej (*Remediations of Polish Literary Bibliography: Towards a Lossless and Sustainable Retro-Conversion Model for Bibliographical Data*). Krzysztof Nowak (Instytut Języka Polskiego PAN, Kraków) przedstawił aplikację-agregator cyfrowych zasobów dla łaciny średniowiecznej (*Medialatinitas.eu. Towards Meaningful Integration and Retrieval of Resources for Medieval Latin*).

Duża grupa prac polskich badaczy dotyczyła zagadnienia stylometrii. W ramach panelu *Literary Concepts: The Past and the Future*, Maciej Eder (UP, Kraków) przedstawił swoje badania dotyczące pojęcia tekstu, a Jan Rybicki (UJ) pojęcia przekładu. Maciej Eder (UP, Kraków), Maciej Piasecki i Tomasz Walkowiak (obaj z Politechniki Wrocławskiej) zaprezentowali WebSty – system wspomagający badania stylometryczne polskich dokumentów (*WebSty – an Open Web-based System for Exploring Stylometric Structures in Document Collections*). Jeremi Ochab (UJ) podjął temat wykorzystania analizy szeregów czasowych przy ustalaniu autorstwa tekstów (*Time Series Analysis Enhances Authorship Attribution*). Agata Hołobut, Jan Rybicki (oboje z UJ) i Monika Woźniak (Sapienza – Università di Roma, Włochy) przedstawili analizę stylometryczną dialogów filmowych (*Stylometry on the Silver Screen: Authorial and Translatorial Signals in Film Dialogue*). Maciej Eder (UP, Kraków) i Jan Rybicki (UJ) zaprezentowali badanie dotyczące autorstwa powieści pt. *Go Set a Watchman* (*Go Set A Watchman while we Kill the Mockingbird in Cold Blood, with Cats and Other People*). M. Eder (UP, Instytut Języka Polskiego PAN, Kraków) i Rafał Górski (UJ, Instytut Języka Polskiego PAN, Kraków) omówili wykorzystanie stylometrii do badania zmian zachodzących w języku na przestrzeni lat (*Historical Linguistics' New Toys, or Stylometry Applied to the Study of Language Change*). Michał Choiński i Jan Rybicki (obaj z UJ) pokazali badanie dotyczące wpływu redaktora na kształt dzieła

na przykładzie prac Jonathana Edwardsa (*Jonathan Edwards and Thomas Foxcroft: In Pursuit of Stylometric Traces of the Editor*).

Polskich naukowców interesowała także kwestia korpusów i ich komputerowej eksploracji. Anna Skolimowska (UW) zaprezentowała korpus tekstów i korespondencji Jana Dantyszka (*Corpus of Ioannes Dantiscus' Texts and Correspondence dantiscus.al.uw.edu.pl*), Jagoda Marszałek (Instytut Języka Polskiego PAN, Kraków) omówiła wyniki badań przeprowadzonych na korpusie tekstów Mateusza z Krakowa (*The Latin of Matthew of Cracow (c. 1345–1410) – a corpus based study of his language and style*), a Anna Ledzińska (Instytut Języka Polskiego PAN, Kraków) wyniki badań przeprowadzonych na korpusie polskiej średniowiecznej hagiografii (*The Holy and the Godless – Cultural Stereotypes Featured in the Language of the Polish Medieval Hagiography. A Corpus-based Study*). Adam Tomasz Pawłowski (UWr) scharakteryzował ChronoPress – korpus polskich tekstów prasowych (*Chronological corpora: Challenges and opportunities of sequential analysis. The example of ChronoPress corpus of Polish*).

