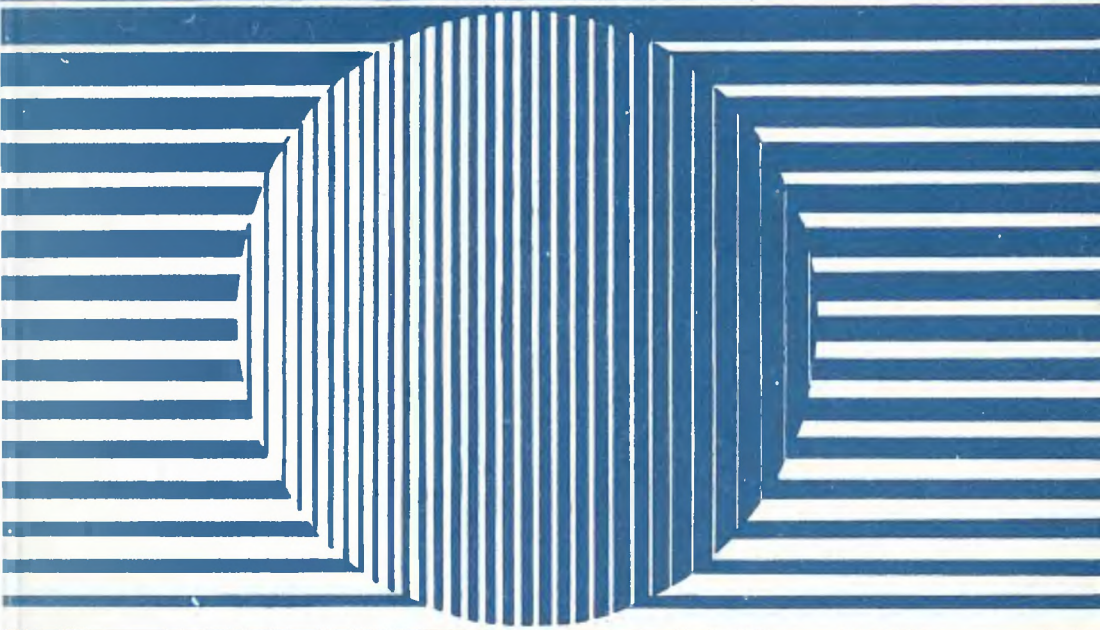


POLSKA AKADEMIA NAUK



OŚRODEK INFORMACJI NAUKOWEJ

PL ISSN 0324-8194

ZAGADNIENIA INFORMACJI NAUKOWEJ

1992

OSSOLINEUM

Nr 1 (59)

POLSKA AKADEMIA NAUK

OŚRODEK INFORMACJI NAUKOWEJ

ZAGADNIENIA INFORMACJI NAUKOWEJ

1992

Nr 1 (59)

ZAKŁAD NARODOWY IMIENIA OSSOLIŃSKICH
WYDAWNICTWO POLSKIEJ AKADEMII NAUK

Rada Redakcyjna

Bożenna Bojar (redaktor naczelny), Ewa Chmielewska-Gorczyca,
Alina Golińska, Andrzej Gromek, Anna Leśniewicz, (sekretarz redakcji),
Bronisław Ługowski, Mieczysław Muraszkiewicz, Hanna Popowska,
Jadwiga Sadowska, Henryk Szarski, Maria Szomańska

Do roku 1971 czasopismo ukazywało się pod tytułem
„BIULETYN ODIIN PAN”
W roku 1991 ukazał się tylko nr 1(58)

Adres Redakcji

Ośrodek Informacji Naukowej PAN
00-330, ul. Nowy Świat 72 (Pałac Staszica)

Maszynopis niniejszego numeru przekazano do Wydawnictwa w czerwcu 1992

Wykonano za składopisu dostarczonego przez OIN PAN

PL ISSN 0324-8194

Zakład Narodowy im. Ossolińskich. Wydawnictwo Wrocław 1992

Objętość: ark. wyd. 9,5, ark. druk. 13,5

Drukarnia Uniwersytetu Wrocławskiego.

WIEŚLAW BABIK

Ośrodek Informacji Naukowej PAN

MATERIAŁY KARTOGRAFICZNE JAKO ŹRÓDŁO INFORMACJI

Podstawowe określenia i rodzaje materiałów kartograficznych. Cechy dystynktywne materiałów kartograficznych.

PODSTAWOWE OKREŚLENIA I RODZAJE MATERIAŁÓW KARTOGRAFICZNYCH

W dziedzinie, która będzie przedmiotem naszych rozważań, panuje pewien nieład terminologiczny. Przejawia się to w różnym rozumieniu takich podstawowych terminów, jak: przedmiot, treść i temat mapy, a nawet samego terminu "mapa".^{1/} W związku z tym podajemy określenia podstawowych rodzajów materiałów kartograficznych.

^{1/} Komitet Normalizacyjny Stowarzyszenia Bibliotek Specjalnych (Committee on Standards, Geography and Map Division, Special Libraries Association (SLA) w "Standards for University Libraries" (College and Research Libraries News, Nr 4 (April 1979), s. 101-110 (Prepared by a joint committee established by the Association of Research Libraries and the Association of College and Research Libraries) proponuje używanie terminu "map" (mapa) w celu oznaczenia każdego elementu zbioru kartograficznego, na przykład takich elementów jak: mapy, atlasy, globusy, zdjęcia przestrzenne, obrazy satelitarne, bibliografie kartograficzne, literatura kartograficzna. Por. "Standards for university map collections". SLA GandM Bulletin nr 148 (June 1987).

Za "materiały kartograficzne" (ang. cartographic materials) zgodnie z International Standard Bibliographic Description (Cartographic Materials - ISBD(CM) uznano "wszystkie materiały przedstawiające całość lub część Ziemi lub innego ciała niebieskiego w jakiejś skali, takie jak: dwu- i trójwymiarowe mapy, globusy, diagramy blokowe, mapy selekcyjne, zdjęcia lotnicze i satelitarne, atlasy, spojrzenia z lotu ptaka itd."^{2/}.

Definicja ta, jakkolwiek ma charakter deskryptywny i wyliczający, nie podaje jednak kompletnego zestawu materiałów, które można uznać za materiały kartograficzne. W tym sensie jest więc ona definicją otwartą. Określenie materiałów kartograficznych w kategoriach "genus proximum et differentia specifica" sprowadza się do uznania ich za jeden z rodzajów materiałów bibliotecznych, różniących się od innych tym, że stanowią one obrazy części lub całości Ziemi lub innego ciała niebieskiego w jakiejś skali^{3/}.

^{2/} Por. International Standard Bibliographic Description for Cartographic Materials. Recommended by the Joint Working Group on the International Standard Bibliographic Description for Cartographic Materials set up by the IFLA Committee on Cataloguing and the IFLA Sub-section of Geography and Map Libraries. 2nd ed. London: IFLA International Office for UBC (draft). 1981. Po raz pierwszy normę tę opublikowało w 1977 roku Międzynarodowe Biuro ds. Uniwersalnej Rejestracji Bibliograficznej IFLA (IFLA International Office for UBC).

^{3/} Materiały kartograficzne, nazywane też "obiektami kartograficznymi" (por. O. Borowik, Z. Haczewska: Klasyfikacja geograficzna. Warszawa 1954) lub "kartografikami" (Por. H. Robótka: Rozwój form kancelaryjnych i współczesne rodzaje dokumentów archiwalnych. Dokumentacja geodezyjno-kartograficzna. Toruń 1985), tworzą zbiory kartograficzne, zaliczane do zbiorów specjalnych. Zbiory te obejmują dokumenty przedstawiające w sposób graficzny (lub plastyczny) za pomocą linii, barw i znaków umownych pomniejszone obrazy rzeczywistości. Są to: mapy, atlasy i globusy. Obrazy te mogą dotyczyć różnych aspektów rzeczywistości ziemskiej lub przestrzeni kosmicznej. Por. B. Sordylowa: Informacja naukowa w Polsce. Problemy teoretyczne, źródła, organizacja. Wrocław 1987.

Międzynarodowa Federacja Stowarzyszeń Bibliotekarzy - IFLA (Division of Special Libraries. Section of Geography and Map Division) w ankiecie^{4/} rozesłanej do poszczególnych kolekcji kartograficznych z prośbą o podanie wielkości zbiorów kartograficznych wymienia następujące rodzaje materiałów gromadzonych w oddziałach kartograficznych bibliotek: mapy (drukowane i rękopiśmienne), plany, atlasy, globusy, modele reliefowe, fotografie przestrzenne i inne rodzaje obrazów oraz monografie i czasopisma. Te ostatnie stanowią raczej rodzaj materiałów pomocniczych ułatwiających interpretację materiałów kartograficznych^{5/}.

Materiały kartograficzne stanowią typ dokumentów znany w bibliotekarstwie i informacji naukowej jako dokument niepiśmienniczy (non-book materials)^{6/}, chociaż mogą również przybierać postać dokumentu piśmienniczego (book materials), jak widać to na przykładzie atlasów.

Istnieje nie tylko wielka różnorodność nośników informacji kartograficznej, jakimi są pergamin, papier, film, mikrofilm, mikrofisz, taśma magnetyczna, ekran monitora, lecz także ich produktów finalnych, takich jak rysunki, projekty, plany, mapy, fotografie przestrzenne, zdjęcia lotnicze i satelitarne oraz inne.

4/ Ankieta była rozsyłana w związku z przygotowaniem przez IFLA kolejnych wydań międzynarodowego informatora o zbiorach kartograficznych. Por. World Directory of Map Collection. Compil. by the Geography and Map Sub-section. Ed. by W. W. Ristow. München 1976 (IFLA Publications 8) oraz J.A. Wolter, R.E. Grim, D.K. Carrington: World Directory of Map Collections. München 1986 (IFLA Publications 31).

5/ Por. M. Piekuth: Pomocnicze materiały informacyjne w systemie dokumentacji kartograficznej. Polski Przegląd Kartograficzny 1981 T. 13 nr 2 s. 72-84.

6/ Por. H.P. Harrison: Non-book materials: a decade of development? Journal of Documentation 1979 Vol. 35 nr 3 s. 207-248.

Wymienione cechy, to jest różnorodność nośników i produktów końcowych, sprawiają, że materiały kartograficzne są gromadzone w zależności od konkretnych potrzeb przez różne instytucje, a mianowicie przez archiwa, biblioteki ogólne, biblioteki fachowe, ośrodki informacji oraz inne instytucje o charakterze państwowym, stanowiące swoistego rodzaju publiczne depozytoria. Przykładem takich instytucji mogą być urzędy geodezyjne i kartograficzne. Większość najstarszych materiałów kartograficznych w formie manuskryptów jest przechowywana w zależności od treści w archiwach lub w bibliotekach (zwykle w specjalnych oddziałach zwanych oddziałami kartograficznymi)^{7/}. W bibliotekach specjalnych oraz depozytoriach różnych instytucji państwowych i prywatnych są często gromadzone materiały kartograficzne o wysokim stopniu specjalizacji, co wynika z jakości i struktury zawartych w nich informacji i determinuje możliwości ich wyszukiwania i udostępniania^{8/}.

^{7/} Materiały kartograficzne stare stanowią osobny problem. Dostępność do nich zależy od stopnia opracowania kolekcji rękopisów i starych druków. Niektóre grupy tych materiałów są w ogóle do tej pory nie opracowane. Informacja o nich na szczeblu zbiorów specjalnych jest wybiórcza i niekompletna.

^{8/} Obecnie prowadzony jest i publikowany centralny katalog zbiorów kartograficznych w Polsce. Por. M. Łodyński: Centralny katalog zbiorów kartograficznych w Polsce. Z. 1-4, Warszawa 1961-1968 (obejmuje zbiory 33 największych bibliotek polskich) oraz Centralny katalog zbiorów kartograficznych w Polsce. Polska Akademia Nauk. Instytut Geografii i Przestrzennego Zagospodarowania. Z. 5 Cz.1-2. Warszawa 1982-1983 (Jest to kontynuacja centralnego katalogu M.Łodyńskiego). Publikowane są również katalogi zbiorów kartograficznych poszczególnych bibliotek, jak na przykład: K. Szukuła: Zbiory kartograficzne Biblioteki Uniwersyteckiej we Wrocławiu. Wrocław 1978.

Do podstawowych materiałów kartograficznych gromadzonych przez biblioteki naukowe w zależności od specjalizacji są zaliczane mapy, plany, atlasy, globusy, zdjęcia lotnicze i satelitarne^{9/}. Wymienione rodzaje materiałów można sprowadzić do trzech podstawowych kategorii: mapy (plany i atlasy)^{10/}, globusy i zdjęcia przestrzenne (lotnicze i satelitarne)^{11/}.

Wymienione rodzaje materiałów kartograficznych różnią się między sobą specyfiką sposobów odwzorowywania rzeczywistości w formie modeli przestrzennych, zawierających ilościową i jakościową charakterystykę obiektów lub zjawisk. Poszczególne rodzaje materiałów kartograficznych posiadają charakterystyczne cechy, które nazywać będziemy cechami dystynktywnymi.

^{9/} Według B. Kmiecikowej "Zbiory kartograficzne obejmują, obok globusów, nie tylko mapy i atlasy geograficzne, których temat stanowi fizyczne wyobrażenie nieba lub Ziemi i co do których przynależności nie mamy wątpliwości, ale również takie opracowania, w których założeniem autora było przedstawienie jakiegoś zagadnienia w ujęciu kartograficznym". Por. B. Kmiecikowa: Uwagi na temat sposobów opracowywania zbiorów kartograficznych. Roczniki Biblioteki Polskiej Akademii Nauk w Krakowie 1969 R. 15 s. 5.

^{10/} Materiały kartograficzne, które nie są mapami L. Ratajski nazywa wytworami mapopodobnymi. Według niego należą do nich globusy, panoramy, anaglifu i ortofotomapy. Por. L. Ratajski: Kartologia. Polski Przegląd Kartograficzny 1970 T. 2 nr 3 s. 97-110.

^{11/} Por. W. Babik: Języki informacyjno-wyszukiwawcze stosowane do opisu zbiorów kartograficznych - map. W: Materiały konferencyjne II Międzynarodowego Sympozjum nt. "Zastosowanie mikrokomputerów w inżynierii", Katowice 4-6 październik 1989, Katowice 1989, s. 126-137.

Mapy

Obecnie około 95% materiałów kartograficznych gromadzonych przez biblioteki stanowią mapy^{12/}. Można nawet zaryzykować stwierdzenie, że podstawowym materiałem kartograficznym są właśnie mapy.

W literaturze przedmiotu można znaleźć wiele różnych definicji mapy^{13/}. Najogólniej mapami nazywa się matematycznie uogólnione przedstawienia na płaszczyźnie obrazowo-znakowej powierzchni Ziemi

12/ Mapa była znana już w starożytności. Grecy nazywali ją "pinaks", tj. tablica, lub "geografija" - opis Ziemi, Rzymianie - "orbis pictus tabula", od której to nazwy pochodzi prawdopodobnie niemieckie "Landtafel" używane do XVI wieku. Nazwa "mapa", od łac. "mappa" (serweta) lub "mappa mundi" (płótno z obrazem świata), która pojawiła się w średniowieczu około IX wieku, oznaczała początkowo mapy malowane na tkaninach, w przeciwieństwie do karty (greckie - "chartes", łac. "charta" = papier), którą to nazwą określono mapy żeglarskie, rysowane na pergaminie. Stąd też pochodzą nazwy mapy: carte, carta etc. w wielu językach. Warto dodać, że w języku angielskim do dzisiaj stosuje się rozróżnienie map właściwych (lądowych), które określa się jako "maps", i map żeglarskich, nazywanych "charts". W Polsce nazwa mapy w dzisiejszym znaczeniu pojawiła się dopiero w drugiej połowie XVIII wieku (I. Krasicki). Dawniej były u nas używane nazwy: "wypisanie", "opisanie" wywodzące się prawdopodobnie od średniowiecznej nazwy mapy "pictura", dalej "malowanie krajów" (St. Łaski, około 1545 rok). Warto również przytoczyć określenie mapy sprzed 150 lat według słownika S.B. Lindego (1814 rok): "Mapa, mapka, mapeczka, wyobrażenie królestw, prowincji etc., z wymiarem ich obszerności na kartach szyćchowania. Rysowanie map na tym zawisło, żeby na papierze ustanowić punkta, któreby tak były między sobą położone, jak są położone na polu przedmioty, mające być oznaczone. Sam termin "mapa" pierwotnie oznaczał zmniejszone uogólnione przedstawienie na płaszczyźnie całości lub fragmentu powierzchni Ziemi. Z biegiem czasu termin ten nabrał dodatkowego znaczenia. Współczesne mapy nie ograniczają się do przedstawienia powierzchni Ziemi (zdjęcia), lecz przedstawiają najróżniejsze charakterystyki zjawisk przyrodniczych, społeczno-gospodarczych (na przykład temperaturę, ciśnienie, skład narodowościowy itp.), jednak zawsze w odniesieniu do określonej powierzchni. Por. J. Szafłarski: Zarys kartografii. Warszawa 1965.

13/ Por. J. Gołaski: Zagadnienia ogólnej definicji mapy w świetle dotychczasowych sformułowań. Przegląd Geodezyjny 1967 R. 39 nr 8 s. 319-322.

lub innej planety (lub ciała kosmicznego)^{14/}, ukazujące rozmieszczenie, stan i powiązanie różnych obiektów i zjawisk (przyrodniczych, społeczno-gospodarczych, kulturowych i innych) wybranych i scharakteryzowanych stosownie do przeznaczenia danej mapy^{15/}.