Kilka wystąpień było poświęconych e-literaturze. Katarzyna Maria Bazarnik (UJ) i Jakub Wróblewski (ASP, Warszawa) pokazali multimedialną wersję powieści *Finnegans Wake* Jamesa Joyce'a (*First We Feel Then We Fall – Multimedia Adaptation of Joyce's Finnegans Wake*). Katarzyna Maria Bazarnik (UJ) i Zenon Fajfer (poeta niezależny, Kraków) prezentowali multimedialną poezję autorstwa Zenona Fajfera (*Playable Books at Electronic Literature's Interface*). Monika Górską-Olesińską (Uniwersytet Opolski) i Mariusz Pisarski (UW) analizowali problem przekładu literatury elektronicznej z języka angielskiego na język polski na przykładzie dzieła (nazywanego też generatorem poezji) *Sea and Spar Between* autorstwa Stephanie Strickland i Nick Montfort (*Translating Electronic Literature. Multicultural, Multilingual and Cross-Platform Encounters*).

Wystąpienia wygłoszone podczas tegorocznej konferencji „Digital Humanities” nakreślają współczesny obraz cyfrowej humanistyki, jej obszary zainteresowań oraz kierunki rozwoju. Zaprezentowane prace odznaczały się wysokim poziomem merytorycznym. Księga abstraktów dostępna jest na stronie internetowej konferencji pod adresem <http://dh2016.adho.org/abstracts/>. Konferencja cieszyła się dużym zainteresowaniem ze strony międzynarodowego środowiska naukowego – wzięło w niej udział ponad dziewięćset osób. Dzięki pracy i zaangażowaniu organizatorów badacze z całego świata mogli uczestniczyć w ciekawym i inspirującym wydarzeniu naukowym, ważnym dla humanistyki XXI wieku.

Anna Kamińska

Katedra Informatologii  
Wydział Dziennikarstwa, Informacji i Bibliologii  
Uniwersytet Warszawski

Nadesłano: 11 sierpnia 2016.

## Wskazówki dla autorów

Redakcja *Zagadnień Informatyki Naukowej – Studiów Informatycznych* przyjmuje wyłącznie teksty wcześniej nieopublikowane i niezłożone do druku w innych czasopismach lub pracach zbiorowych. Przyjmowane są: oryginalne rozprawy i prace badawcze, recenzje oraz sprawozdania z konferencji i innych wydarzeń naukowych.

Teksty artykułów są recenzowane zgodnie z zasadami double-blind peer review. Zapewnienie anonimowości tekstów przekazywanych do recenzji wymaga, aby w tekście artykułu w żadnym miejscu nie była umieszczona informacja umożliwiająca identyfikację autora.

Każdy artykuł recenzowany jest na podstawie jednolitego formularza przez dwóch recenzentów dobieranych spośród specjalistów problematyki w nim poruszanej. Każda recenzja zawiera jednoznaczne wskazanie czy tekst rekomendowany jest do publikacji w *Zagadnieniach Informatyki Naukowej*. Podstawowymi kryteriami oceny artykułu są: zgodność tematu z profilem czasopisma, wartość merytoryczna, organizacja logiczna i forma językowa tekstu.

O przyjęciu tekstu do publikacji autorzy informowani są w ciągu 10 tygodni od otrzymania go przez Redakcję. Redakcja przyjmuje wyłącznie teksty przygotowane zgodnie z zasadami przedstawionymi poniżej. Teksty należy nadsyłać na adres e-mail: [zin.iinsb@uw.edu.pl](mailto:zin.iinsb@uw.edu.pl)

### 1. Zasady ogólne

#### 1.1. Format

Wszystkie pliki (tekst artykułu, materiały ilustracyjne) należy przysyłać jako dokumenty edytora MS WORD w formacie RTE. Zaleca się stosować w tekście czcionkę Times New Roman 12 pkt, interlinię 1.5. Tytuł artykułu należy wyróżnić czcionką Times New Roman 16 pkt. **Nie należy używać automatycznych stylów.**

Materiały ilustracyjne, wstawione w treść artykułu, dodatkowo należy przysyłać również w formacie JPG. Załączniki powinny być ponumerowane według kolejności występowania w tekście oraz zawierać nazwę, np.: *1. Tabela 1. Poziomy metadanych* albo *3. Rysunek 1. Mapa myśli*.

#### 1.2. Długość tekstu

Artykuł nie powinien przekraczać 40 000, a recenzja lub sprawozdanie 14 000 znaków (ze spacjami).

#### 1.3. Strona tytułowa

Autorzy artykułów proszeni są o przygotowanie odrębnej strony tytułowej, zawierającej:

- tytuł artykułu (w językach polskim i angielskim)
- dane autora (imię i nazwisko, afiliacja – w językach polskim i angielskim)
- adres e-mail
- adres do korespondencji
- notę biograficzną autora (patrz niżej)
- abstrakt ustrukturyzowany (patrz niżej)
- słowa kluczowe (patrz niżej)
- oświadczenie o oryginalności tekstu (patrz niżej).

Zgodnie z zasadami przeciwdziałania zjawiskom *ghostwritingu* i *guest authorship* Redakcja prosi również, aby na tej stronie ujawnione zostały nazwiska i afiliacje wszystkich osób, które przyczyniły się do powstania artykułu, ich rola i udział w przygotowaniu publikacji (kto jest autorem koncepcji, założeń, metod itp. wykorzystywanych w pracy zgłoszonej do druku; procentowy udział w przeprowadzonych badaniach i opracowaniu artykułu). Redakcja prosi także o podanie informacji o źródłach finansowania publikacji, wkładzie instytucji naukowo-badawczych, stowarzyszeń i innych podmiotów (*financial disclosure*).

#### 1.4. Nota biograficzna autora / autorów

Na stronie tytułowej należy umieścić zwięzłą notę biograficzną (ok. 70 słów) każdego autora artykułu. Nota powinna zawierać następujące informacje: tytuł / stopień naukowy lub zawodowy autora, aktualne miejsce pracy i zajmowane stanowisko; specjalności naukowe lub zawodowe, najważniejsze publikacje (max. 3).

### 1.5. Abstrakt ustrukturyzowany

Na stronie tytułowej należy umieścić abstrakt w języku polskim o objętości ok. 100 słów (ok. 1 tys. znaków) oraz jego przekład na język angielski. W abstrakcie należy wyróżnić co najmniej cztery spośród następujących kategorii informacji:

- Cel/teza | Purpose/thesis (obowiązkowo)
- Koncepcja/metody badań | Approach/methods (obowiązkowo)
- Wyniki i wnioski | Results and conclusions (obowiązkowo)
- Ograniczenia badań | Research limitations (opcjonalnie)
- Zastosowanie praktyczne | Practical implications (opcjonalnie)
- Oryginalność/wartość poznawcza | Originality/value (obowiązkowo)

### 1.6. Słowa kluczowe

Na stronie tytułowej artykułu należy umieścić od 4 do 10 słów kluczowych, w formie fraz nominalnych w mianowniku liczby pojedynczej, których pierwszy wyraz zapisany jest wielką literą, uporządkowanych alfabetycznie, rozdzielonych kropkami. Słowa kluczowe należy podać w językach polskim i angielskim.

### 1.7. Oświadczenie o oryginalności tekstu

Na stronie tytułowej artykułu należy umieścić oświadczenia autora /autorów, że tekst przedstawiany Redakcji *Zagadnień Informacji Naukowej – Studiów Informacyjnych* nie był dotychczas opublikowany ani zgłoszony do publikacji w żadnym innym czasopiśmie lub pracy zbiorowej. Jeśli tekst był prezentowany na konferencji, należy podać jej szczegółowe dane wraz z ewentualnymi informacjami o publikacji materiałów konferencyjnych. Jeśli artykuł jest częścią przygotowywanej do druku książki, należy podać jej dane oraz planowany termin publikacji.

## 2. Zasady opracowania artykułu

### 2.1. Organizacja i podział tekstu

Tekst artykułu powinien być podzielony na podrozdziały zaopatrzone w tytuły. W pierwszej części pod nagłówkiem **Wprowadzenie** zaleca się umieścić informacje wprowadzające w problematykę prezentowaną w artykule. W części ostatniej – pod nagłówkiem **Wnioski** lub **Zakończenie** – wnioski końcowe i podsumowanie przedstawionych rozważań.