Określenie mapy jako "graficznego przedstawienia powierzchni jakiejś planety" może wydawać się czymś oczywistym, podanie jednak precyzyjnego określenia tak powszechnego i pozornie łatwego do określenia materiału bibliotecznego nie jest zadaniem łatwym. Znacznie łatwiejsze wydaje się wymienienie charakterystycznych (istotnych) cech mapy. Mapa jest bowiem obrazem całości albo wybra-

^{14/} W dawnej topografii Ziemi (i innych planet) przywiązywano szczególną uwagę do przedstawiania różnych własności powierzchni Ziemi, co miało umożliwić dokładny opis świata i zaspokoić potrzeby informacyjne ówczesnych użytkowników tych materiałów. Dlatego mapy i inne materiały kartograficzne najczęściej wykorzystywano - co ma miejsce także i dzisiaj - jako metodę przedstawiania różnych danych i zjawisk w ich aspekcie geograficznym. Wraz ze wzrostem znaczenia kartografii tematycznej akcent przesunął się z ujęć topograficznych, kładących nacisk na lokalizację informacji, na ujęcia tematyczne (przedmiotowe). Obecnie lokalizacja danych nie stanowi już najważniejszego zadania materiałów kartograficznych. Współcześnie opracowane materiały kartograficzne umożliwiają dokonanie różnego rodzaju porównań danych i zjawisk także w ujęciu strukturalnym i ich wzajemną korelację. Przykładem tego kierunku rozwoju wykorzystywania materiałów kartograficznych są mapy i atlasy tematyczne oraz atlasy narodowe i regionalne. Materiały kartograficzne są wykorzystywane na każdym etapie badań naukowych, a także w procesie gromadzenia informacji, ponieważ kartografia umożliwia przedstawienie w sposób uporządkowany przestrzennie i strukturalnie dużej ilości danych i oferuje w tym celu różne sposoby i narzędzia.

^{15/} E. Reński określa mapę jako "rysunek na płaszczyźnie będący pomniejszonym obrazem powierzchni Ziemi, wykonanym z zachowaniem określonych warunków matematycznych, uwzględniających jej krzywiznę i przedstawiający znakami umownymi szczegóły terenowe lub na tym tle zjawiska przyrody i życia społeczno-gospodarczego, dobrane i scharakteryzowane zgodnie z przeznaczeniem mapy". Por. E. Reński: Charakterystyka map oraz ich wykorzystania do planów zagospodarowania przestrzennego. Warszawa 1963.

nego fragmentu przestrzeni (trójwymiarowym) albo płaszczyzny (dwuwymiarowym).

Zdaniem wybitnego kartografa radzieckiego K.A. Saliszczewa są to trzy cechy:^{16/}

1. matematycznie określona konstrukcja;
2. zastosowanie specjalnych systemów znaków (symboli) kartograficznych;
3. wybór i uogólnianie przedstawianych treści.

Określona matematycznie konstrukcja map podaje z góry ścisłą zależność funkcyjną pomiędzy współrzędnymi punktów powierzchni Ziemi a współrzędnymi tychże punktów na płaszczyźnie. Umożliwia to przedstawianie za pomocą map stosunków przestrzennych i kształtów wybranych elementów rzeczywistości. Przejście od fizycznej powierzchni Ziemi do jej przedstawienia na płaszczyźnie jest możliwe dzięki odwzorowaniu kartograficznemu, które pozwala uwzględnić i określić zniekształcenia odwzorowania obiektów na płaszczyźnie ziemi. Stosowanie odwzorowań kartograficznych pozwala na czerpanie z map danych o rzeczywistym położeniu, rozmiarach i kształtach przedstawianych obiektów. Ów proces wyboru i uogólnienia przedstawianych zjawisk i obiektów nazywany jest w kartografii "generalizacją kartograficzną"^{17/}.

W celu zrozumienia istoty i roli map w procesie odzwierciedlenia rzeczywistości można rozpatrywać mapy jako modele, to jest konstrukcje odtwarzające w sposób uproszczony i schematyczny

^{16/} Por. K.A. Saliszczew: Kartografia ogólna. Warszawa 1984, s. 26.

^{17/} Badania informatyków nad ogólnymi problemami gromadzenia, przechowywania i udostępniania informacji pozwalają określić mapy jako specyficzną formę prezentacji, przetwarzania i analizy informacji. Tu niewątpliwą zaletą mapy jest zdolność jednoczącego i całościowego ujęcia przedstawień kartograficznych.

wybrane fragmenty lub aspekty rzeczywistości^{18/}. Materiały kartograficzne - w tym i mapy - jako przedstawienia kartograficzne posiadają podstawowe cechy modeli w ogóle: abstrahowanie od całości w celu badania części (konkretnego obszaru, konkretnych zjawisk i procesów); uproszczenie polegające na rezygnacji z uwzględnienia wielu charakterystyk i związków, a zachowaniu niektórych najistotniejszych; uogólnienie mające na celu wydzielenie ogólnych cech i właściwości.

Omawiając mapy należy również wspomnieć o planiglobach i planisferach, używanych przede wszystkim w celach dydaktycznych^{19/}. Planigloby to mapy świata dwóch półkul, zaś planisfery są mapami półkuli Ziemi lub nieba wykonanymi w odpowiednich odwzorowaniach.

Za materiały mapopodobne uważa się plany, którymi są zmniejszone obrazy przedmiotów usytuowanych na niewielkiej części powierzchni, wyrysowane na płaszczyźnie w skali 1:50, 1:100 lub mniejszej^{20/}.

Zbiory map lub planów nazywane są atlasami.

^{18/} Jest to model obrazowo-znakowy i pojęciowy. O mapach jako modelach traktuje bardzo obszernie artykuł angielskiego geografa C. Boarda: *Maps as models*, zamieszczony w pracy zbiorowej "Models in geography". London 1967. Obszerne streszczenie tego artykułu w języku polskim ukazało się w Przeglądzie Zagranicznej Literatury Geograficznej 1969. Z. 3/4. Por. także: A. Czerny: *Struktura i własności mapy jako modelu rzeczywistości*. Praca doktorska. Warszawa 1989 (mps).

^{19/} Idea wspomnianych tu planiglobów i planisfer wywodzi się z obrazów świata lub Ziemi z okresu, kiedy to całą wiedzę o świecie lub Ziemi można było zamknąć w jednym kole.

^{20/} Historycznie nazwa "plan" wiąże się z rozporządzeniem carskim wydanym po upadku Księstwa Warszawskiego, zakazującym polskim mierniczym wykonywania map z rysunkiem warstwic lub innych oznaczeń wysokości i rzeźby terenu.

Inne rodzaje materiałów kartograficznych

W ostatnim czasie nastąpił ogromny wzrost zapotrzebowania na bardzo szczegółowe informacje i dane kartograficzne. Mogą one być zdobywane i dostarczane użytkownikowi między innymi dzięki wykorzystaniu najnowszych technologii informacji^{21/}, dzięki którym informacja może zostać zapamiętana przez odpowiednie medium w oryginalnej formie. Użytkownik, korzystając z tych technologii, będzie miał możliwość selekcji i wyboru odpowiednich informacji i danych po to, aby - stosownie do zapotrzebowania - móc dokonywać ich analizy, korelacji i interpretacji^{22/}.

Wykorzystywanie fotografii przestrzennej przy sporządzaniu map zaowocowało w postaci różnego rodzaju zdjęć przestrzennych, nazywanych, w zależności od urządzenia, z którego je wykonano, zdjęciami lotniczymi lub zdjęciami satelitarnymi^{23/}.

Zdjęcia przestrzenne uznano za materiał kartograficzny sensu stricto, umożliwiający opracowywanie tak zwanych fotomap,

21/ Ten nowy sposób udostępniania informacji kartograficznej staje się możliwy między innymi dzięki osiągnięciom technicznym w dziedzinie gromadzenia, przetwarzania, przechowywania i udostępniania informacji. Za produkty nowych technologii w interesującej nas dziedzinie można uznać między innymi różnego rodzaju materiały kartograficzne w formie filmów, slajdów, obrazów (lub wydruków) komputerowych.

22/ Tego typu zapotrzebowania użytkowników mogą być w pełni zaspokojone tylko przez systemy faktograficzne lub ekspertowe. Ich omówienie mogłoby być przedmiotem odrębnej pracy.

23/ Technika ta po II wojnie światowej wywarła ogromny wpływ na sposób opracowywania map, zwłaszcza map topograficznych. Fotografia przestrzenna czyni bowiem możliwym utrwalenie i poszerzenie obrazu terytorium bez konieczności kompilowania go z wielu fragmentów. Interpretacja zdjęć przestrzennych dała asumpt między innymi do nowego sposobu opracowywania map tematycznych.

które w zasadzie są komplementarne do map topograficznych dla tego samego obszaru stanowiąc ich mutacje^{24/}.

Wśród materiałów kartograficznych gromadzonych przez biblioteki wymieniane są również globusy, kartogramy, reliefy i segmenty globusowe.

Globusy to trójwymiarowe modele Ziemi lub innych planet^{25/}. Pojawiły się one po raz pierwszy w okresie Odrodzenia^{26/} i od tej pory są gromadzone przez biblioteki wraz z innymi materiałami kartograficznymi^{27/}.

Obecnie gromadzone przez biblioteki globusy służą przeważnie do celów dydaktycznych (nauka geografii i astronomii). Oprócz globusów Ziemi spotyka się także globusy Księżyca i innych ciał niebieskich. Najczęściej są one w skali 1:40 000 000. Za pewną odmianę globusów można uznać globusy niebą, na których w rzucie centralnym na kulę jest przedstawiany obraz firmamentu z uwidoczonym obrazem południków i równoleżników oraz zaznaczeniem zbiorów gwiazd według ich wielkości. Analogicznie jak mapy, globusy mogą być ogólnogeograficzne i tematyczne. Ich skale są zwykle bardzo małe (od 1:30 000 000 do 1:89 000 000). Biblioteki gromadzą też globusy plastyczne i globusy odtwarzające rzeźbę powierzchni Ziemi.

24/ Ogromne zbiory zdjęć przestrzennych znajdują się w Narodowym Instytucie Geograficznym w Paryżu. W Polsce materiały tego typu są gromadzone przez Bibliotekę Politechniki Warszawskiej.

25/ Globusy (od łac. globus - kula) stanowią kuliste modele Ziemi lub kosmosu.

26/ Pierwsze globusy przedstawiały raczej idee kulistości Ziemi, a nie dokładną lokalizację lądów. Były na nich wyimaginowane obszary antypodów i błędne kontury lądów.

27/ Przyczyną gromadzenia globusów przez ówczesne biblioteki była snobistyczna moda posiadania bardzo wielkich globusów w bibliotekach dworów królewskich.

Gromadzone przez biblioteki kartogramy w sensie ścisłym nie są mapami. Mapa stanowi tylko ich podkład, na której tle są umieszczane odpowiednie wykresy. Przy ich pomocy przedstawia się w sposób graficzny dane statystyczne dotyczące poszczególnych obiektów, zjawisk lub regionów. Podkładem kartogramu jest zwykle mapa ogólnogeograficzna, a więc mapa z ubogą treścią. Służy ona do najogólniejszej orientacji i do umieszczania na niej określonych danych statystycznych. Dane statystyczne są przedstawiane w formie odpowiednich wykresów^{28/}.

Wykaz cech dystynktywnych poszczególnych rodzajów materiałów kartograficznych zawiera Załącznik nr 1. Są to cechy zarówno tekstów materiałów kartograficznych, jak i ich charakterystyk jako dokumentów. Mają one decydujący wpływ na relewancję w procesie wyszukiwania tych materiałów, dlatego powinny być uwzględniane w informacji pochodnej o formie i treści materiałów kartograficznych.

CECHY DYSTYNKTYWNE MATERIAŁÓW KARTOGRAFICZNYCH

Materiały kartograficzne jako dokumenty posiadają szereg cech formalno-wydawniczych, które mogą być relewante w procesie wyszukiwania. Są to przede wszystkim cechy wyznaczające formę oraz kompozycję materiałów kartograficznych. Cechy te związane są z procesem opracowania redakcyjnego materiałów kartograficz-

^{28/} Do najczęściej stosowanych rodzajów wykresów należą m.in.

- wykres słupkowy, służący do porównywania danych statystycznych z różnych przedziałów czasowych;
- wykres powierzchniowy;
- wykres sektorowy;
- wykres obrazkowy;
- zwykły wykres liniowy;
- wykres wektorowy.

nych, to jest organizacją i sposobem przedstawienia treści, techniką wykonania, rodzajem materiału, jego kształtem i wielkością^{29/}. Dotyczy to także odwzorowania kartograficznego, siatki kartograficznej, orientacji oraz elementów opisu bibliograficznego. Elementy te wprawdzie są wykładnikami relewancji formalnej, mają jednak wpływ na relewancję semantyczną^{30/}, mogą bowiem określać nie tylko sposób odwzorowywania, lecz także zakres odwzorowywanej wiedzy o rzeczywistości oraz stopień redukcji

29/ Założenia redakcyjne dokładnie precyzują przeznaczenie mapy (szkolna, turystyczna, samochodowa itp.), jej charakter (ścienne, podręczna, atlasowa itp.) oraz skalę.

Przy opracowywaniu założeń brane są pod uwagę:

- tytuł materiału;
 - skala wydania (z podaniem skali roboczej);
 - rodzaj, przeznaczenie mapy;
 - odwzorowanie;
 - zasięg mapy;
 - sieć wodna;
 - rzeźba terenu z podaniem gęstości cięcia i sposobu generalizacji;
 - osadnictwo z podaniem kryterium klasyfikacji i wyboru;
 - koleje z podaniem klasyfikacji;
 - drogi z podaniem klasyfikacji;
 - granice polityczne i administracyjne;
 - lasy, parki, rezerваты przyrody itp.;
 - nazewnictwo;
 - treść zasadnicza dla map tematycznych;
 - sposób opracowania pierwowysu (podkład), czystorys.
- (Por. Z. Brunner, J. Kosmowski, J. Lopatto, A. Meljon: Organizacja produkcji wydawniczej PPWK. Polski Przegląd Kartograficzny 1976 T. 8 Nr 3 s.113-122).

30/ Uzyskanie spójności organizacyjnej pola semantycznego na osi paradygmatycznej języka umożliwia precyzja odwzorowania kryteriów relewancji. Przyczyną wprowadzenia podziału relewancji na semantyczną i formalną była potrzeba rozróżnienia elementów relewantnych dotyczących treści materiałów kartograficznych i elementów odnoszących się do formy materiałów kartograficznych. W przypadku treści dokumentów wyróżnia się relewancję przedmiotową i relewancję interpretacyjną dotyczącą sposobu ujęcia przedmiotu, które to relewancje są odwzorowywane na osi syntagmatycznej języka i wpływają na zdolność selekcji informacji.

informacji, a także rodzaj i wielkość zniekształceń powstałych przy redakcyjnym opracowywaniu materiałów kartograficznych^{31/}.

Skala

W opisie materiałów kartograficznych za cechę dystynktywną może być również uznana skala. Określa ona stopień zmniejszenia długości przy przejściu od rzeczywistości do jej obrazu. Skala wyraża stosunek długości linii na materiale kartograficznym do długości odpowiedniej linii w terenie, dokładniej - do długości poziomego rzutu linii na powierzchnię elipsoidy.

Wielkość skali jest stała tylko na planie, czyli na wielkoskalowym obrazie niedużego fragmentu powierzchni Ziemi, przy którym można nie uwzględniać jej krzywizny. W różnych punktach przedstawienia kartograficznego skala przyjmuje różne wartości i zmienia się (z wyjątkiem odwzorowań wiernokątnych) w zależności od kierunku. Na mapach, podobnie jak i na planach, odwzorowywana jest tylko jedna skala, zwana skalą główną, odpowiadająca skali modelu elipsoidy ziemskiej odtwarzanej na płaszczyźnie. Odwzorowywanie tylko jednej skali głównej jest charakterystyczne przede wszystkim dla map topograficznych, praktycznie nie odzwierciedlających dostrzegalnych zniekształceń, oraz dla map małoskalowych, na których zniekształcenia długości są małe lub niezauważalne.

^{31/} Opracowanie materiałów kartograficznych odbywa się według następującego algorytmu. Dane są zwykle: obszar, temat, przeznaczenie. Pytamy o: (to znaczy mamy wybrać lub ustalić) skalę, mapę podstawową, tematyczne podziały przedmiotowe, powierzchnię odniesienia, jednostki miary lub ilości, grupowania i gradacje ilościowe, interwały i relacje, formę, rozmiary i kolorystykę. Do rozważenia pozostają jeszcze: odwzorowanie, treść, generalizacja, normalizacja, opis treściowy, językowy i graficzny, kształtowanie tytułów, legenda. Por. też: F. Piątkowski: Kartografia. Redakcja map i reprodukcja kartograficzna. Warszawa 1969.

Na mapach obejmujących większe obszary występują znaczne odchylenia skal miejscowych od skali głównej, dlatego zaznacza się punkt lub linię siatki kartograficznej zachowującą skalę główną.

W różnych krajach są stosowane różne skale. W krajach, w których przyjęto metryczny system miar, stosowane są przeważnie następujące skale: 1:5 000, 1:10 000, 1:25 000, 1:50 000, 1:100 000, 1:200 000, 1:300 000, 1:500 000, 1:1 000 000. W celu ułatwienia pracy na mapach często umieszczane są jednocześnie skale: liczbowa, mianowana i liniowa, zwana podziałką.

Oprócz skali długości na materiałach kartograficznych niekiedy bywa stosowana skala powierzchni określająca stopień zmniejszenia powierzchni płaszczyzny przy przejściu od powierzchni elipsoidy do mapy. Wymienione rodzaje skal są zaliczane do skal poziomych. Oprócz tego są stosowane skale pionowe określające stopień zmniejszenia wysokości, co jest ważne w przypadku map plastycznych i reliefów.

Odwzorowanie kartograficzne

Odwzorowanie kartograficzne jest to określony matematycznie sposób odzwierciedlania powierzchni elipsoidy na płaszczyźnie. Ustala on analityczną zależność pomiędzy współrzędnymi geograficznymi (lub innymi) punktów elipsoidy ziemskiej a współrzędnymi prostokątnymi (lub innymi) tychże punktów na płaszczyźnie^{32/}.