Dopuszcza się stosowanie do trzech poziomów podziału tekstu, każdy wyodrębniony własnym śródtytułem i opatrzonego oznaczeniem numerycznym zgodnie z następującymi regułami:

#### 1. Pierwszy poziom podziału

##### 1.1. Drugi poziom podziału

###### 1.1.1 Trzeci poziom podziału

### 2.2. Przypisy

Nie stosuje się przypisów bibliograficznych. Odesłania do wykorzystanej literatury należy przygotować zgodnie z edytorskimi standardami tekstu naukowego APA 6<sup>th</sup> (patrz niżej).

Przypisy zawierające komentarze, dygresje, objaśnienia i inne dodatkowe informacje należy umieszczać na dole strony i numerować liczbami arabskimi; zaleca się ograniczenie liczby przypisów do niezbędnego minimum.

### 2.3. Pisownia tytułów w tekście artykułu

Tytuły wystaw, konferencji, programów itp. powinny być zapisane w cudzysłowie. Tytuły publikacji (książek, czasopism, artykułów itp.) należy wyróżnić kursywą.

### 2.4. Wyróżnienia w tekście

W tekście można stosować wyróżnienia za pomocą czcionki półgrubej (bold).

## 2.5. Materiały ilustracyjne i ich oznaczanie w tekście

Materiały ilustracyjne (tabele, wykresy itp.) powinny być przygotowane w odcieniach szarości lub kolorystyce czarno-białej. Wszystkie tego typu materiały należy oznaczyć wskazaniem rodzaju materiału (np. Tabela, Rysunek, Fotografia, Wykres), jego numeru w tekście oraz jego tytułu (np. Tabela 1. Poziomy metadanych). W odpowiednich miejscach tekstu artykułu należy umieścić odesłania do informacji prezentowanych w formie ilustracji, używając w tym celu skrótu określenia rodzaju ilustracji oraz jej numeru (np. zob. Tabela 1, zob. Wykres 5).

## 2.6. Cytowanie wykorzystanej literatury w tekście i bibliografia załącznikowa

Cytowania w tekście i bibliografię załącznikową należy przygotować zgodnie ze standardami edytorskim publikacji naukowych APA 6<sup>th</sup>. W bibliografii załącznikowej mogą być umieszczone wyłącznie opisy publikacji cytowanych w tekście artykułu.

Publikacje należy cytować w tekście używając odsyłaczy w formie: (nazwisko, rok wydania), np. (Dembowska, 1991); gdy publikacja ma dwóch autorów należy podać obydwa nazwiska połączone znakiem ampersand (nazwisko1 & nazwisko2, rok), np. (Cisek & Sapa, 2007); gdy publikacja ma trzech i więcej autorów należy podać nazwisko pierwszego autora, skrót *et. al.* i rok wydania (nazwisko1 et al., rok), np. (Berners-Lee et al., 2001); gdy publikacja jest pracą zbiorową, należy podać nazwisko redaktora, skrót *red.* i rok wydania (nazwisko, *red.*, rok), np. (Kocójowa, *red.*, 2005). Jeśli w publikacji nie wskazano nazwiska autora lub redaktora, należy podać pierwszy wyraz tytułu zapisany kursywą, trzy kropki i rok wydania (*Wyraz...*, rok), np. (*Biblioteki...*, 1976). Odwołania do określonych stron cytowanych tekstów należy podawać w formie: (Dembowska, 1991, 15), albo (Cisek & Sapa, 2007, 40–42), (Dervin & Nilan, 1986, 3) albo (Kocójowa, *red.*, 2005, 18).

Opisy bibliograficzne wykorzystanych publikacji należy umieścić na końcu tekstu w układzie alfabetycznym, bez numeracji pozycji, pod nagłówkiem **Bibliografia**.

Opisy autorskich książek i artykułów umieszcza się pod nazwiskiem pierwszego autora. Opisy prac zbiorowych należy umieszczać pod nazwiskiem redaktora, po którym podaje się skrót *red.* lub *ed.* Jeśli w publikacji nie wskazano autora lub redaktora pracy zbiorowej, jej opis należy umieścić pod pierwszym wyrazem tytułu.

Tytuły książek i czasopism należy zapisać kursywą, tytuły artykułów w czasopismach i artykułów lub rozdziałów w książkach – czcionką prostą.

W opisach artykułów w pracach zbiorowych stosuje się oznaczenie skrótu „W” dla publikacji w języku polskim i „In” dla publikacji w językach obcych.

Opisy prac tego samego autora powinny być uporządkowane według chronologii wstępującej, a w każdym z nich należy powtórzyć nazwisko i inicjał (inicjały) autora. Prace tego samego autora opublikowane w tym samym roku należy uporządkować w kolejności alfabetycznej tytułów i oznaczać wg zasady:

Dembowska, M. (1976a) ...

Dembowska, M. (1976b) ..., itd.

### 2.6.1 Przykłady redagowania opisów bibliograficznych

#### KSIĄŻKA

Breslin, J.G.; Passant, A.; Decker, S. (2009). *The Social Semantic Web*. Berlin: Heidelberg: Springer Verlag.

Dembowska, M. (1991). *Nauka o informacji naukowej: organizacja i problematyka badań w Polsce*. Warszawa: IINTE.

#### PRACA ZBIOROWA

Bellardo Hahn, T.; Buckland, M. (eds.). (1998). *Historical Studies in Information Science*. Medford, NJ: Information Today.

Biblioteki (1976). *Biblioteki publiczne województwa toruńskiego: informator*. Toruń: Wojewódzka Biblioteka Publiczna i Książnica Miejska im. M. Kopernika.

Kocójowa, M. (red.). (2005). *Profesjonalna informacja w Internecie*. Kraków: Wydaw. UJ.

#### ARTYKUŁ W CZASOPIŚMIE

Dervin, B.; Nilan, M. (1986). Information Needs. *Annual Review of Information Science and Technology*, 21, 3–31.

Osińska, V. (2010). Rozwój metod mapowania domen naukowych i potencjał analityczny w nim zawarty. *Zagadnienia Informatyki Naukowej*, 96(2), 41–51.

## ARTYKUŁ W PRACY ZBIOROWEJ

- Rayward, W.B. (1998). Visions of Xanadu: Paul Otlet (1868–1944) and Hypertext. In: T. Bellardo Hahn & M. Buckland (eds.). *Historical Studies in Information Science*. Medford, NJ: Information Today, 65–80.
- Gawrysiak, P. (2000). W stronę inteligentnych systemów wyszukiwawczych. W: Cz. Daniłowicz (red.) *Multimedialne i sieciowe systemy informacyjne*. Wrocław: Oficyna Politechniki Wrocławskiej, 59–69.

## ARTYKUŁ W CZASOPISIE ELEKTRONICZNYM

- Berners-Lee, T.; Hendler, J.; Lassila, O. (2001). The Semantic Web. *Scientific American* [online], May, [30.06.2013], <http://www.scientificamerican.com/article.cfm?id=the-semantic-web>
- Hollender, H. (2013). SYNAT: dziesiątki dużych i małych pomysłów na informację naukową. *Biuletyn EBIB* [online], 135(8), [15.07.2013], [http://www.ebib.pl/?page\\_id=413#art6](http://www.ebib.pl/?page_id=413#art6)
- Miller, H. (2013). Big-data in cloud computing: a taxonomy of risks. *Information Research* [online], 18(1), [15.07.2013], <http://informationr.net/ir/18-1/paper571.html>

## HASŁA ENCYKLOPEDYCZNE

- Psychology of culture contact (1926). In: *Encyclopaedia Britannica*, Vol. 1, 13th ed., London and New York, NY: Encyclopaedia Britannica, 765–771.
- Iluminatorstwo (1971). W: *Encyklopedia Wiedzy o Książce*. Wrocław – Warszawa – Kraków: Zakł. Narod. im. Ossolińskich, 911–952.
- Big data. (2013, November 12). In: *Wikipedia, The Free Encyclopedia* [online] [12.11.2013], [http://en.wikipedia.org/w/index.php?title=Big\\_data&oldid=581347727](http://en.wikipedia.org/w/index.php?title=Big_data&oldid=581347727)

Autorskie artykuły encyklopedyczne należy opisywać tak jak artykuły w pracach zbiorowych.