Podstawową cechą każdego odwzorowania kartograficznego jest zniekształcenie wyrażające się zmiennością skali. Odwzorowania różnią się między sobą charakterem zniekształceń, ich wielkość-

^{32/} Por. K.A. Saliszczew: Kartografia ogólna. Warszawa 1984, s. 28.

cją i rozkładem. W każdym odwzorowaniu istnieją pewne punkty lub linie (lub układy linii), w których zachowana jest skala główna. Takie punkty i linie są nazywane punktami i liniami zerowych zniekształceń. Wielkość zniekształceń rośnie w miarę oddalania się od punktów i linii zerowych zniekształceń. Innymi słowy, rosną one wraz ze zwiększaniem się przedstawianej powierzchni.

O wyborze odwzorowania dla konkretnej mapy decyduje przeznaczenie mapy lub innego materiału kartograficznego, potrzeby użytkowników oraz geograficzne własności przedstawianego obszaru.

Mając na względzie przeznaczenie konkretnego materiału kartograficznego ustala się dopuszczalny charakter zniekształceń. Uwzględnienie czynników geograficznych, to jest wielkości, kształtu i położenia danego obszaru pozwala na wybór odpowiedniego odwzorowania, to znaczy takiego, które odznacza się najmniejszymi zniekształceniami lub ich korzystnym rozkładem albo też innymi właściwościami cennymi dla postaci redagowanej mapy. Dlatego na przykład dla map świata stosuje się przede wszystkim odwzorowania walcowe i pseudowalcowe, z siatkami o prostych i równoległych do siebie równoleżnikach, co jest bardzo cenne przy badaniu zjawisk o równoleżnikowej strefowości. W wypadku map kontynentów (Europy, Azji, Ameryki Północnej i Południowej, Australii z Oceanią) stosuje się przeważnie odwzorowania azymutalne ukośne wiernopolowe Lamberta, z punktem zerowym zniekształceń w centrum przedstawianego kontynentu. W przypadku Afryki zamiast ukośnego stosuje się odwzorowanie równikowe.

Przy wyborze odwzorowania dużą rolę odgrywa również czynnik matematyczny, jakim jest wielkość zniekształceń, lecz nie zawsze jest to czynnik decydujący. W wielu przypadkach ważniejsze jest

uwzględnienie wymagań geograficznych, takich jak całościowe przedstawienie wzajemnie powiązanych obiektów (na przykład Ocean Atlantycki wspólnie z Morzem Arktycznym) oraz najkorzystniejsze pokazanie obszarów istotnych dla konkretnego przedstawienia kartograficznego.

Siatka współrzędnych

Ważnym elementem materiałów kartograficznych jest również sieć linii współrzędnych - płaski obraz sieci odpowiednich linii na elipsoidzie ziemskiej. W procesie wykonywania map i innych materiałów kartograficznych taka siatka odgrywa rolę szkieletu przy konstruowaniu obrazu kartograficznego, natomiast przy posługiwaniu się konkretnym materiałem pozwala określać współrzędne punktów elipsoidy ziemskiej, nanosić na odpowiedni materiał kartograficzny punkty o określonych współrzędnych, mierzyć kierunki linii w stosunku do stron świata, w dowolnym miejscu materiału wyliczać skalę i zniekształcenia^{33/}.

Znaczenie siatki kartograficznej jest uwarunkowane przydatnością geograficzną południków i równoleżników^{34/}. Południki odpowiadają kierunkowi północ-południe, równoleżniki zaś kierunkowi wschód-zachód. Kierunki te po wyznaczeniu ich w terenie służą do orientowania mapy w czasie prac polowych. Ta właściwość siatki kartograficznej jest szczególnie istotna w wypadku map topograficznych.

^{33/} Wielkość zniekształceń w dowolnym punkcie mapy można określić na podstawie siatki kartograficznej, nawet jeśli nie jest znany rodzaj odwzorowania.

^{34/} Siatki kartograficzne stanowiące sieć równoleżników i południków należą do najczęściej stosowanych siatek, które na mapach małoskalowych występują jako jedyne.

Na mapach małoskalowych siatka kartograficzna jest podstawowym elementem orientacji geograficznej i podstawą różnorodnych uogólnień i wniosków wynikających z równoleżnikowej strefowości występowania wielu zjawisk przyrodniczych. Ponadto różnice długości geograficznej wyrażają różnicę czasu.

U podstaw każdego obrazu kartograficznego leży określone odwzorowanie. Siatka kartograficzna, charakteryzująca je matematycznie, czasem nie jest na mapie przedstawiona. Pozbawione siatek bywają mapy schematyczne, mapy niewielkich obszarów o nieznacznych zniekształceniach, mapy nie przeznaczone do prac kartometrycznych oraz obrazowe przedstawienia powierzchni ziemskiej.

Orientacja mapy

Orientacją mapy nazywa się określenie stron świata w stosunku do ramek mapy. Metody ustalania i zaznaczania kierunków na mapie były różne^{35/}. Na starszych mapach jest wytyczona północ przez wskazówkę busoli umieszczoną na rózży wiatrów, odpowiedni opis poza ramkami lub strzałkę lub przez prostopadłe skrzyżowanie dwóch linii. Współczesne mapy mają orientację północną, która umownie wypada na górnej ramce arkusza^{36/}.

^{35/} W średniowieczu większość map świata, tak zwanych "mappae mundi" (XII-XV w.), była zorientowana na wschód, to znaczy wschód znajdował się u góry mapy. Taką też orientację posiada wiele map starożytnych.

^{36/} Zasady orientowania map ciągle się zmieniały. Na wielu mapach średniowiecznych wschód znajdował się u góry mapy. Stąd też termin "orientacja" (od słowa łacińskiego "oriens" = wschód). Orientację północną wprowadzono w XIV wieku na portolanach, a w XVI wieku została ogólnie przyjęta w praktyce.

Podobnie jak skale orientacja nie miała stałego miejsca na arkuszu. Umieszczano ją w dowolnym miejscu w świetle ramek lub poza nimi.

Elementy opisu bibliograficznego

Określają one zakres informacji identyfikujących materiały kartograficzne. Opis ten - stosowany od wieków w bibliografiach i sporządzany według utartych wzorów - nie został w Polsce do tej pory w pełni znormalizowany^{37/}. Podstawowe elementy opisu materiałów kartograficznych zawarto w PN-78/N-2345. Opis bibliograficzny^{38/}. Istnieje natomiast międzynarodowa norma opisu bibliograficznego materiałów kartograficznych, opracowana przez IFLA w 1977 roku^{39/}. Wymienione normy autor artykułu poddał dokładnej analizie pod kątem zawartości informacji o materiałach kartograficznych. Okazało się, że dane zawarte w tytułaturze charakteryzują nie tylko materiały kartograficzne jako dokumenty, lecz także zawierają informacje o tekstach tych dokumentów.

37/ Wprowadzie elementy opisu materiałów kartograficznych podaje norma PN-78/N-2345, ale od lat trwają prace nad przystosowaniem normy międzynarodowej, opracowanej przez IFLA, do warunków polskich. Prace te są prowadzone przez Bibliotekę Narodową w Warszawie i polegają na adaptacji opisu międzynarodowej ISBD (CM). W ramach polskiej normy opisu bibliograficznego, która to norma jest normą arkuszową, zaplanowano arkusz 05. Dokumenty kartograficzne.

38/ W PN-73/N-01152. Opis bibliograficzny, na s.19 wymieniono następujące elementy opisu bibliograficznego materiałów kartograficznych (przy pomocy [*] oznaczono niezbędne elementy opisu): nazwa kartografa [*], tytuł [*] wraz z określeniem rodzaju materiału, oznaczenie wydania [*], miejsce wydania [*], nazwa instytucji wydawniczej, rok wydania [*], liczba map lub arkuszy, format, nazwa instytucji sprawczej, nazwa i numer serii.

39/ Norma ta nosi akronim ISBD(CM), który pochodzi od: International Standard Bibliographic Description for Cartographic Materials. ISBD(CM). Recommended by the Joint Working Group on the International Standard Bibliographic Description for Cartographic Materials set up by the IFLA Committee on Cataloguing and the IFLA Subsection of Geography and Map Libraries. London IFLA International Office for UBC. London 1981.

Cechy tekstów omawianych dokumentów odzwierciedlone w tytule stają się elementami identyfikującymi materiały kartograficzne. Stąd szczególne zainteresowanie autora bibliograficznymi charakterystykami materiałów kartograficznych.

Opis bibliograficzny materiałów kartograficznych może ujawniać te elementy informacji o tekstach, które są obecne w ich tytułach, co jest zależne od podstaw normalizacyjnych przyjmowanych przy sporządzaniu opisów bibliograficznych. W przypadku ISBD(CM) elementy te są zależne od decyzji indeksatora, do którego należy decyzja o wyborze stopnia szczegółowości opisu, zależnie od przeznaczenia opisu, to jest m.in. od potrzeb jego potencjalnych użytkowników. Zakres, rodzaj i formy zapisu danych bibliograficznych o tekstach materiałów kartograficznych przedstawiono w Załącznikach nr 2 i 3, w których połączono elementy opisu z określeniem, w jakiej mierze charakteryzują one teksty materiałów kartograficznych i same materiały. Na podstawie tych tabel możemy stwierdzić, że w opisach bibliograficznych materiałów kartograficznych zawarte są *implicite* nie tylko cechy identyfikacyjne samych dokumentów, lecz także zawarte są *explicite* informacje o tekstach tych materiałów. Obecność tych danych w tytułaturze materiałów kartograficznych oraz stopień szczegółowości opisu decydują o umieszczeniu tych informacji w głównej strefie opisu, z której przede wszystkim generuje się klucze wyszukiwawcze (tj. hasła porządkujące zbiory danych). Informacje znajdujące się poza tytułaturą materiałów kartograficznych często nie są uwzględniane przy doborze kluczy wyszukiwawczych.

Tytuł należy do najważniejszych elementów informacji o materiałach kartograficznych^{40/}. Składa się on z następujących elementów: nazwa geograficzna przedstawianego obszaru, temat, czasem przeznaczenie. Na starszych materiałach kartograficznych tytuł występuje razem z imieniem i nazwiskiem autora, datą powstania materiału, datą wydania i skalą. W tytułach materiałów kartograficznych często występują określenia rch rodzaju, na przykład w przypadku map: "ogólna", "przeglądowa", "zbiorcza".

Tytuł umieszczano w różnych miejscach map, zazwyczaj w obrębie ramek^{41/}. W celu wyeksponowania tytułu stosowano dawniej ozdobne ramki, zwane kartuszami^{42/}. Na nowszych mapach tytuł

40/ Tytuł jest umieszczany tylko na mapach i materiałach mapopodobnych.

41/ Spełniają one funkcję rozgraniczającą rysunek mapy od miejsca przeznaczonego na jej opis. Starsze mapy najczęściej nie posiadały ramek. Powierzchnię całego arkusza pokrywał rysunek kartograficzny. W miarę rozwoju kartografii (XV i XVI wiek) zaczęto stosować ramki, które stanowiły również element zdobniczy. Umieszczano na nich widoki miast, widoki ważniejszych budowli, sceny batalistyczne i rodzajowe, herby, inne ornamenty. W wiekach XVIII i XIX ramki stały się miejscem, w którym umieszczano współrzędne katastralne danego obszaru (mapy katastralne), albo współrzędne geograficzne lub astronomiczne, niekiedy również elementy siatki kartograficznej.

Obecnie stosuje się ramki, poza którymi umieszcza się opis informacyjny mapy. Ramkę mapy często tworzy kilka linii. Linie, które bezpośrednio ograniczają obraz kartograficzny, nazywane są ramką wewnętrzną. Często jest ona uzupełnianą ramką minutową (lub stopniową), zazwyczaj składającą się z dwu blisko położonych linii równoległych, na których zaznaczono wyjścia pośrednich południków i równoleżników, uzupełniające siatkę kartograficzną. Ramka minutowa pozwala określać na podstawie mapy współrzędne geograficzne prościej i dokładniej niż sama siatka kartograficzna. Często jest wykreślana ramka zewnętrzna, której zadaniem jest nadanie mapie estetycznego wyglądu i wykończenia. Niekiedy rezygnuje się zupełnie z wykreślania ramki.

42/ Odnosi się to jednak tylko do pierworysów, bowiem inne rodzaje materiałów kartograficznych (map) raczej nie posiadają ramek (szkice polowe, odręcznie odrysy).

jest umieszczony zwykle na dolnym marginesie arkusza lub w górnej jego części.

Wykonawcy

Pod tym terminem rozumie się wszystkie osoby, które brały udział w tworzeniu materiału kartograficznego, to jest autor-kartograf, sztycharz, wydawca, nakładca.

Miejsce występowania danych o wykonawcach było różne. Czasem umieszczano je obok tytułu. Jeżeli chodzi o obecnie produkowane mapy, to dane o autorach umieszcza się pod tytułem, poza dolną ramką lub w dolnej części arkusza.

Data i miejsce wydania

Elementy te są najczęściej umieszczane przy informacjach o tytule i wydawcach^{43/}. Jako elementy formalne określają one datę i miejsce publikacji (wydania), a czasem sytuacji na mapie. Ten ostatni element nie posiada charakteru formalno-wydawniczego, lecz wchodzi w zakres treści materiałów kartograficznych.

Kształt i format

Wymienione tu elementy stanowią podstawę formalnych podziałów materiałów kartograficznych. Kryterium kształtu pozwala dzielić je według ich wymiarowości na materiały kartograficzne dwuwymiarowe: mapy, zdjęcia lotnicze i satelitarne, oraz materiały kartograficzne trójwymiarowe: globusy, reliefy i mapy reliefowe.

^{43/} W dawnych mapach elementy te należą do najczęściej pomijanych. Ich ustalenie wymaga wiele czasu i wiedzy bibliotekarza-kartografa.

Same przedstawienia kartograficzne umieszczone na nośniku mogą przybierać różne kształty: wielobok, owal, trapez, prostokąt, koło (planigloby). Współczesne mapy mają najczęściej kształt prostokątów.

Kryterium formatu pozwala na opis materiałów kartograficznych w kategoriach wymiarów fizycznych (wielkości), podawanych w formie symbolu formatu lub w centymetrach. Formaty map współcześnie wydawanych są z reguły unormowane przepisami technicznymi, pozwalającymi wyróżnić formaty: A1, A2, A3.

Mapy o dużych wymiarach są sporządzane na wielu arkuszach. System podziału map na arkusze nazwano cięciem. Rozróżnia się podział prostokątny i podział trapezowy.

Kształt i wielkość warunkują sposób korzystania z materiałów kartograficznych. Dlatego wyróżnia się m.in. mapy wiszące, mapy podręczne, mapy stojące, atlasy.

Materiał (nośnik)

Materiały kartograficzne, a zwłaszcza mapy, wykonywano zwykle na najlepszym materiale, jakim dysponowano. Najpierw były to tabliczki gliniane lub woskowe, potem papierus. W nowszych czasach używano grubego papieru podklejonego płótnem lub samego płótna. Obecnie pierworwisy map powstają na białym kartonie, niekiedy naklejonym na aluminiową folię (blachę), tak zwaną planszę. Wzmocnienie papieru metalem ma zapewnić odporność na kurczenie lub rozciąganie się papieru pod wpływem warunków atmosferycznych. Papier jest coraz częściej zastępowany przez tworzywa sztuczne (folie), które charakteryzują się dużą stabilnością powierzchni, przezroczystością, nie reagują na wilgoć. Wykorzystuje się je najczęściej do wykonania rysów pomiarowych i matryc pierworysów.

Najnowsze techniki pomiaru - wywodzące się z fotogrametrii - umożliwiają przygotowywanie materiałów kartograficznych na papierze fotograficznym w postaci fotomap i fotoszkiców. Materiały kartograficzne można również spotkać na mikrofilmach i na ekranie monitora w komputerze.

Technika wykonania

Współcześnie opracowywane i wydawane mapy i inne materiały kartograficzne są wielobarwne. Do sporządzania map stosuje się już od XIX wieku fotografię. W wyniku zastosowania fotogrametrii uzyskuje się fotogramy, to jest fotografie powierzchni Ziemi, które są następnie uczytelniane, to znaczy następuje identyfikacja i oznaczanie za pomocą znaków kartograficznych obiektów zarejestrowanych przez obiektyw fotograficzny.

Do reprodukcji stosuje się często metodę światłokopii lub druku (litografię). Druk jako szybki sposób otrzymania odrysów jest znany od dawna. Najpierw wykorzystywano drzeworyt, następnie miedzioryt, a od XIX wieku litografię. Rysunek przeszedł ewolucję w kierunku precyzji i dokładności: od grubych linii drzeworytu aż po siateczkę cienkich linii miedziorytu. Wśród najnowszych map można spotkać coraz więcej materiałów sporządzonych za pomocą komputerów.

Wykaz wyróżnionych przez nas cech dystynktywnych materiałów kartograficznych zawiera Załącznik nr 4. Wymienione tam cechy mogą zostać uznane za elementy relewantne dopiero po skonfrontowaniu ich z potrzebami użytkowników materiałów kartograficznych. Potrzeby te ujawniają się w procesie wyszukiwania informacji w formie zapytań informacyjnych. Konfrontacja wyróżnionych cech

dystynktywnych z realnymi potrzebami użytkowników będzie przedmiotem kolejnego artykułu.

Podsumowanie

Proponowany przez nas zbiór cech dystynktywnych jest zbiorem cech opisujących tylko takie rodzaje materiałów, które tradycyjnie są zaliczane do materiałów kartograficznych. Przedstawione cechy dystynktywne poszczególnych rodzajów materiałów kartograficznych mogą stanowić elementy wyszukiwawcze w procesie wyszukiwania informacji kartograficznej. Analiza norm opisu bibliograficznego pokazała, że repertuar tych środków jest bardzo ubogi. Powoduje to niepotrzebne straty informacji na poziomie j/w, co sprawia, że użytkownicy zbiorów (kartograficznych) często nie otrzymują takich materiałów kartograficznych, jakie pragnęliby otrzymać, i w takiej formie, jaka byłaby pożądana ze względu na zaspokojenie ich potrzeb informacyjnych.

Przeprowadzona charakterystyka materiałów kartograficznych umożliwiła wyodrębnienie ich zasadniczych cech, którymi są: rodzaj materiału kartograficznego, terytorium, temat, data wydania, format, rodzaj siatki kartograficznej i rodzaj odwzorowania i cechy fizyczne.

Wyróżnione cechy dystynktywne materiałów kartograficznych uznane za elementy relewantne w procesie wyszukiwania mogą stanowić materiał m.in. do budowy klasyfikacji fasetowej materiałów kartograficznych. Już cechy samej informacji kartograficznej predestynują ją do tego, aby była ona porządkowana i wyszukiwana za pomocą klasyfikacji fasetowej. Opis materiałów kartograficznych jest niejako fasetowy z samej swej natury.

Terytorium zawsze stanowi podstawowe kryterium przy wyszukiwaniu materiałów kartograficznych i jest ono najczęściej wykorzystywane przy stawianiu pytań. Drugim w kolejności stosowania kryterium klasyfikacyjnym może być przedmiot. Kryterium to stwarza dalsze możliwości podziałów stosownie do tematyki. Użytkownicy materiałów kartograficznych interesują się w pierwszym rzędzie aspektem terytorialnym przedmiotu, natomiast sam przedmiot jest traktowany jako element drugorzędny.

Zarówno charakter samych tekstów materiałów kartograficznych, jak i kompleksowość i wieloaspektowość informacji przedstawianych na materiałach kartograficznych wskazuje na potrzebę przyjęcia wielowymiarowej struktury pola semantycznego języka informacyjno-wyszukiwawczego dla materiałów kartograficznych. Przyjęcie wielowymiarowości jest niezbędne ze względu na konieczność odzwierciedlenia przy indeksowaniu i wyszukiwaniu wielości i różnorodności wyróżnionych cech dystynktywnych tych materiałów.

Literatura

1. BABIK W. Języki informacyjno-wyszukiwawcze stosowane do opisu zbiorów kartograficznych - map. W: Materiały konferencyjne II Międzynarodowego Sympozjum nt. "Zastosowanie mikrokomputerów w inte". Katowice 4-6 październik 1989, Katowice 1989, s.126-137.
2. BABIK W. Wykorzystanie mikrokomputerów w Polsce do gromadzenia i wyszukiwania informacji o materiałach kartograficznych. Zagadnienia Informacji Naukowej 1990 nr 2(57) s. 165-174.
3. BABIK W., BZINKOWSKA J. The state and trends in the computerization of cartographic collections in Poland. Referat wygłoszony na 7th Conference of Liber Maplibrarians (Ligue des Bibliothèques Europeennes de Recherche. Group de Cartothecaires) Paris

- 24-28 Sept. 1990 (mps). [Złożono do druku w materiałach konferencyjnych].
4. BOARD C. Maps as models. W: Models in geography. London 1967 (Obszerne streszczenie tego artykułu w języku polskim ukazało się w Przeglądzie Zagranicznej Literatury Geograficznej 1969 Z. 3/4).
 5. BOROWIK O., HACZEWSKA Z. Klasyfikacja geograficzna. Warszawa 1954.
 6. BRUNNER Z., KOSMOWSKI J., LOPATTO J., MELJON A. Organizacja produkcji wydawniczej PPWK. Polski Przegląd Kartograficzny 1976 T. 8 Nr 3 s. 113-122.
 7. CHARRE J.P. Essai d'analyse des documents geo-iconographiques. Paper presented on IFLA General Conference, Paris 1989. W: [Materiały konferencyjne] IFLA General Conference, Paris 1989. Paris 1989 s. 49-61.
 8. CZERNY A. Struktura i własności mapy jako modelu rzeczywistości. Praca doktorska wykonana pod kierunkiem prof. dr.hab. Wiktora Grygorenki w Instytucie Geografii i Przestrzennego Zagospodarowania PAN w Warszawie. Warszawa 1989 (mps).
 9. DRABEK J., PIĄTKOWSKI F. 1000 słów o mapach i kartografii. Warszawa: Wydawnictwo Ministerstwa Obrony Narodowej 1989.
 10. ENCYKLOPEDIA wiedzy o książce. Red. nacz. A. Binkermajer, B. Kocowski, J. Trzynaśkowski. Wrocław: Zakład Narodowy im. Ossolińskich 1971.
 11. GOŁASKI J. Zagadnienia ogólnej definicji mapy w świetle dotychczasowych sformułowań. Przegląd Geodezyjny 1967 R.39 Nr 8 s.319-322.
 12. GRABOWSKA B. Centralna Informacja Kartograficzna. Polski Przegląd Kartograficzny 1983 T. 15 Nr 3 s. 133-134.
 13. GRABOWSKA-WARZECHA H. Komunikat o rejestracji map i atlasów w bibliografii narodowej, W: Materiały z pierwszej Ogólnopolskiej Konferencji Bibliotekarzy Kartografów. Warszawa 7-8 październik 1977. Oprac. B. Krassowski. Warszawa 1979, s. 55-61.

14. HARRISON H.P. Non-book materials: a decade of development? *Journal of Documentation* 1979 Vol. 35 Nr 3 s. 207-248.
15. INTERNATIONAL Standard Bibliographic Description for Cartographic Materials. ISBD(CM). Recommended by the Joint Working Group on the International Standard Bibliographic Description for Cartographic Materials set up by the IFLA Committee on Cataloguing and the IFLA Subsection of Geography and Map Libraries. London: IFLA International Office for UBC 1981 (draft).
16. KARTOGRAFIA. Opracowanie map. Podstawowe normy i określenia. PN-73/N-02260. Warszawa 1979.
17. KARTOGRAFIA. Opracowanie map. Terminologia (PN-87/N-02260). Warszawa 1987.
18. KMIĘCIKOWA B. Uwagi na temat sposobów opracowywania zbiorów kartograficznych. *Roczniki Biblioteki Polskiej Akademii Nauk w Krakowie* 1969 R. 15 s. 5-16.
19. ŁODYŃSKI M. Centralny katalog zbiorów kartograficznych w Polsce. Z. 1-4 Warszawa 1961-1968.
20. OPIS bibliograficzny (PN-73/N-01152). Warszawa 1973.
21. PIĄTKOWSKI F. Kartografia. Redakcja map i reprodukcja kartograficzna. Warszawa: PWN 1969.
22. PIEKUTH M. Pomocnicze materiały informacyjne w systemie dokumentacji kartograficznej. *Polski Przegląd Kartograficzny* 1981 T.13 Nr 2 s. 72-84.
23. POLARCZYKOWA A. Dokumenty pergaminowe, druki, ikonografia, kartografia w archiwach prywatnych - ich ewidencja i przechowywanie. *Archeion* 1973 T. 59 s. 15-30.
24. RATAJSKI L. Kartologia. *Polski Przegląd Kartograficzny* 1970 T. 2 Nr 3 s. 97-110.
25. REŃSKI E. Charakterystyka map oraz ich wykorzystania do planów zagospodarowania przestrzennego. Warszawa 1963 (Prace własne Komitetu Urbanistyki i Architektury, Nr 73).
26. ROBÓTKA H. Rozwój form kancelaryjnych i współczesne rodzaje dokumentów archiwalnych. Dokumentacja geodezyjno-kartograficzna. Toruń 1985.

27. SALISZCZEW K.A. Kartografia ogólna. Warszawa: PWN 1984.
28. SITARSKA A. Informacja o tekście w strukturze danych bibliograficznych i katalogowych. W: Od kodu do kodu. Prace ofiarowane Profesorowi Olgierdowi Adrianowi Wojtasiewiczowi na 70-lecie urodzin. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego 1987, s. 211-221.
29. SMITS J. Report on the "Inquiry into Map-Use and User-Habits in Europe". Paper presented on the 7th Conference of Librarians, Paris 24-28 Sept. 1990.
30. SORDYŁOWA B. Informacja naukowa w Polsce. Problemy teoretyczne, źródła, organizacja. Wrocław 1987.
31. SZAFŁARSKI J. Zarys kartografii. Warszawa: PPWK 1965.
32. WALLIS H. Maps as a medium on scientific communication. W: Monografie z Dziejów Nauki i Techniki, Warszawa 1987 T. 87 s. 251-262.
33. WOLTER J.A., GRIM R.E., CARRINGTON D.K. World Directory of Map Collections. Munchen-New York-Paris 1986 (IFLA Publications 31).
34. WORLD Directory of Map Collection. Compil. by the Geography and Map Sub-section. Ed. by W.W. Ristow. München-London-Paris 1976 (IFLA Publications 8).
35. YOUNGMANN C.E. A linguistic Approach to Map Description. In: First International Advanced Study Symposium on topological data structures for geographic information systems. Ed. G. Dutton, Vol. 7: Spatial Semantics: Understanding and Interacting with Map Data. Cambridge, Mass. 1978.

1.02.1992 r.

Rodzaj materiału kartograficznego	Cechy formalno-wytwórcze	Źródło informacji
Mapa, plan, atlas	Rodzaj materiału Forma wydawnicza Skala Sposób przedstawienia Format Autorstwo Tytuł Przeznaczenie Barwność Data wydania Język opisu Odzorowanie Układ współrzędnych Orientacja	Temat Terytorium Data sytuacji
Głobus	Rodzaj Skala Wielkość Wykonawca Przeznaczenie Data produkcji	Temat
Zdjęcie przestrzenne	Rodzaj Skala Technika wykonania Cechy fizyczne Wykonawca Data wykonania Data wywołania Entuzjatyka	Obszar/Terytorium Data sytuacji

Załącznik nr 2 Dane o tekstach w opisie bibliograficznym
 materiałów kartograficznych według
 PH-73/H-01152

Lp.	Nazwa elementu opisu bibliograficznego	Dane o tekście
1.	Nazwa kartografa	Nazwiska autorów tekstu
2.	Tytuł i dodatki do tytułu	Typ, rodzaj materiału Temat Nazwa obszaru Przeznaczenie
3.	Skala	Rodzaj skali Wielkość skali
4.	Nazwy współtwórców	Nazwiska współtwórców
5.	Oznaczenie wydania	Zródła opracowania
6.	Mniejsze wydania	
7.	Nazwa instytucji wydawniczej	
8.	Rok wydania	Data opublikowania (aktualność danych)
9.	Uwagi	
10.	Nazwa instytucji sprawczej	
11.	Nazwa serii oraz numer w jej obrębie	
12.	Dokumentacja	
13.	Mapy boczne	
14.	Wysokość nakładu	
15.	Cena	

Załącznik nr 2 Dane o tekstach w opisie bibliograficznym
 materiałów kartograficznych według
 PN-73/N-01152

Lp.	Nazwa elementu opisu bibliograficznego	Dane o tekście
1.	Nazwa kartografa	Nazwiska autorów tekstu
2.	Tytuł i dodatki do tytułu	Typ, rodzaj materiału Temat Nazwa obszaru Przeznaczenie
3.	Skala	Rodzaj skali Wielkość skali
4.	Nazwy współtwórców	Nazwiska współtwórców
5.	Oznaczenie wydania	Zródła opracowania
6.	Mniejsze wydania	
7.	Nazwa instytucji wydawniczej	
8.	Rok wydania	Data opublikowania (aktualność danych)
9.	Uwagi	
10.	Nazwa instytucji sprawczej	
11.	Nazwa serii oraz numer w jej obrębie	
12.	Dokumentacja	
13.	Mapy boczne	
14.	Wysokość nakładu	
15.	Cena	

Załącznik nr 3 : Dane o tekstach w opisie bibliograficznym
materiałów kartograficznych według ISDD (CH).

l.p.	Strefa	Nazwa elementu opisu	Dane o tekście
1.	Tytuł i strefa odpowiedzialności autorskiej	1.1. Tytuł właściwy 1.2. Przeznaczenie materiałów kartogr. (opcjonalnie) 1.3. Tytuł równoległy 1.4. Dodatki do tytułu 1.5. Odpowiedzialność autorska	Rodzaj (typ) materiału kartograficznego Temat Przeznaczenie Język opisu Nazwiska twórców
2.	Strefa wydania	2.1. Oznaczenie wydania 2.2. Oznaczenie odpowiedzialności w stosunku do wydania	Wariant publikowanego tekstu Fundacja publikacji
3.	Strefa danych matematycznych	3.1. Skala 3.2. Projekcja (opcjonalnie) 3.3. Współrzędne (opcjonalnie)	Rodzaj skali Wielkość skali
4.	Strefa publikacji i dystrybucji	4.1. Miejsce publikacji 4.2. Nazwa wydawcy 4.3. Funkcja dystrybutora 4.4. Data wydania 4.5. Miejsce druku 4.6. Nazwa drukarni 4.7. Data druku	
5.	Strefa opisu fizycznego	5.1. Przeznaczenie 5.2. Inne cechy fiz. 5.3. Wymiary 5.4. Dodatki	Przeznaczenie
6.	Strefa serii		
7.	Strefa uwag		Język legendy
8.	Strefa numeru		

Załącznik Nr 4: Wykaz cech dystyngtywnych materiałów kartograficznych

(Cechy wymieniono w kolejności wynikającej z częstotliwości ich wykorzystywania do opisu materiałów kartograficznych)

1. Rodzaj materiału kartograficznego.
2. Przestrzeń (Terytorium/Obszar):
 - lokalizacja względna,
 - podział geograficzno-polityczny (administracyjny),
 - podział geograficzno-fizyczny,
 - podział geograficzno-historyczny,
 - podział regionalny.
3. Treść:
 - przedmiot,
 - ujęcie.
4. Czas:
 - data wydania (publikacji),
 - data sytuacji (pomiarów/aktualności).
5. Skala:
 - rodzaj skali,
 - wielkość skali.
6. Autorstwo (autor, kartograf, ...).
7. Język opisu.
8. Forma wydawnicza.
9. Format (wielkość).
10. Nakład.
11. Metoda prezentacji.
12. Sposób użytkowania/przechowywania.
13. Przeznaczenie.
14. Nośnik.
15. Barwność.
16. Orientacja.
17. Odwzorowanie.
18. Siatka kartograficzna.
19. Współrzędne geograficzne.
20. Dostępność.
21. Technika wykonania.
22. Technika druku.

CARTOGRAPHIC MATERIALS AS THE SOURCE OF INFORMATION

Summary

The cartographic materials are discussed as textual documents. The main qualifications and types of cartographic materials are presented. There are distinguished those distinctive features of the separate cartographic materials which have crucial influence on the relevance in the process of information retrieval. There were recognized as such: type of cartographic material, territory, subject, scale, cartographic mapping, net of co-ordinates, orientation and elements of bibliographic description.

КАРТОГРАФИЧЕСКИЕ МАТЕРИАЛЫ КАК ИСТОЧНИК ИНФОРМАЦИИ

Резюме

Картографические материалы обсуждены как текстовые документы. Названы основные определения и виды картографических материалов. Выделены те различающие признаки отдельных видов картографических материалов, которые имеют решающее влияние на релевантность в процессе поиска информации. Такими признаками автор считает в частности: вид картографического материала, территорию, тему, масштаб, картографическое отображение, сетку координат, ориентировку, а также элементы библиографического описания.

HANNA POPOWSKA, JERZY SOLAK

Instytut Informacji Naukowej,
Technicznej i Ekonomicznej

**UCZENIE SIĘ JĘZYKA NATURALNEGO PRZEZ SYSTEM KOMPUTEROWY
(NA PRZYKŁADZIE SYSTEMU DABINAL)**

Komputerowe systemy komunikacji w języku naturalnym. Uczenie się maszyn. Ogólna charakterystyka systemu DABINAL. Struktura systemu. Przykład uczenia się. Baza wiedzy językowej (słownik, wzorce zdań, reguły zmiany końcówek, definicje predyktorów). Język pośrednik - formuły interpretujące. Moduł analizy. Moduł uogólniający. Korzystanie z systemu DABINAL. Rezultaty dotychczasowych prac.

Tytuł artykułu sugeruje, że będzie on dotyczył zagadnień leżących na styku dwóch dziedzin badań nad zastosowaniami systemów komputerowych: maszynowego uczenia się i przetwarzania języka naturalnego. Dziedziny te należą do szerszej dyscypliny naukowej zwanej sztuczną inteligencją. Wprawdzie w przeszłości zwracano niekiedy uwagę na potencjalne zalety stosowania uczących się programów komputerowych do analizy języka naturalnego (np. /32/), stosunkowo niewiele prac poświęcono dotychczas łączeniu obu wskazanych dziedzin.

"Zagadnienia Informacji Naukowej" 1992 nr 1(59)

W tym artykule przedstawimy opis realizacji projektu, którego celem jest stworzenie systemu zdolnego do uczenia się języka naturalnego w zakresie niezbędnym do komunikacji z bazami danych. System komputerowy powstający w wyniku realizacji tego projektu nosi roboczą nazwę DABINAL (Data Base Interface in Natural Language). DABINAL należy do tzw. mobilnych (albo przenośnych) systemów komunikacji w języku naturalnym (SKJN), tzn. powinien zapewnić możliwość porozumiewania się w języku naturalnym z różnymi bazami danych, także dotyczącymi różnych dziedzin tematycznych. W odróżnieniu od innych SKJN, takich jak np. EUFID czy INTELLECT, DABINAL nie bazuje na wiedzy językowej przygotowanej przez ekspertów (językoznawców i specjalistów z danej dziedziny tematycznej), lecz sam - poprzez uczenie się - gromadzi potrzebną wiedzę językową.

DABINAL jest opracowywany w Instytucie Informacji Naukowej, Technicznej i Ekonomicznej (IINTE) w Warszawie. W 1991 r. zakończono prace nad wersją 1.0 tego systemu /27, 28/.