## DOKUMENT Z WITRYNY INSTYTUCJI, ORGANIZACJI LUB OSOBY PRYWATNEJ

- Aristotle (2009). *Organon*. From 1a to 164 a according to Bekker numbers [online]. Translated under the editorship of W.D. Ross. Internet archive [29.10.2013], [http://archive.org/stream/AristotleOrganon/AristotleOrganon-collectedWorks\\_djvu.txt](http://archive.org/stream/AristotleOrganon/AristotleOrganon-collectedWorks_djvu.txt)
- MNiSW (2011). *Narodowe Centrum Nauki w Krakowie. Nadchodzi czas nauki* [online]. Ministerstwo Nauki i Szkolnictwa Wyższego, [15.07.2013], <http://www.nauka.gov.pl/?id=2268>
- Smith, B. (2004). *Ontology and information systems* [online]. The Buffalo University, Department of Philosophy, [15.07.2013], <http://ontology.buffalo.edu/ontology.doc>
- US NLM (2004). *Medical Subject Headings* [online]. US National Library of Medicine. National Institutes of Health, [15.07.2013], <http://www.nlm.nih.gov/mesh/meshhome.html>

## Guidelines for Authors

ZIN – *Studia Informacyjne* (ZIN – *Information Studies*) accepts only manuscripts that have not been published before and are not under consideration for publication anywhere else. Following types of paper may be submitted for publication: original papers, book reviews, conference (and other events) reports.

Each manuscript is reviewed under a double-blind peer review process. In order to ensure the anonymity of the review process, please do not place any information in the text that could be used to identify the author.

Each manuscript is reviewed by two referees, selected on the basis of necessary expertise in the subject area under review. The review report is based on standard form containing a statement whether the manuscript is recommended for publication. Criteria for acceptance include appropriateness to the field of the Journal, scientific merit, proper text organization and correct language use.

The final decision about publication of manuscript will be sent to Author within 10 weeks after text submission. Manuscript should be formatted according to guidelines listed below and submitted via e-mail: [zin.iinsb@uw.edu.pl](mailto:zin.iinsb@uw.edu.pl)

### 1. General guidelines

#### 1.1. Format

All files should be submitted in RTF (Rich Text Format) files, including text and illustrative content. All pages must be typed and 1.5 spaced using 12-point Times New Roman font. The title of the manuscript should be typed 14-point font. Please do not use any preformatted styles.

Illustrative content inserted in the article, should be send also in JPG format. Attachments should be numbered in order of occurrence and include the title, for example: *1. Table 1. List...* or *3. Figure 1. System....*

#### 1.2. Extent

Manuscript should be no longer than 40,000 characters (including spaces), review and report no longer than 14,000 characters.

#### 1.3. Title page

Authors should prepare **separate title page**, which include:

- title of the paper,
- the name(s) of the author(s) with appropriate affiliations,
- the e-mail address of the corresponding author,
- address for correspondence,
- biographic note (see below),
- structured abstract (see below),
- keywords (see below),
- statement of originality (see below).

According to the Journal policy against *ghostwriting* and *guest authorship*, authors are requested to list on title page names and affiliations of each person that contributed to the text (author of the idea, methods, etc. used in the submitted manuscript; percentage of contribution to the research process and text compilation). Authors are also requested to describe sources of founding that have supported the work and the financial involvement of research institutes, associations and other entities (*financial disclosure*).

#### 1.4. Author(s) biographic note

Title page should include concise biographic notes (about 70 words) of each author : academic degree or professional position, current place of work and position, area of interest, the most important publications (max. 3).

#### 1.5. Structured abstract

An abstract (about 100 words or 1000 characters) should be included with each submission and placed on the title page. Abstract should be formatted according to categories listed below. Author should identify at least four mandatory sections:

- **Purpose/thesis** (*mandatory*)
- **Approach/methods** (*mandatory*)
- **Results and conclusions** (*mandatory*)
- **Research limitations** (*optional*)
- **Practical implications** (*optional*)
- **Originality/value** (*mandatory*)

### **1.6. Keywords**

Title page should include keywords (4 to 10) as a noun phrases in singular form, where first element is capitalized. Keywords in alphabetical order should be delimited by full stop.