KOMPUTEROWE SYSTEMY KOMUNIKACJI W JĘZYKU NATURALNYM

Na nośnikach maszynowych zgromadzono już olbrzymie zbiory informacji. Informacje przechowywane w systemach skomputeryzowanych są sformatowane i uporządkowane zgodnie z wymaganiami określonymi dla poszczególnych systemów. Użytkownik korzystający z wielu systemów musi się więc nauczyć posługiwania się różnymi formalizmami, niezbędnymi do prawidłowego formułowania pytań, a opanowanie zasad porozumiewania się choćby z jednym systemem zazwyczaj bywa uciążliwe dla użytkownika. Na przykład, załóżmy, że w pewnym systemie informacyjnym do zadawania pytań używane są formuły podobne jak w rachunku predykatów pierwszego rzędu. Użytkownik, który chce

się np. dowiedzieć, jak się nazywa stolica Wybrzeża Kości Słoniowej, musi wiedzieć, że moduł wyszukiwawczy potrafi w bazie danych systemu zinterpretować predykt STOLICA, musi znać typy argumentów tego predykatu, a także musi znać reguły tworzenia formuły zapytania. W tym przypadku trzeba zapytać następująco:

STOLICA (\$X, WYBRZEŻE KOŚCI SŁONIOWEJ)?

Podobnie, pytania CZY TEGUCIGALPA JEST STOLICĄ HONDURASU? lub CZY STOLICA HONDURASU NAZYWA SIĘ TEGUCIGALPA? są reprezentowane przez formułę:

STOLICA (TEGUCIGALPA, HONDURAS)?

Konieczność stosowania nawet tak prostego formalizmu w celu komunikacji z systemem komputerowym stanowi dla wielu użytkowników barierę trudną do pokonania. Dlatego zapewnienie przyjaznego dla użytkowników łącza z komputerem jest jednym z tematów najbardziej absorbujących informatyków w ostatnich latach. O wadze tego problemu świadczy np. mnogość języków programowania, języków opisu danych oraz języków manipulowania danymi, które opracowano w celu umożliwienia i ułatwienia człowiekowi porozumienia z systemami komputerowymi. Z pewnością nie należy się spodziewać rozwiązania problemu komunikacji z komputerem poprzez stworzenie uniwersalnego narzędzia, które równocześnie zaspokoiłoby potrzeby np. programistów systemowych i osób korzystających ze skomputeryzowanej książki telefonicznej: zawodowi informatycy nadal korzystać będą z języków programowania (uniwersalnych i specjalizowanych), wspomaganych przez narzędzia ułatwiające pisanie, testowanie i poprawianie programów. Dla nieprofesjonalistów szukać jednak trzeba bardziej naturalnych i łatwiejszych do opanowania sposobów porozumiewania się z komputerem. W tej dziedzinie można wyróżnić trzy kierunki działania:

- umożliwienie użytkownikowi prowadzenia dialogu w języku naturalnym.
- zastosowanie schematu dialogu o strukturze hierarchicznej, z działaniem użytkownika ograniczonym do wyboru jednego z wariantów wskazanych w wykazie (menu - dialog jadłospisowy),
- wykorzystanie metod grafiki komputerowej do wprowadzania i wyprowadzania danych.

Metody graficzne i programy dialogowe stanowią niewątpliwie istotny krok naprzód w dziedzinie tworzenia przyjaznego dla użytkownika łącza z systemami komputerowymi, ale zakres ich zastosowań jest ograniczony: nie wszystkie informacje dają się przedstawić graficznie, a konwersacja z użyciem menu pozwala jedynie na wybór jednej z kilku podanych przez system możliwości. Można wskazać liczne sytuacje, w których obie te metody zawodzą i język naturalny pozostaje najlepszym środkiem przekazywania informacji, pozwalającym użytkownikowi uniknąć stosowania narzędzi formalnych.

Język naturalny jest i pozostanie najbardziej naturalnym środkiem porozumiewania się dla człowieka, trudno się więc dziwić temu, że próby dialogu z komputerem w języku naturalnym towarzyszą rozwojowi informatyki prawie od początków istnienia tej dyscypliny. Tym samym historia zautomatyzowanych systemów przetwarzania języka naturalnego liczy sobie obecnie ponad czterdzieści lat: prekursorskie dokonania teoretyczne i praktyczne w tej dziedzinie dotyczyły przekładu maszynowego, a rozpoczęcie prac teoretycznych nad wykorzystaniem komputera do tłumaczenia (z jednego języka naturalnego na inny) datowane jest w literaturze przedmiotu na koniec lat czterdziestych. W połowie lat pięćdziesiątych - dzięki stałemu ulepszaniu sprzętu komputerowego - zarysowała się realna możliwość sprawdzenia w praktyce koncepcji teoretycznych, a to z

kolei okazało się bodźcem inspirującym różnorodne badania teoretyczne, które następnie owocowały nowymi eksperymentami praktycznymi. Pierwsze tłumaczenie (z rosyjskiego na angielski) dokonane przez maszynę zaprezentowano w 1954 r., a pierwsze próby konwersacji z komputerem w języku naturalnym odbyły się w 1958 r. Tak więc, rozwój techniki komputerowej rozbudził w połowie lat pięćdziesiątych nadzieje na realizację (i to w stosunkowo nieodległej przyszłości) przy pomocy maszyny takich czynności, które dotychczas wymagały udziału inteligencji ludzkiej. I tak, znaczna część środowiska informatycznego uważała wówczas za możliwe przygotowanie w ciągu kilku lat programów zdolnych do dowodzenia twierdzeń matematycznych czy odpowiadania na pytania zadawane w języku naturalnym, dotyczące informacji zawartych w dostępnym zbiorze danych. Następne lata wykazały jednak, że w dziedzinie przetwarzania języka naturalnego okresy optymistycznych hipotez i sensacyjnych eksperymentów przeplatały się z okresami skrajnego pesymizmu i prawie zupełnej stagnacji. Na szczęście, porażki nie doprowadziły do zaniechania prac w tej dziedzinie, wpłynęły natomiast na istotne przeformułowanie ich celów. Skoncentrowano się na rozwiązywaniu problemów pragmatycznych, takich jak np. porozumiewanie się użytkowników z konkretnym systemem komputerowym za pomocą zdań "mieszczących się" (pod względem składni i tematyki) w określonym wycinku języka naturalnego. Do tego nurtu badań zaliczamy wszystkie prace mające na celu stworzenie systemów baz danych akceptujących dane wejściowe i pytania w języku naturalnym, a także prezentujących wyniki wyszukiwania w języku naturalnym. Systemy o takich właściwościach nazywane są systemami komunikacji w języku naturalnym (SKJN) lub operującymi językiem naturalnym łączami (interfejsami) do baz danych (lub - do systemów komputerowych), albo

krócej - "łączami naturalnymi" (ang. natural interfaces, natural front-ends).

Typowa metoda postępowania, stosowana w pracach nad tworzeniem SKJN polega na zadawaniu systemowi "z zewnątrz" całej potrzebnej mu wiedzy o języku naturalnym (reguł gramatycznych, słowników, zależności semantycznych między jednostkami leksykalnymi), przy czym wiedza ta w zasadzie nie ulega podczas działania systemu istotnym zmianom, choć zazwyczaj przewidywane jestokresowe uzupełnianie słowników. Większość znanych wykorzystywanych w praktyce systemów przetwarzania języka naturalnego, a więc także i SKJN, opracowano według powyższej metody. W przypadku SKJN oznacza to, że zmiana przedmiotowej bazy danych (z którą użytkownik ma się komunikować poprzez dany SKJN) wymaga dokonania wielu zmian w SKJN, a niekiedy nawet całkowitej przeróbki systemu. Trudność tę usunąć może jedynie stworzenie mobilnych (przenośnych) SKJN, tj. systemów łatwo adaptowalnych do nowych warunków, co jest możliwe do uzyskania poprzez wykorzystanie w systemie komputerowym algorytmów uczenia się. Główną zasadą przyjętą w systemie DABINAL jest właśnie zastosowanie mechanizmów uczenia się maszyn do stworzenia mobilnego SKJN, mającego służyć do komunikacji z bazami danych i z bazami wiedzy.

Problematyce przetwarzania języka naturalnego poświęcono bardzo wiele opracowań, skromny wybór literatury na ten temat przytoczono w bibliografii.

Na zakończenie tej części artykułu warto podkreślić, że istnieje wiele uproszczonych sposobów pisania programów, które dają użytkownikowi przyjemne uczucie uczestnictwa w dialogu z komputerem, prowadzonym w języku naturalnym: bardzo łatwo spowodować, by system komputerowy zadawał pytania w języku naturalnym. Pytania

mogą być nawet bardzo długie i złożone strukturalnie, pod jednym wszakże warunkiem: odpowiedzi można wybierać tylko z podanego przez system zestawu możliwości (jest to wspomniane już wyżej porozumiewanie się z komputerem za pomocą menu). Oczywiście, systemy operujące menu nie należą do systemów przetwarzania języka naturalnego i choć w wielu sytuacjach okazują się całkowicie wystarczające, to jednak zakres ich zastosowań jest ograniczony.

UCZENIE SIĘ MASZYN

"Obecne systemy sztucznej inteligencji, za wyjątkiem tych, których budowa ma na celu badanie maszynowego uczenia się, przejawiają niewielkie lub nie przejawiają żadnych umiejętności uczenia się. Cała wiedza, jaką dysponują, musi być uprzednio przygotowana i zaprogramowana. Jeśli w trakcie pracy wystąpi błąd, to nie potrafią samodzielnie go poprawić. Jest on powtarzany tyle razy, ile razy dana procedura jest wykonywana. Systemy te nie potrafią też doskonalić się w miarę zbierania doświadczeń ani gromadzić wiedzy ze swej dziedziny drogą eksperymentowania. Nie są w stanie automatycznie generować dla własnych potrzeb nowych algorytmów, formułować nowych abstrakcyjnych pojęć ani proponować nowych rozwiązań przez analogię do dotychczasowych czy też drogą swoistego procesu odkrywania (ang. discovery). Ogólnie rzecz biorąc, systemom tym brakuje umiejętności indukcyjnego wyciągania wniosków z podanej im informacji. Można powiedzieć, że niemal wszystkie istniejące systemy są dedukcyjne, gdyż wyciągają wnioski opierając się na dostępnej wiedzy i nie potrafią samodzielnie przyswajać ani generować wiedzy. Dla porównania, jedną z najbardziej uderzających cech w inteligentnym zachowaniu się człowieka jest zdolność przy-

swajania nowej wiedzy, nabierania wprawy i doskonalenia przez praktykę" /21/.

Komputerowe przetwarzanie języka naturalnego można powiązać z badaniami nad systemami zdolnymi do uczenia się. Herbert Simon, jeden z klasyków sztucznej inteligencji, raczej sceptycznie nastawiony do prac związanych z uczeniem się maszyn, uważa jednak analizę języka naturalnego za jedną z nielicznych dziedzin, w których systemy uczące się mogą znaleźć rzeczywiste zastosowanie /32/. Dotyczy to szczególnie komunikacji w języku naturalnym między różnego rodzaju systemami użytkowymi i ich użytkownikami. Przykładem tego rodzaju podejścia może być system NANOKLAUS /11/, zdolny do gromadzenia wiedzy z nowych dziedzin poprzez dialog z użytkownikami. System wyposażony jest w ustalony zbiór reguł syntaktycznych i semantycznych, pokrywających niewielki podzbiór języka angielskiego. NANOKLAUS ma także "zarodek" słownika, który może być rozszerzany w miarę zdobywania przez system wiedzy o nowej dziedzinie. Gramatyka systemu składa się z niewielkiej liczby wyspecjalizowanych reguł, odpowiadających zdaniom różnych typów. Uczenie się systemu polega na wyjaśnianiu nowych wyrażeń poprzez wskazywanie ich relacji z innymi wyrażeniami, już wprowadzonymi do słownika.

Istnieją dwa podstawowe rodzaje uczenia się:

- gromadzenie wiedzy, tzn. budowanie abstrakcyjnego modelu obserwowanych sytuacji,
- doskonalenie umiejętności (motorycznych i poznawczych) przez powtarzanie ćwiczeń.

Oba rodzaje uczenia się są wykorzystywane w systemie DABINAL. Z założenia DABINAL jest systemem uczącym się języka, musi więc traktować język jako część rzeczywistości, której model próbuje

tworzyć. Jednocześnie DABINAL ma służyć konkretnemu praktycznemu celowi, jakim jest ułatwienie przepływu informacji między użytkownikiem i komputerem, musi więc operować także na modelu rzeczywistości, będącej przedmiotem zastosowania systemu. Tworzenie modelu rzeczywistości językowej nie jest jednak celem samoistnym, ma bowiem służyć doskonaleniu zdolności systemu do analizowania zdań języka naturalnego, a zatem mamy w tym wypadku do czynienia także i z drugim z wymienionych rodzajów uczenia się.

Podstawową techniką uczenia się systemu DABINAL jest uczenie się indukcyjne. Celem uczenia się jest wypracowanie, na podstawie podawanych przykładów, metod przekształcania zdań w języku naturalnym do postaci formuł języka sformalizowanego, będącego rozszerzeniem języka rachunku predykatów pierwszego rzędu.

Na temat uczenia się maszyn można zebrać pokaźną literaturę, przykładowo wymienimy następujące opracowania: /9, 20, 21, 22, 23, 24/.

OGÓLNA CHARAKTERYSTYKA SYSTEMU DABINAL

System DABINAL ma do spełnienia podobne cele, jak mobilne systemy komunikacji z bazami danych w języku naturalnym (SKJN) - chodzi w nim o zapewnienie możliwości porozumiewania się w języku naturalnym z różnymi bazami danych, dotyczącymi różnych dziedzin tematycznych - jednakże środki zastosowane w systemie DABINAL są odmienne niż w innych projektach z tego zakresu, opiera się on bowiem na wiedzy językowej, zgromadzonej poprzez samodzielne uczenie się, a nie na wiedzy przygotowanej przez ekspertów językowych.

Głównym celem projektu DABINAL jest sprawdzenie skuteczności metod, stosowanych w dziedzinie uczenia komputerów, w odniesieniu

do systemów komunikacji w języku naturalnym z bazami danych. Aby utrzymać klarowne warunki eksperymentu, przyjęto, że system nie będzie dysponował innymi rodzajami informacji językowych, jak tylko te, które samodzielnie zgromadzi drogą uczenia się. Założenie to nie oznacza bynajmniej, że technika samodzielnego zdobywania przez system wiedzy językowej powinna być uważana za jedyne narzędzie przydatne do konstruowania SKJN. W tak złożonej dziedzinie, jaką jest przetwarzanie języka naturalnego, trzeba brać pod uwagę równoległe stosowanie narzędzi różnego rodzaju (na przykład, wiedza dotycząca użycia spójników logicznych jest łatwiejsza do przekazania systemowi w formie reguł opracowanych przez człowieka, niż za pomocą przykładów, jednak w systemie DABINAL świadomie ograniczono dodatkowe źródła wiedzy językowej.

DABINAL należy do kategorii systemów uczących się na podstawie przykładów. Typowymi zadaniami takich systemów jest klasyfikacja obiektów na kategorie i znajdowanie ogólnych opisów tych kategorii. Klasyfikacja jest jednym z zadań stojących przed DABINAL-em: zdania o podobnej strukturze powinny być łączone w jedną kategorię. System musi się także nauczyć wybierać ze zdań frazy odpowiadające różnym elementom formuły interpretującej i przekształcać je do podstawowej formy gramatycznej. Aby ten cel osiągnąć, system musi stosować w szerszym zakresie narzędzia analityczne, odwołujące się do wiedzy wykraczającej poza zasób poznanych empirycznie faktów. W przypadku DABINAL-a wiedza tego rodzaju ogranicza się do niewielu prostych reguł, w rodzaju "poszukując odpowiedniej reguły zamiany końcówek, weź pod uwagę ostatnią literę tematu". Dzięki istnieniu elementów wiedzy analitycznej, możliwe jest wyciąganie pewnych wniosków uogólniających, nawet na podstawie pojedynczego przykładu.

DABINAL był projektowany z myślą o systemach komunikujących się z bazami danych w języku polskim. Założenia dotyczące języka naturalnego, przyjęte w trakcie realizacji systemu, są jednak tak ogólne, że pozostają w mocy także dla wielu innych języków europejskich, których gramatyka wywodzi się z wzorca łacińskiego. W wersji 1.0 systemu dodatkowo, aby ułatwić jego zastosowanie do analizy zdań w takich językach jak angielski, francuski i niemiecki, wydzielono przedimki i rodzajniki jako osobną kategorię słów, traktowaną w specjalny sposób. Dzięki operowaniu tylko ogólnymi założeniami dotyczącymi języków naturalnych możliwe jest uczenie systemu DABINAL "rozumienia" zdań w różnych językach, odnoszących się do różnych dziedzin zastosowań. Eksperymenty, które przeprowadzono dotychczas z wersją prototypową i wersją 1.0, używając, oprócz języka polskiego, także angielskiego i niemieckiego, nie wykazały zasadniczych różnic w zdolności przyswajania przez system tych języków, choć oczywiście różna była sprawność działania w różnych sytuacjach. Na przykład, opanowanie prawidłowego przekształcania różnych form fleksyjnych do postaci podstawowej, które w przypadku języka polskiego jest zadaniem pracochłonnym, wymagającym analizy wielu przykładów, dla języka angielskiego jest o wiele prostsze i już po kilkunastu przykładach można spodziewać się prawidłowego działania systemu w tym zakresie.

Realizacja wersji 1.0 nie była ostatecznym celem prac nad systemem DABINAL. W zamierzeniach prace nad projektem DABINAL mają przebiegać w dwóch etapach. Celem pierwszego z nich, którego efektem jest powstanie wersji 1.0, było stworzenie takich procedur, które pozwalałyby systemowi uczyć się rozróżniania zdań o różnej strukturze i przekształcania tych zdań do postaci formuł pewnego języka formalnego w sposób podobny jak w przykładach treningowych.

W tej wersji systemu jego zdolności do uogólniania wiedzy są dosyć skromne: dotyczą one głównie przenoszenia cech zaobserwowanych dla pewnych słów, zwrotów i konstrukcji zdaniowych na inne słowa, zwroty i zdania, należące do tej samej kategorii.

Etap drugi będzie się koncentrował na procedurach poszukujących podobieństw między elementami wiedzy zgromadzonej w bazie wiedzy językowej i na procedurach dokonujących uogólniania tej wiedzy. Uogólnienia te mogą przebiegać na różnych poziomach. Pierwszy z nich dotyczy słów i zwrotów występujących w podobnych kontekstach. Może to np. prowadzić do wydzielenia nowych klas słów i zwrotów. Jeśli dla pewnej grupy wyrazów zostanie zaobserwowany wspólny zbiór reguł zamiany końcówek, wyrazy te mogą zostać wydzielone jako odrębna kategoria deklinacyjna lub koniugacyjna, a przypisane do tej kategorii reguły zamiany końcówek utworzą nowy wzorzec odmiany. W ten sposób powstają w systemie nowe kategorie, które mogą ułatwić analizę nowych przykładów zdań. Proces uogólniania wiedzy jest wielopoziomowy, tzn. procedury uogólniania mogą się odwoływać do reguł i kategorii, które same są rezultatem wcześniejszych uogólnień. Można powiedzieć, że system wyposażony w reguły działania tego rodzaju dysponuje nie tylko umiejętnością porządkowania faktów według zadanych z góry kategorii, ale potrafi również tworzyć nowe kategorie, pozwalające uogólniać wiedzę o poznanych faktach.

W uczących się systemach komputerowych wiedzą ogólną kształtuje się stosunkowo późno - dopiero wtedy, gdy system zgromadzi już dostatecznie dużo wiedzy faktograficznej. Wiedza ogólna jest ze swej natury w znacznym stopniu niezależna od dziedziny zastosowań systemu, dzięki czemu może stanowić trwałą bazę systemu

wówczas, gdy jest on przenoszony do nowych zastosowań. Wiedza faktograficzna jest w takiej sytuacji znacznie mniej przydatna.

Działanie systemu DABINAL opiera się na analizie przykładów przedstawionych przez nauczyciela. Każdy przykład składa się ze zdania w języku naturalnym i odpowiadającej mu formuły w języku pośredniku. Rozważane są zarówno zdania oznajmujące, jak i zapytania. Przyjęty język pośrednik jest odmianą języka rachunku predykatów. Przyjęto założenie, że nawet w przypadkach, gdy zdanie ma złożoną strukturę logiczną, odpowiednia formuła zawiera tylko jeden predykt, który z kolei może być zdefiniowany poprzez złożoną formułę. Umożliwia to m.in. wyraźne odgraniczenie zadania "zrozumienia" pytania od zadania odpowiedzi na to pytanie. Np. zdanie "Do jakiego morza wpada rzeka, która przepływa przez Warszawę?" może być zinterpretowane za pomocą formuły " $P(\text{Warszawa}, \$x)$?", gdzie P jest predykatem dwuargumentowym, którego pierwszy argument należy do kategorii miasto, a drugi do kategorii morze. Predykat P może posiadać interpretację odwołującą się do innych predykatów, np.:

$$P(x, y) \Leftarrow (\exists z) (\text{przepływa}(z, x) \wedge \text{wpada}(z, y))$$

ale dla DABINAL-a interesująca jest jedynie zależność między zdaniem i predykatem P , a nie predykatami przepływa i wpada. Ten drugi rodzaj zależności może być użyteczny dla modułu tłumaczącego formuły języka pośrednika na formuły języka zapytań lub języka operowania danymi systemu zarządzania bazą danych. Należy podkreślić, że DABINAL jest tylko łączem komunikacyjnym, nie stanowi natomiast systemu odpowiadającego na pytania, który powinien wnioskować z danych faktów o innych faktach i odpowiadać na skomplikowane pytania. W przeciwieństwie do tego typu systemów,

DABINAL wyraźnie rozgranicza zadanie "rozumienia" pytania od zadania odpowiadania na nie. W przytoczonym wyżej przykładzie "rozumienie" pytania polega na kojarzeniu go z predykatem P, podczas gdy odpowiedź na pytanie wymaga dodatkowo wzięcia pod uwagę zależności predykatu P od predykatów przepływa i wpada.

Podejście przyjęte w systemie DABINAL może być owocne wówczas, gdy spodziewamy się, że użytkownicy będą się posługiwać zdaniami zbudowanymi według ograniczonej liczby schematów. W takim przypadku należy systemowi przedstawić po kilka przykładów każdego spodziewanego schematu, łącznie z odpowiednimi formułami interpretacyjnymi. Wiedza czerpana z tych przykładów nie zawsze całkowicie wystarcza do rozróżnienia dowolnych dwóch schematów. W sytuacjach wątpliwych rozstrzygnięcia mogą być dokonywane na podstawie danych zaczerpniętych z przedmiotowej bazy danych lub uzyskanych od użytkownika.

W przypadku języków o zróżnicowanej fleksji stosunkowo dużego wysiłku wymaga nauczanie systemu przekształcania form fleksyjnych w formę podstawową. Jednakże dzięki temu, że tego rodzaju wiedza jest niezależna od dziedziny zastosowań systemu, może być ona zachowana przy zmianie środowiska działania systemu, a zatem proces uczenia fleksji nie musi być powtarzany dla każdego nowego zastosowania.

W systemie DABINAL przyjęto jako główne założenie, że wiedza językowa jest gromadzona wyłącznie poprzez uczenie się na przykładach przedstawionych przez nauczyciela i użytkowników w trakcie dialogu z nimi. Procedury uczenia się są jedynym instrumentem zdobywania sprawności językowej i tworzenia wiedzy językowej. System rozpoczyna działanie w stanie zupełnej niewiedzy na temat języka, którego ma się uczyć. Oczywiście istnieje wiele wartościowo-

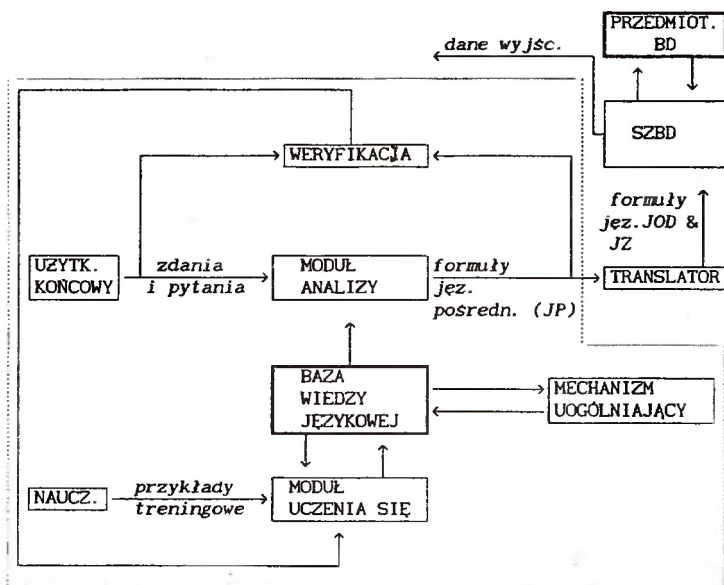
wych źródeł informacji językowej, włączając w to przedmiotową bazę danych, które mogłyby zostać wykorzystane do celów analizy językowej. W systemie DABINAL z tych dodatkowych źródeł zrezygnowano, aby nie zakłócać czystości eksperymentu.

STRUKTURA SYSTEMU DABINAL

System DABINAL może być łatwo adaptowany do różnych języków naturalnych dzięki temu, że rozpoczyna działanie bez wstępnych założeń na temat języka, który będzie używany do komunikacji z bazą danych. Elastyczność tę ograniczają jedynie ogólne zasady leżące u podstaw aparatu uczenia się. Zasady te są słuszne dla języka polskiego, angielskiego i innych języków europejskich, ale ich znaczenie może być różne dla różnych języków. W DABINAL-u zastosowano więc pewną liczbę parametrów numerycznych, zmieniających heurystyki uczenia się i analizy w zależności od języka, którego system się uczy.

Rysunek 1 przedstawia ogólną strukturę systemu DABINAL. Bloki objęte kropkowaną linią stanowią najistotniejszą część systemu i dlatego poświęcono im najwięcej uwagi. Pozostałe bloki związane są z konkretną przedmiotową bazą danych i z tego względu są one dla systemu DABINAL mniej interesujące.

Zakłada się, że przedmiotowa baza danych, z którą system współpracuje, jest skonstruowana według modelu relacyjnego. Nie przyjęto jednak żadnych wstępnych założeń co do zawartości tej bazy ani co do systemu zarządzania bazą danych (SZBD). Moduł uczenia się i moduł analizy są najważniejszymi elementami systemu. Pierwszy gromadzi wiedzę na podstawie analizy przykładów dostarczanych przez nauczyciela, a drugi stosuje tę wiedzę do przetwa-



Rys. 1. Ogólna budowa systemu DABINAL.

rzania zdań języka naturalnego, wprowadzanych przez użytkownika. Ponieważ istnieje bliski związek między relacyjnymi bazami danych a rachunkiem predykatów pierwszego rzędu, przyjęto pewne rozszerzenie tego rachunku jako język pośrednik (JP), który oddziela moduł analizy od konkretnego SZBD. Translator przekształca formuły tego języka na komendy języka zapytań (JZ) lub języka operowania danymi (JOD), zgodnie z zastosowanym systemem zarządzania bazą danych (SZBD).

DABINAL zdobywa zdolność operowania na zdaniach języka naturalnego, ucząc się z przykładów dostarczanych przez nauczyciela. Każdy przykład składa się ze zdania w języku naturalnym i odpowiadającej mu formuły w języku pośredniku. Zadaniem modułu uczenia się jest wykrycie relacji między zdaniami a formułami oraz połączenie elementów zdań w języku naturalnym z predykatami i ich argu-

mentami. Wiedza zdobywana przez ten moduł jest przechowywana w bazie wiedzy językowej. Wiedza ta jest przedstawiona jako zbiór faktów określonych poprzez analizę przykładów. Mechanizm uogólniający ma za zadanie budowanie bardziej ogólnych stwierdzeń na podstawie podobnych faktów zaobserwowanych w bazie wiedzy językowej. Pewne zasady uogólnień, np. odpowiednie uporządkowanie faktów, są realizowane przez moduł uczenia się, ale procedury wnioskujące oparte na subtelniejszych przesłankach zostały zgromadzone w ramach modułu uogólniającego.

Baza wiedzy językowej jest głównym źródłem danych dla modułu analizy. Moduł ten przekształca zdania w języku naturalnym, wprowadzone przez użytkownika, na formuły języka pośrednika. Podstawowa droga łącząca moduł uczenia się z modułem analizy biegnie poprzez bazę wiedzy językowej. Jest też drugie połączenie, przez blok weryfikacji. Blok ten uaktywnia się albo w przypadku wykrycia niespójności między wejściem i wyjściem modułu analizy, albo gdy blok ten nie jest w stanie samodzielnie podać żadnej formuły wyjściowej odpowiadającej zdaniu wejściowemu. W takim przypadku nauczyciel musi podać modułowi uczenia się prawidłową formułę w języku pośredniku.

Schemat połączeń między blokami funkcjonalnymi, przedstawiony na rys. 1, odpowiada docelowej organizacji systemu DABINAL. W zrealizowanej wersji 1.0 system traktuje współpracującą z nim osobę jednocześnie jako nauczyciela i zwykłego użytkownika. Realizacja modułu uogólniającego przewidziana jest na dalsze etapy pracy nad systemem. W dotychczas przeprowadzonych eksperymentach nie prowadzono współpracy z żadnymi przedmiotowymi bazami danych. W związku z tym wszelkie informacje o przedmiotowej bazie danych, które można by odczytywać z niej samej (np. jakie znajdują się w niej relacje,

jakie są typy ich atrybutów) w wersji 1.0 muszą być wprowadzane przez użytkownika w momencie, gdy staje się one systemowi potrzebne. Dokładniejszy opis bloków funkcjonalnych z rys. 1 znajduje się w następnych częściach tego artykułu. Zanim jednak przejdziemy do ich omówienia, celowe będzie prześledzenie przykładu sesji uczenia się, ilustrującego działanie systemu.

PRZYKŁAD UCZENIA SIĘ

Poniższy dialog dotyczy języka polskiego i pokazuje, między innymi, jak system radzi sobie z różnymi formami fleksyjnymi wyrazów. W początkowej fazie dialogu używane są jedynie jednowyrazowe identyfikatory obiektów. Zwroty wielowyrazowe zostały zademonstrowane na końcu dialogu.

Zakładamy, że system "startuje" bez jakiejkolwiek wiedzy o języku komunikowania się i o przedmiotowej bazie danych. W takim przypadku pierwszą jego akcją jest zainicjowanie bazy wiedzy językowej. Następnie sterowanie jest przekazywane do głównej pętli programu, składającej się z następujących kroków:

- wprowadzenie zdania w języku naturalnym przez nauczyciela,
- próba interpretacji zdania przez moduł analizy, wspomagany przez nauczyciela,
- jeśli próba ta nie powiedzie się, przyjęcie przez moduł uczący się prawidłowej interpretacji od nauczyciela.

Przyjmijmy, że pierwszym zdaniem wprowadzonym do systemu jest:

Warszawa leży nad Wisłą.

System nie rozporządza jeszcze żadną wiedzą językową, nie może więc samodzielnie zanalizować tego zdania. Wobec tego nauczyciel jest proszony o wprowadzenie formuły interpretującej to

zdanie w języku pośredniku. W tym przypadku będzie to formuła:

leżeć nad rzeką (Warszawa, Wisła).

Ponieważ "leżeć nad rzeką" jest predykatem nieznanym jeszcze systemowi, nauczyciel musi zdefiniować typy jego argumentów:

argument 1: miasto

argument 2: rzeka

Na podstawie powyższego przykładu DABINAL może wyciągnąć pewne wnioski na temat:

- kategoryzacji leksykograficznej słów,
- zależności między strukturą zdania i ogólnym schematem formuły interpretującej to zdanie,
- przekształceń form fleksyjnych.

Niżej przytoczono przykłady wypracowanych danych szczegółowych i reguł decyzyjnych:

- Wisła należy do kategorii rzeka,
- Wisłą jest formą fleksyjną słowa Wisia,
- Wisła występuje jako wartość drugiego argumentu predykatu "leżeć nad rzeką",
- zdanie o ogólnej budowie "A' leży nad B'" wiąże się z formułą "leżeć nad rzeką (A, B)", gdzie A i B są frazami należącymi odpowiednio do kategorii miasto i rzeka, A' i B' są formami fleksyjnymi tych fraz,
- frazy A' i A, w powyższej regule decyzyjnej są identyczne,
- jednosłowna fraza B może być uzyskana z B' poprzez zamianę końcówki ą na a, jeśli ł jest ostatnią literą tematu.

Lista podobnych wniosków, wprowadzonych z przykładu treningowego, jest znacznie dłuższa.

Dalszy ciąg dialogu przedstawiono niżej.

DABINAL swoje komunikaty podaje w języku angielskim i w poniższym zapisie dialogu zostało to zachowane. Dialog rzeczywisty nieco się różni od przedstawionego tutaj, ze względu na zastosowanie wielu okien na ekranie monitora.

Enter the next sentence.

Paryż leży nad Sekwaną.

Does PARYŻ belong to a category MIASTO?

Y

Is SEKWANA a basic form of SEKWANA?

Y

Does SEKWANA belong to a category RZEKA?

Y

My interpretation of the sentence is:

LEŻEĆ NAD RZEKĄ (PARYŻ, SEKWANA).

Is it right?

Y

Celem tego fragmentu dialogu było sprawdzenie, czy system jest w stanie rozpoznać wzorzec zdaniowy, wyuczony na podstawie wcześniejszego przykładu, oraz czy ten wzór może być poprawnie przeanalizowany. W przedstawionym zdaniu DABINAL rozpoznaje tę samą strukturę ogólną, jaką już zna, i kojarzy ją z predykatem "leżeć nad rzeką". Pozwala to systemowi postawić hipotezę, że Paryż należy do tej samej kategorii co Warszawa (miasto), oraz Sekwaną należy do tej samej kategorii co Wisła (rzeka). Ponadto, ponieważ Sekwaną ma tę samą końcówkę co Wisłą, jest bardzo prawdopodobne, że reguła zastępowania końcówek zaobserwowana uprzednio (-ą na -a) może być zastosowana także do tego przykładu. Wszystkie te hipotezy wymagają akceptacji użytkownika, ale zostały utworzone bez żadnej pomocy z jego strony.

Do tej chwili DABINAL przeanalizował dwa przykłady o tej samej strukturze zdania. Następny przykład treningowy przedstawia zdanie o nowej strukturze i jego formalną interpretację. Tym razem będzie to pytanie.

Enter the next sentence.

Nad jaką rzeką leży Londyn?

I cannot interpret this sentence. Enter your interpretation:

leżeć nad rzeką (Londyn, \$x)?

Enter the next sentence.

Nad jaką rzeką leży Paryż?

My interpretation of the sentence is:

LEŻEĆ NAD RZEKĄ (PARYŻ, \$X)?

Pierwsze z powyższych dwóch pytań wystarcza systemowi do nauczenia się analizy zdań o schemacie "Nad jaką rzeką leży x?"¹¹ Celem następnego pytania, podanego przez nauczyciela, było sprawdzenie, czy nauczanie przyniosło spodziewany efekt. W tym wypadku wiedza językowa zdobyta już przez system była dostateczna do analizy postawionego pytania bez żadnej pomocy nauczyciela.

Warto zademonstrować, w jaki sposób DABINAL uczy się rozróżniania zdań o podobnej budowie, które powinny być interpretowane w odmienny sposób.

Enter the next sentence.

Gdańsk leży nad Bałtykiem.

Does GDAŃSK belong to a category MIASTO?

Y

I cannot find the basic form of BAŁTYKIEM. Enter it:

Bałtyk

Does BAŁTYK belong to a category RZEKA?

N

I cannot interpret this sentence. Enter your interpretation:

leżeć nad morzem (Gdańsk, Bałtyk).

LEŻEĆ NAD MORZEM is a new predicate. What are its arguments?

argument 1: miasto

argument 2: morze

Enter the next sentence.

Triest leży nad Adriatykiem.

Does TRIEST belong to a category MIASTO?

Y

Is ADRIATYK a basic form of ADRIATYKIEM?