### **1.7. Statement of originality**

Author(s) should include on title page statement that submitted text has not been published before and is not under consideration for publication anywhere else. If the paper was presented at a scientific meeting, provide detailed information about the event and the conference proceedings. If the paper will be the part of the author's book, provide its details and planned publishing date.

## **2. Manuscript Format and Preparation**

### **2.1. Body of the Paper**

The text should be organized into entitled sections and subsections. Text should start with **Introduction**, giving an overview and stating the purpose and end with **Conclusion**, giving the summary of the author contributions to the study.

Author may use three levels of headings. Each heading should have its own title and number according to the following pattern:

#### **1. First-level heading**

##### **1.1. Second-level heading**

###### **1.1.1 Third-level heading**

### **2.2. References**

Bibliographic citations are not allowed in footnotes. The reference list should be prepared according to APA 6-th Edition citation style (see below). Footnotes can be used only to give additional information or commentary. Footnotes to the text are numbered consecutively with Arabic numerals. It is recommended to limit the amount of footnotes per page.

### **2.3. Titles in the body of the text**

Titles of exhibitions, conferences, programmes, etc should be written within double quotation marks. Use italics for publication titles (books, journals, papers, etc.).

### **2.4. Emphasis**

Bold face should be used to emphasize certain words or passages.

### **2.5. Illustrative content**

All illustrations (tables, charts, figures etc.) should be converted to greyscale. All illustrations should be cited in the text properly to their form (Table, Figure, Photograph, etc.) and have title and consecutive number (e.g. Table 1. Metadata levels). Use abbreviation in the text when refereeing to the illustrative content (e.g. see Table 1, see Figure 5).

### **2.6. Citations and reference list**

Use APA 6-th Edition as a citation and reference list format. The references list should only include works that are cited in the text.

Cite references in the text by name of the author(s) and year of publication in parentheses: (Name, Year of publication), eg. (Dembowska, 1991). If there are two authors, put their names with ampersand (&) mark between:

(Name & Name, Year of publication), eg. (Cisek & Sapa, 2007). If there are more than two authors, put the name of the first one followed by abbreviation *et al.*: (Name et al., Year of publication), eg. (Berners-Lee et al., 2001). Edited books are cited by the name of the editor followed by abbreviation *Ed.*: (Name, Ed., Year of publication), eg. (Kocojowa, Ed., 2005). If there is no author or editor information, put the first word from the title in italics followed by ellipsis (...) and the year of publication : (Word..., Year of publication), eg. (*Libraries...*, 1995). Use the following pattern when referring to specific pages in the cited publications: (Dembowska, 1991, 15) or (Cisek & Sapa, 2007, 40–42) or (Kocojowa, Ed., 2005, 18).

Place the reference list at the end of the text under the heading **References**. Reference list should be in alphabetical order without numbering.

List the references (books and journal articles) in alphabetical order by authors' last names. Citations of edited books list under the name of editor followed by abbreviation *Ed.*. If there is no author or editor information, list the publication under the first word from the title.

Use italics for book titles and regular font for titles of papers and book chapters. Use abbreviation *In*: when referring to book chapters in citations.

If there are two or more items by the same author(s), list them in order of year of publication (reverse date order). If two or more works are by the same author(s) within the same year, list them in alphabetical order by title and distinguish them by adding the letters a, b, c, ... to the year of publication:

Dembowska, M. (1976a) ....

Dembowska, M. (1976b) ...., etc.

### 2.6.1 References List Examples

#### BOOK

Breslin, J.G.; Passant, A.; Decker, S. (2009). *The Social Semantic Web*. Berlin: Heidelberg: Springer Verlag.

Dembowska, M. (1991). *Nauka o informacji naukowej: organizacja i problematyka badań w Polsce*. Warszawa: IINTE.