Y

Does ADRIATYK belong to a category MORZE?

Y

My interpretation of the sentence is:

LEŻEĆ NAD MORZEM (TRIEST, ADRIATYK).

Pierwsze ze zdań podanych w tym fragmencie ma identyczną strukturę jak zdania znane z wcześniejszych przykładów, ale jego interpretacja jest inna, bo inna jest kategoria semantyczna słowa Bałtyk. Nie należy ono do kategorii rzeka, lecz do kategorii morze, ale DABINAL nie ma jeszcze informacji pozwalającej na zrobienie tego rozróżnienia i próbuje interpretować to zdanie za pomocą predykatu "leżeć nad rzeką". Próba się nie udaje i nauczyciel musi zdefiniować nowy predykat "leżeć nad morzem".

Analizując następny przykład, system ma już do wyboru dwa schematy pasujące do przedstawionego mu zadania i interpretowane za pomocą dwóch różnych predykatów: "leżeć nad rzeką" i "leżeć nad morzem". Musi zatem dokonać wyboru jednego z nich. W tym przypadku kategoria słownikowa wyrazu Adriatyk mogłaby być czynnikiem różniącym: jeśli należy on do kategorii rzeka to powinno się

wybrać predykat "leżeć nad rzeką", w przeciwnym przypadku, jeśli wyraz ten należy do kategorii morze, to prawidłowym wyborem będzie "leżeć nad morzem". Ale system nie zna jeszcze klasyfikacji słowa Adriatyk, wobec tego próbuje zastosować drugorzędne kryteria rozpoznawcze (w tym przypadku końcówka -iem, napotkana wcześniej przy słowie Bałtyk, usprawiedliwia założenie, że Adriatyk, podobnie jak Bałtyk, należy do kategorii morze i wobec tego należy użyć predykatu "leżeć nad morzem").

Na koniec przedstawimy przykład, w którym pojawia się zwrot wielowyrazowy.

Enter the next sentence:

Ustronie Morskie lezy nad Morzem Bałtyckim.

Is USTRONIE MORSKIE a meaningful phrase?

Y

Is MORZEM BAŁTYCKIM a meaningful phrase?

Y

Enter the basic form of MORZEM BAŁTYCKIM:

Morze Bałtyckie

Does MORZE BAŁTYCKIE belong to a category MORZE?

Y

My interpretation of the sentence is:

LEŻEĆ NAD MORZEM (USTRONIE MORSKIE, MORZE BAŁTYCKIE).

Is it right?

Y

Zdanie wprowadzone przez użytkownika strukturalnie pasuje (podobnie jak w poprzednim przykładzie) do dwóch znanych systemowi schematów, reprezentowanych przez predykaty "leżeć nad rzeką" i "leżeć nad morzem", pod jednym wszakże warunkiem: zwroty "Ustronie Morskie" i "Morzem Bałtyckim" nie mogą być przypadkowymi zbitkami

wyrazów, lecz muszą stanowić logiczne całości. Aby się o tym upewnić, system zadaje odpowiednie pytanie użytkownikowi i uzyskuje na nie odpowiedź pozytywną. Następnie system próbuje odgadnąć podstawową formę zwrotu "Morzem Bałtyckim", ale nie ma na to jeszcze dostatecznie szerokiej wiedzy. W związku z tym prosi użytkownika o wprowadzenie formy podstawowej tego zwrotu. Wreszcie, podobnie jak w poprzednim przykładzie, musi dokonać wyboru między dwoma schematami interpretacyjnymi. Dzięki występowaniu w zwrocie "Morze Bałtyckie" aż dwóch członów mających powiązanie z kategorią morze, system skłonny jest zakładać, że Morze Bałtyckie należy raczej do kategorii morze niż rzeka, a zatem właściwym schematem interpretacyjnym powinien być ten związany z predykatem "leżeć nad morzem". Po upewnieniu się, że tak jest w istocie, DABINAL odtwarza formułę interpretacyjną, właściwą dla wprowadzonego zdania.

Należy zwrócić uwagę, że w ostatnim przykładzie, w celu analizy zdania zawierającego zwroty wielowyrazowe, system posłużył się schematem wyuczonym na przykładach zawierających tylko zwroty jednowyrazowe.

BAZA WIEDZY JĘZYKOWEJ

Wiedza językowa systemu DABINAL dzieli się na kilka części:

- słownik,
- wzorce zdań,
- reguły zamiany końcówek,
- definicje predykatów.

Słownik

Wszystkie wyrażenia występujące w zdaniach i formułach wejściowych są pamiętane w słowniku. Przechowuje on wszystkie znane

systemowi formy fleksyjne wyrazów i poszczególnych słów. Elementy słownika należą do czterech kategorii: nazwa predykatu, nazwa atrybutu (argumentu predykatu), wartość atrybutu i słowo pomocnicze. Element słownika może należeć jednocześnie do kilku kategorii, np. w pewnych kontekstach może spełniać rolę słowa pomocniczego, a w innych - być wartością atrybutu. Element słownika będący wartością atrybutu może być związany z więcej niż jednym atrybutem, np. Warszawa może należeć zarówno do wartości atrybutu "miasto", jak i atrybutu "pojęcie geograficzne" (czy "obszar"), którego zakres jest szerszy.

Jak już podano wyżej, słownik systemu DABINAL przechowuje zarówno pojedyncze słowa, jak i zwroty wielowyrazowe, we wszystkich napotkanych formach gramatycznych. Jedną z tych form, nazywaną formą podstawową, jest wyróżniona. W przypadku rzeczowników powinien być to mianownik liczby pojedynczej, w przypadku czasowników - bezokolicznik. Oczywiście forma podstawowa zwrotu wielowyrazowego może zawierać słowa, które występują w formach różnych od podstawowej.

W przypadku przymiotników decyzja nie jest tak jednoznaczna jak dla rzeczowników i czasowników. Problem polega na tym, czy wybrać wspólną formę podstawową dla wszystkich rodzajów przymiotnika (np. mianownik liczby pojedynczej rodzaju męskiego), czy uznać odrębność formy podstawowej dla każdego rodzaju. W tym drugim przypadku wyrazy wysoki, wysoka i wysokie będą traktowane jako osobne jednostki leksykalne, a nie jako formy tego samego przymiotnika. Obydwa wymienione rozwiązania mają swoje dobre i złe strony, które mogą okazać się istotne dla modułu uogólniającego wiedzę językową. W obecnej wersji systemu, w niektórych przypadkach, może się okazać, że korzystniejsze jest operowanie

jednostkami leksykalnymi o mniejszej liczbie form fleksyjnych. W takiej sytuacji zalecanym rozwiązaniem jest uznanie różnych rodzajów tego samego przymiotnika za osobne słowa, mające różne formy podstawowe.

Podobny problem, jak w przypadku przymiotników, pojawia się w odniesieniu do imiesłów i przysłówków. W pewnych sytuacjach korzystniejsze może być uważanie imiesłów za formy czasownika, a przysłówków - za formy przymiotnika, w innych lepsze efekty może dawać uznanie imiesłowu i przysłówka za samoistne kategorie leksykalne.

Szczególny problem dotyczący klasyfikacji słów pojawia się w języku angielskim. Spotyka się w nim dosyć często wyrazy, które w identycznej formie mogą być czasownikiem i rzeczownikiem (np. call), albo czasownikiem i przymiotnikiem (np. open) itd. W tej chwili DABINAL nie ma środków, które pozwalałyby na oddzielenie w słowniku np. wyrazu open w funkcji czasownika od tego samego wyrazu w funkcji przymiotnika. Dla obecnej wersji systemu nie stanowi to istotnej przeszkody w działaniu, ale dla modułu uogólniania wiedzy może być pewnym utrudnieniem.

Wzorce zdań

Jedną z procedur przetwarzających przykłady treningowe polega na kojarzeniu oddzielnych elementów zdania z argumentami (atributami) odpowiedniego predykatu. W rezultacie tworzony jest wzorzec zdania, zawierający nazwy argumentów w miejscach zawierających poprzednio wartości tych argumentów, np. dla zdaniaⁿ "Warszawa jest stolicą Polski"ⁿ utworzony wzorzec zdania będzie miał postaćⁿ "⟨miasto⟩ jest stolicą ⟨kraj⟩". Dodatkowo DABINAL wypracowuje reguły przekształcania poszczególnych elementów wzorca zdania w wartości

argumentów odpowiedniego predykatu. Wśród tych reguł szczególną rolę spełniają reguły zamiany końcówek fleksyjnych. Są one związane z tymi elementami wzorca zdania, dla których zostały zaobserwowane.

Reguły zamiany końcówek

Wyrażenia, będące wartościami argumentów, przed wstawieniem w odpowiednie miejsca formuły języka pośrednika, podlegają transformacjom od formy fleksyjnej, w której wystąpiły w zdaniu, do formy podstawowej. Reguły zamiany końcówek opisują te transformacje w skrótowy sposób. Każda reguła zamiany końcówek zawiera końcówkę pewnej formy fleksyjnej, końcówkę odpowiedniej formy podstawowej słowa i ostatnią literę tematu. Z regułami zamiany końcówek związane są liczniki, wskazujące liczbę udanych zastosowań reguły. Dzięki temu reguły częściej używane zyskują pierwszeństwo przed regułami stosowanymi tylko w rzadkich przypadkach. Priorytet reguł jest także związany z kontekstem słowa, które ma być poddane przekształceniu. Reguły zaobserwowane w podobnym kontekście są brane pod uwagę w pierwszej kolejności.

Definicje predykatów

Oprócz danych czysto językowych, baza wiedzy językowej zawiera informacje dotyczące języka pośrednika. Nazwy predykatów i nazwy ich argumentów są przechowywane we wspólnym słowniku z innymi rodzajami jednostek leksykalnych. Definicje predykatów są natomiast umieszczane w odrębnym pliku.

Baza wiedzy językowej systemu DABINAL nie ma jednolitej struktury, dlatego też nie można było zastosować do tworzenia i obsługi tej bazy dostępnych na rynku systemów obsługi baz da-

nych. Niezbędne było przygotowanie własnego zestawu środków programowych, projektowanych specjalnie dla systemu DABINAL.

JĘZYK POŚREDNIK - FORMUŁY INTERPRETUJĄCE

Do przedstawiania treści zdań języka naturalnego używa się w systemie DABINAL języka formalnego, przypominającego język rachunku predykatów. Dzięki temu system jest niezależny od organizacji przedmiotowej bazy danych i systemu zarządzania tą bazą. Zmiana systemu zarządzania bazą danych wymaga odpowiedniego przystosowania modułu translatora (por. rys. 1), lecz nie ma wpływu na pozostałe istotne moduły systemu.

W systemie DABINAL poczyniono założenie, że interesują nas zdania oznajmujące, definiujące relacje między pewnymi obiektami, oraz zdania pytające, wymagające podania informacji o obiektach będących w danej relacji z innymi obiektami. Każda relacja wyrażona przy pomocy zdania w języku naturalnym jest związana z pewnym predykatem - elementem języka formalnego. Każdy predykat ma argumenty należące do określonych dziedzin. Formuła języka formalnego zawiera nazwy obiektów lub zmiennych w miejscach przeznaczonych na argumenty. Formuły mogą być zakończone kropką lub znakiem zapytania, zależnie od tego, czy reprezentują zdanie oznajmujące, czy pytanie. Na przykład, następujące formuły:

stolica (Polska, Warszawa),

stolica (Polska, Warszawa)?

stolica (Polska, \$x)?

odpowiadają następującym zdaniom:

Warszawa jest stolicą Polski.

Czy Warszawa jest stolicą Polski?

Jakie miasto jest stolicą Polski?

Pytania szczegółowe, na które odpowiada się wymieniając nazwy obiektów, o które chodzi w pytaniu, są interpretowane przy pomocy formuł zawierających w miejscu odpowiedniego argumentu oznaczenie zmiennej. Oznaczenie to składa się z dowolnego słowa rozpoczynającego się od znaku \$. Formuła odpowiadająca pytaniu powinna być zawsze zakończona znakiem zapytania.

Zarówno nazwa predykatu, jak i nazwa i wartości jego argumentów, mogą być zwrotami wielowyrazowymi. Nie jest wówczas wymagane używanie specjalnych znaków do rozdzielania wyrazów wchodzących w skład zwrotu.

Wyrazy i zwroty występujące w formule w charakterze argumentów predykatu muszą znajdować swoje odpowiedniki w tekście zdania. Różnice między argumentem predykatu a odpowiednią frazą zdania mogą się przejawiać jedynie poprzez użycie różnych form gramatycznych. Ograniczenie to nie dotyczy nazw predykatów i nazw typów ich argumentów - mogą być one dobrane dowolnie. W formule interpretującej zdanie nazwy predykatów i argumenty powinny występować w formach podstawowych.

Siła semantyczna języka naturalnego jest znacznie większa niż siła semantyczna jakiegokolwiek języka formalnego. Oznacza to, że pewne zdania w języku naturalnym nie mogą być w pełni oddane przez formuły języka formalnego. Dotyczy to między innymi zdań wieloznacznych, które mogą mieć wiele różnych interpretacji, podczas gdy przyporządkowane im formuły są jednoznaczne i mają jedno, dobrze określone znaczenie. Z drugiej strony, ta sama myśl może być sformułowana w różny sposób, a między różnymi sformułowaniami tej samej idei mogą istnieć subtelne różnice semantyczne, ale system automatyczny, taki jak DABINAL, nie zawsze ma możliwość ich uchwycenia.

DABINAL jest projektowany z myślą o zastosowaniu go jako łącznika z przedmiotową bazą danych, powinien więc umieć ze zdania w języku naturalnym wybierać tylko te dane, które są interesujące z punktu widzenia tej bazy. Jeśli np. w bazie danych gromadzone są wyłącznie imiona ojców, to ze zdania "Jan jest synem Marii i Karola" DABINAL powinien wygenerować formułę OJCIEC (JAN, KAROL), tzn. całkowicie pominąć fragment zdania w języku naturalnym, nieistotny z punktu widzenia konkretnej bazy danych ("Marii i").

MODUŁ UCZĄCY SIĘ

Zdolność systemu DABINAL do radzenia sobie z wypowiedziami w języku naturalnym jest oparta na trzech rodzajach reguł decyzyjnych:

- regułach kojarzących ogólną strukturę zdania w języku naturalnym ze schematem odpowiedniej formuły języka pośrednika,
- regułach wiążących fragmenty zdania w języku naturalnym z argumentami predykatów występujących w odpowiednich formułach interpretacyjnych,
- regułach transformacji słów i zwrotów do postaci standardowej (języka pośrednika) przechowywanej w bazie wiedzy językowej.

Moduł uczący się wypracowuje wymienione trzy rodzaje reguł drogą porównywania zdań w języku naturalnym z odpowiadającymi im formułami języka pośrednika. Okazuje się, że w ten sposób można zgromadzić bardzo wiele danych językowych, nawet jeśli opieramy się tylko na pojedynczym przykładzie. Dane te można podzielić na kilka grup:

- dane słownikowe, z uwzględnieniem klasyfikacji słów i hierarchicznych powiązań między nimi,

- dane kontekstowe, dotyczące współwystępowania w tekście różnych jednostek słownikowych,
- dane strukturalne, dotyczące budowy całego zdania, porządku wyrazów itp.,
- dane morfologiczne, odnoszące się do form fleksyjnych wyrazów, i współzależności form fleksyjnych wyrazów tego samego zdania.

Znaczenie wyżej wymienionych grup informacji językowych dla analizy językowej jest różne w przypadku różnych języków naturalnych. Na przykład, w języku polskim charakterystyka morfologiczna wyrazu ma często większe znaczenie niż pozycja wyrazu w zdaniu, w przeciwieństwie do języka angielskiego (zazwyczaj sztywny szyk wyrazów w zdaniu i znacznie mniej zróżnicowane formy fleksyjne).

Moduł uczący się systemu DABINAL (w zrealizowanej wersji 1.0) dysponuje jedynie ograniczoną zdolnością do uogólniania zdobytej wiedzy językowej. Wiedza zgromadzona przez ten moduł znajduje zastosowanie przy analizie zdań mających tę samą strukturę, co zdania napotkane w materiale treningowym. Ogólniejsze reguły decyzyjne będą w kolejnych wersjach systemu wypracowywane przez procedury heurystyczne, należące do modułu uogólniania. Moduł ten będzie mógł dokonywać uogólnień faktów językowych zgromadzonych w bazie wiedzy językowej, lecz jego zadaniem nie będzie ujawniania logicznych zależności między danymi należącymi do przedmiotowej bazy danych.

MODUŁ ANALIZY

Analiza zdań języka naturalnego jest w systemie DABINAL oparta na informacjach zgromadzonych przez moduł uczący się i

przechowywanych w bazie wiedzy językowej. System nie korzysta z żadnych innych źródeł informacji językowej. Wykorzystaniem dostępnej wiedzy sterują procedury heurystyczne oparte na bardzo ogólnych założeniach. Są one używane w określonym z góry porządku: najpierw dokonuje się klasyfikacji słownikowej, następnie próbuje się rozpoznać strukturę zdania, w dalszej kolejności stosowane są reguły określające zależność między fragmentami zdania i argumentami interpretujących je predykatów, a na końcu następują przekształcenia form fleksyjnych do postaci podstawowej.

Moduł analizy jest zdolny do rozpoznawania podobieństwa fragmentów zdania w języku naturalnym i elementów wzorca zdania w bazie wiedzy językowej, nawet jeśli nie są one całkiem identyczne. Istnieje kilka typów różnic wpływających na funkcję podobieństwa, określającą stopień zgodności porównywanych fragmentów tekstów:

- różnice porządku słów,
- różnice w występowaniu słów pomocniczych,
- różnice form fleksyjnych,
- różnice w klasyfikacji poszczególnych elementów zdania.