#### BOOK (EDITED)

Bellardo Hahn, T.; Buckland, M. (eds.). (1998). *Historical Studies in Information Science*. Medford, NJ: Information Today.

Biblioteki (1976). *Biblioteki publiczne województwa toruńskiego: informator*. Toruń: Wojewódzka Biblioteka Publiczna i Książnica Miejska im. M. Kopernika.

#### JOURNAL ARTICLE

Osińska, V. (2010). Rozwój metod mapowania domen naukowych i potencjał analityczny w nim zawarty. *Zagadnienia Informatyki Naukowej*, 96(2), 41–51.

Dervin, B.; Nilan, M. (1986). Information Needs. *Annual Review of Information Science and Technology*, 21, 3–31.

#### BOOK CHAPTER

Rayward, W.B. (1998). Visions of Xanadu: Paul Otlet (1868-1944) and Hypertext. In: T. Bellardo Hahn & M. Buckland (eds.). *Historical Studies in Information Science*. Medford, NJ: Information Today, 65–80.

#### ELECTRONIC JOURNAL ARTICLE

Berners-Lee, T.; Hendler, J.; Lassila, O. (2001). The Semantic Web. *Scientific American* [online], May, [30.06.2013], <http://www.scientificamerican.com/article.cfm?id=the-semantic-web>

Hollender, H. (2013). SYNAT: dziesiątki dużych i małych pomysłów na informację naukową. *Biuletyn EBIB* [online], 135 (8), [15.07.2013], [http://www.ebib.pl/?page\\_id=413#art6](http://www.ebib.pl/?page_id=413#art6)

Miller, H. (2013). Big-data in cloud computing: a taxonomy of risks. *Information Research* [online], 18(1), [15.07.2013], <http://informationr.net/ir/18-1/paper571.html>

#### ARTICLE IN ENCYCLOPEDIA

Psychology of culture contact (1926). In: *Encyclopaedia Britannica*, Vol. 1, 13th ed., London and New York, NY: Encyclopaedia Britannica, 765–771.

Iluminatorstwo (1971). W: *Encyklopedia Wiedzy o Książce*. Wrocław – Warszawa – Kraków: Zakł. Narod. im. Ossolińskich, 911–952.

Big data. (2013, November 12). In: *Wikipedia, The Free Encyclopedia* [online] [12.11.2013], [http://en.wikipedia.org/w/index.php?title=Big\\_data&oldid=581347727](http://en.wikipedia.org/w/index.php?title=Big_data&oldid=581347727)

Article in encyclopedia with author information describe as book chapter.

#### ELECTRONIC DOCUMENT FROM WEBSITE

MNiSW (2011). *Narodowe Centrum Nauki w Krakowie. Nadchodzi czas nauki* [online]. Ministerstwo Nauki i Szkolnictwa Wyższego, [15.07.2013], <http://www.nauka.gov.pl/?id=2268>

Smith, B. (2004). *Ontology and information systems* [online]. The Buffalo University, Department of Philosophy, [15.07.2013], <http://ontology.buffalo.edu/ontology.doc>

US NLM (2004). *Medical Subject Headings* [online]. US National Library of Medicine. National Institutes of Health, [15.07.2013], <http://www.nlm.nih.gov/mesh/meshhome.html>

---

#### **Adres Wydawnictwa**

ul. Konopczyńskiego 5/7

00-335 Warszawa, tel. 22 827 52 96

#### **Prenumerata i sprzedaż**

Dział Promocji i Kolportażu SBP

Al. Niepodległości 213, 02-086 Warszawa, tel. 22 825 50 24

**Cena prenumeraty na 2016 r. – 114 zł**

Wydawnictwo SBP – Warszawa 2016. Nakład 350 egz.

Ark. wyd. 13,8. Ark. druk. 10,88.

Skład i łamanie: Justyna Grzymała

Druk i oprawa: Mazowieckie Centrum Poligrafii

ul. Piłsudskiego 2A, 05-270 Marki, [www.c-p.com.pl](http://www.c-p.com.pl)

e-mail: [biuro@c-p.com.pl](mailto:biuro@c-p.com.pl), tel. 22 497 66 55