Na obecnym etapie realizacji DABINAL jest w stanie rozpoznawać zdania mające schemat identyczny lub bardzo podobny do schematu zdań spotkanych wcześniej w materiałach treningowych. Na przykład, jeśli system nauczył się już prawidłowej interpretacji pytania "Jaki jest najwyższy szczyt w Tatrach?", będzie on również zdolny prawidłowo zinterpretować pytanie "Jaki jest najwyższy szczyt w Alpach?", ponieważ ma ono tę samą strukturę. Ale wiedza zdobyta dzięki analizie powyższego przykładu nie wystarczy do zinterpretowania pytania "Jak się nazywa najwyższy szczyt w Alpach?", ponieważ system jeszcze nie jest w stanie stwierdzić równoważności zwrotów

KORZYSTANIE Z SYSTEMU DABINAL

Oprogramowanie wersji 1.0 systemu DABINAL zawiera, oprócz programu głównego o nazwie DABINAL wykonującego podstawowe funkcje uczenia się i stosowania zgromadzonej wiedzy językowej, także dwa programy pomocnicze, ułatwiające pracę użytkownikowi. Są to:

- DBPRINT - program drukujący zawartość bazy wiedzy językowej w formacie czytelnym dla użytkownika,
- DBREST - program pozwalający ponownie wykorzystać przykłady użyte wcześniej do kształtowania innej bazy wiedzy językowej, bez konieczności powtórznego wprowadzania tych przykładów z klawiatury.

Program DABINAL jest głównym modułem systemu, realizującym funkcję uczenia się i stosującym zgromadzoną wiedzę językową. Komunikuje się on z użytkownikiem poprzez system okienek ukazujących się na ekranie monitora i znikających, jeśli nie są potrzebne. Dialog z użytkownikiem prowadzony jest po angielsku, co w żadnym stopniu nie ogranicza swobody wyboru języka przedmiotowego, będącego w danym momencie obiektem uczenia się programu.

Rysunek 2 przedstawia sytuację, w której widoczne są prawie wszystkie okna komunikacyjne wykorzystywane przez program. Okna te spełniają następujące funkcje:

- okno Database - widoczne jest na ekranie przez cały czas pracy programu i ukazuje nazwę aktualnie aktywnej bazy wiedzy językowej,
- okno Sentence - widoczne jest przez cały czas pracy programu i służy do wprowadzania zdań w języku naturalnym, przedstawianych systemowi do analizy,
- okno Formula - widoczne jest przez cały czas pracy programu

Database test	
Sentence New York is the biggest town in USA.	
Formula biggest town (USA, New York).	
Lexicon Updates Item: TOWN Basic form:	Predicate's Definition Predicate's name BIGGEST TOWN No of attributes: 2 1st attribute: country 2nd attribute: town
Messages Defining predicate. Enter no of attributes. Declaring attributes' names.	

Rys. 2. Układ okien komunikacyjnych na ekranie monitora.

i służy zarówno do wprowadzania przez użytkownika formuł interpretacyjnych, adekwatnych do zdania wyświetlanego w danym momencie w oknie Sentences, jak również do wyprowadzania formuł wygenerowanych przez system, jeśli jest on w stanie zinterpretować zdanie samodzielnie,

- okno Lexicon Updates - pojawia się na ekranie tylko w takich momentach, kiedy konieczne jest wyjaśnienie pewnych problemów słownikowych (np. określenie podstawowej formy fleksyjnej wyrazu lub zwrotu); w oknie tym pojawiają się hipotezy formułowane przez system, ale także służyć ono może do wprowadzania informacji przez użytkownika,
- okno Predicate's Definition pojawia się na ekranie tylko wówczas, gdy zachodzi konieczność zdefiniowania nowego predykatu;

służy ono użytkownikowi do wprowadzania wszystkich informacji dotyczących tego predykatu,

- okno Messages - jest widoczne przez cały czas działania programu i służy systemowi do wyświetlania komunikatów i poleceń dla użytkownika; może ono także pośredniczyć we wprowadzaniu przez użytkownika krótkich odpowiedzi na pytania systemu.

Oprócz okien widocznych na rys. 2 na początku pracy ze świeżo założoną bazą wiedzy językowej pojawić się może okno służące do definiowania przedimków i rodzajników istniejących w języku, który ma być przedmiotem uczenia się systemu.

Podczas prowadzenia dialogu z systemem trzeba przestrzegać kilku reguł:

1. W każdej sytuacji, gdy system oczekuje na odpowiedź użytkownika, naciśnięcie klawisza ENTER jest równoznaczne z odpowiedzią twierdzącą.
2. W sytuacjach, gdy system oczekuje na wprowadzenie przez użytkownika pewnej formy wyrazu lub zwrotu widocznego w nagłówku okna Lexicon Updates, naciśnięcie klawisza ENTER oznacza, że oczekiwana forma jest tożsama z formą ukazaną w nagłówku.
3. Wprowadzenie kropki jako pierwszego znaku zdania lub formuły interpretacyjnej oznacza żądanie przerwania aktualnie wykonywanej głównej pętli programu i przejście do miejsca, w którym można podjąć decyzję o jego kontynuowaniu, zakończeniu sesji lub skopiowaniu aktualnej wersji bazy wiedzy do katalogu BACKUP.
4. Wprowadzenie znaku kończącego zdanie lub formułę (kropki lub znaku zapytania) jest równoznaczne z zaakceptowaniem jej aktualnej postaci, tzn. system po przyjęciu tego znaku przechodzi do analizy wprowadzonego tekstu. Dlatego, jeśli chcemy sprawdzić i ewentualnie poprawić wprowadzany tekst, należy to uczynić przed postawieniem znaku kończącego.

5. W wersji 1.0 systemu nie istnieją narzędzia pozwalające na edycję danych w bazie wiedzy językowej. W związku z tym, jeśli na skutek pomyłki użytkownika zostaną do bazy wiedzy wprowadzone dane błędne, trudno je później skorygować. Dlatego przy wprowadzaniu tekstu z klawiatury należy być uważnym.

6. Należy przestrzegać systemowych ograniczeń dotyczących rozmiarów wprowadzonych zdań, formuł i jednostek leksykalnych, których przekroczenie może doprowadzić do wprowadzenia błędnych informacji do bazy wiedzy językowej, a nawet do naruszenia jej spójności. Ograniczenia te zostały ustalone na takim poziomie, że nie utrudniają pracy dla przewidywanych zastosowań obecnej wersji systemu, niemniej należy pamiętać o ich istnieniu. Oto ich lista:

- maksymalna długość zdania i formuły: 200 znaków,
- maksymalna liczba elementów zdania (słów i znaków przestankowych): 15,
- maksymalna długość słowa: 40 znaków,
- maksymalna długość frazy: 50 znaków,
- maksymalna liczba elementów frazy: 5,
- maksymalna liczba argumentów predykatu: 5.

Tuż po rozpoczęciu działania programu w oknie komunikatów (Messages) pojawia się żądanie wprowadzenia nazwy bazy wiedzy językowej, z którą zamierzamy współpracować. Cursor przenosi się do okienka Database i tam oczekuje na wprowadzenie odpowiedniej nazwy. Naciśnięcie w tym momencie klawisza ENTER oznacza zakończenie działania programu.

Po wprowadzeniu nazwy bazy wiedzy językowej program odczytuje z niej informacje niezbędne do dalszego działania. Jeśli rozpoczynamy pracę z nową bazą wiedzy, w oknie komunikatów pojawia się

pytanie, czy język, którego system będzie się uczył, zawiera przedimki lub rodzajniki. Jeśli tak, na ekranie ukazuje się okienko, w którym należy zdefiniować wszystkie rodzajniki i przedimki, uwzględniając ich pogrupowanie. Instrukcje wyświetlane w oknie komunikatów opisują sposób wykonania tej operacji.

Po opisanych wyżej wstępnych czynnościach program wchodzi w swoją główną pętlę, na którą się składają następujące fazy:

1. Wprowadzenie przez użytkownika, za pośrednictwem okna Sentence, zdania, które ma zostać przeanalizowane,
2. Próba analizy wprowadzonego zdania przez system. W tej fazie program może zadawać użytkownikowi pytania uzupełniające, dotyczące form podstawowych wyrazów i zwrotów występujących w zdaniu, a także ich klasyfikacji do różnych kategorii słownikowych. Dialog jest prowadzony za pośrednictwem okien Messages i Lexicon Updates.
3. Jeśli w fazie 2 systemowi udało się wygenerować formułę interpretującą wprowadzone zdanie, formuła ta zostaje wyświetlona w oknie Formula z prośbą o jej akceptację lub odrzucenie. Jeżeli dotychczasowa wiedza systemu nie pozwoliła mu na prawidłowe zinterpretowanie zdania, użytkownik jest proszony o podanie swojej interpretacji.
4. Faza ta ma miejsce tylko wtedy, gdy w fazie 3 formuła interpretująca została podana przez użytkownika. Następuje wówczas analiza tej formuły przez system. W trakcie tej analizy, podobnie jak w fazie 2, może wywiązać się dialog pomiędzy użytkownikiem i programem, dotyczący elementów słownika. Dodatkowo, jeżeli w formule interpretującej został użyty predykat dotychczas nieznan systemowi, na ekranie pojawia się okno Predicate's De-

inition, przez które należy wprowadzić definicję nowego predykatu, zgodnie z instrukcjami pojawiającymi się w oknie Messages.

5. Punkt węzłowy pętli, w którym użytkownik może zdecydować o kontynuowaniu programu i przejściu do fazy 1, o zakończeniu działania programu lub o skopiowaniu aktualnej wersji bazy wiedzy językowej do katalogu BACKUP. W przypadku podjęcia decyzji o sporządzaniu kopii, program, po jej wykonaniu, przechodzi automatycznie do fazy 1.

Na zakończenie warto poświęcić kilka uwag ograniczeniom nałożonym na zdania poddawane analizie przez DABINAL.

W tekstach wprowadzanych zdań dopuszczalne jest użycie: małych i dużych liter alfabetu łacińskiego i znaków diakrytycznych alfabetu polskiego, cyfr, spacji, znaków przestankowych: podkreślenia, myślnika, apostrofu i przecinka oraz znaków kończących: kropki i znaku zapytania. Poszczególne wyrazy nie mogą być dłuższe niż 40 znaków, a całe zdanie nie powinno zawierać więcej niż 15 elementów i 200 znaków. Jeśli znak apostrof, myślnik lub podkreślenie występują bezpośrednio po znaku różnym od spacji lub przecinka, to uważa się je za część poprzedzającej jednostki leksykalnej. Delimitatorem jednostki leksykalnej (słowa) jest spacja, przecinek lub znak kończący zdanie.

Poza wspomnianym wyżej ograniczeniem liczby elementów, nie narzuca się ograniczeń na złożoność struktury zdania. W szczególności, może być to zdanie współrzędnie lub podrzędnie złożone.

Zarówno w zdaniach, jak i w formułach interpretujących i w jednostkach słownikowych nie robi się rozróżnienia pomiędzy dużymi i małymi literami. W związku z tym dopuszczalne jest rozpoczynanie zdania od małej litery lub pisanie małą literą nazw własnych.

REZULTATY DOTYCHCZASOWYCH PRAC NAD SYSTEMEM DABINAL

Punktem wyjścia do realizacji wersji 1.0 systemu DABINAL była jego wersja prototypowa, przygotowana w roku 1990. W wersji 1.0 zmieniono struktury danych opisujące reguły zamiany końcówek i procedury posługujące się tymi strukturami, udoskonalono metody analizy zwrotów wielowyrazowych i wydzielono przedimki i rodzajniki jako osobną kategorię słów, traktowaną w szczególny sposób. Dzięki intensywnym testom udało się wykryć i poprawić wiele błędów istniejących w wersji prototypowej systemu.

Zmiana struktur danych opisujących reguły zamiany końcówek w różnych formach fleksyjnych wyrazów umożliwiła osiągnięcie większej zwartości zapisu tych reguł. Udało się także połączyć w grupy podobne reguły, zwiększając tym samym skuteczność ich działania.

Udoskonalone algorytmy analizy zwrotów wielowyrazowych pozwalają m.in. na prawidłowe rozpoznawanie takich zwrotów w sytuacjach, gdy w zdaniu sąsiadują one ze sobą bezpośrednio. W wersji prototypowej, z dwóch sąsiadujących ze sobą zwrotów, nieznanych wcześniej systemowi, tylko jeden mógł być wielowyrazowy - w przeciwnym przypadku analiza zdania prowadziła do błędnych wniosków.

Wersja prototypowa systemu DABINAL była tworzona z myślą głównie o języku polskim. Dlatego nie przewidziano w niej specjalnych środków, uwzględniających istnienie przedimków i rodzajników, które występują np. w językach germańskich i romańskich. Rodzajniki i przedimki były przez wersję prototypową DABINAL-a traktowane jako równoprawne elementy zwrotu wielowyrazowego. Doświadczenia z udziałem zdań w języku niemieckim i angielskim przeprowadzone na prototypowej wersji systemu wykazały, że traktowanie przed-

inition, przez które należy wprowadzić definicję nowego predykatu, zgodnie z instrukcjami pojawiającymi się w oknie Messages.

5. Punkt węzłowy pętli, w którym użytkownik może zdecydować o kontynuowaniu programu i przejściu do fazy 1, o zakończeniu działania programu lub o skopiowaniu aktualnej wersji bazy wiedzy językowej do katalogu BACKUP. W przypadku podjęcia decyzji o sporządzaniu kopii, program, po jej wykonaniu, przechodzi automatycznie do fazy 1.

Na zakończenie warto poświęcić kilka uwag ograniczeniom nałożonym na zdania poddawane analizie przez DABINAL.

W tekstach wprowadzanych zdań dopuszczalne jest użycie: małych i dużych liter alfabetu łacińskiego i znaków diakrytycznych alfabetu polskiego, cyfr, spacji, znaków przestankowych: podkreślenia, myślnika, apostrofu i przecinka oraz znaków kończących: kropki i znaku zapytania. Poszczególne wyrazy nie mogą być dłuższe niż 40 znaków, a całe zdanie nie powinno zawierać więcej niż 15 elementów i 200 znaków. Jeśli znak apostrof, myślnik lub podkreślenie występują bezpośrednio po znaku różnym od spacji lub przecinka, to uważa się je za część poprzedzającej jednostki leksykalnej. Delimitatorem jednostki leksykalnej (słowa) jest spacja, przecinek lub znak kończący zdanie.

Poza wspomnianym wyżej ograniczeniem liczby elementów, nie narzuca się ograniczeń na złożoność struktury zdania. W szczególności, może być to zdanie współrzędnie lub podrzędnie złożone.

Zarówno w zdaniach, jak i w formułach interpretujących i w jednostkach słownikowych nie robi się rozróżnienia pomiędzy dużymi i małymi literami. W związku z tym dopuszczalne jest rozpoczynanie zdania od małej litery lub pisanie małą literą nazw własnych.

REZULTATY DOTYCHCZASOWYCH PRAC NAD SYSTEMEM DABINAL

Punktem wyjścia do realizacji wersji 1.0 systemu DABINAL była jego wersja prototypowa, przygotowana w roku 1990. W wersji 1.0 zmieniono struktury danych opisujące reguły zamiany końcówek i procedury posługujące się tymi strukturami, udoskonalono metody analizy zwrotów wielowyrazowych i wydzielono przedimki i rodzajniki jako osobną kategorię słów, traktowaną w szczególny sposób. Dzięki intensywnym testom udało się wykryć i poprawić wiele błędów istniejących w wersji prototypowej systemu.

Zmiana struktur danych opisujących reguły zamiany końcówek w różnych formach fleksyjnych wyrazów umożliwiła osiągnięcie większej zwartości zapisu tych reguł. Udało się także połączyć w grupy podobne reguły, zwiększając tym samym skuteczność ich działania.

Udoskonalone algorytmy analizy zwrotów wielowyrazowych pozwalają m.in. na prawidłowe rozpoznawanie takich zwrotów w sytuacjach, gdy w zdaniu sąsiadują one ze sobą bezpośrednio. W wersji prototypowej, z dwóch sąsiadujących ze sobą zwrotów, nieznanych wcześniej systemowi, tylko jeden mógł być wielowyrazowy - w przeciwnym przypadku analiza zdania prowadziła do błędnych wniosków.

Wersja prototypowa systemu DABINAL była tworzona z myślą głównie o języku polskim. Dlatego nie przewidziano w niej specjalnych środków, uwzględniających istnienie przedimków i rodzajników, które występują np. w językach germańskich i romańskich. Rodzajniki i przedimki były przez wersję prototypową DABINAL-a traktowane jako równoprawne elementy zwrotu wielowyrazowego. Doświadczenia z udziałem zdań w języku niemieckim i angielskim przeprowadzone na prototypowej wersji systemu wykazały, że traktowanie przed-

imków i rodzajników w taki sam sposób, jak wyrazów innych kategorii, a co za tym idzie, traktowanie wyrazu poprzedzonego przedimkiem lub rodzajnikiem jako pełnoprawnego zwrotu wielowyrazowego, często prowadzi do sytuacji, w których system działa mniej sprawnie niż powinien. Wyraz poprzedzony przedimkiem i wyraz bez przedimka są wówczas uważane za dwie osobne jednostki leksykalne, których system musi się nauczyć na odrębnych przykładach. Dlatego w trakcie realizacji wersji 1.0 zdecydowano się na wydzielenie przedimków i rodzajników jako osobnej kategorii wyrazów, spełniających w zdaniu specjalne role. Opracowując sposoby postępowania z wyrazami tej kategorii, brano pod uwagę ich cechy wspólne dla języka angielskiego, niemieckiego i francuskiego.

Równoległe z pracami programistycznymi nad wersją 1.0 systemu DABINAL przeprowadzano eksperymenty z udziałem wersji prototypowej i z pojawiającymi się w ciągu roku 1991 kolejnymi wersjami poprawionymi i udoskonalonymi. Celem tych eksperymentów było:

- wykrycie błędów programowania istniejących w wersji prototypowej,
- ocena sprawności działania systemu i sformułowanie na jej podstawie celów doraźnych i dalekosiężnych dalszych prac nad projektem DABINAL,
- ocena systemu pod względem wygody posługiwania się nim i określenie zmian i rozszerzeń poprawiających jego cechy użytkowe.

Wiele wyników eksperymentów zostało uwzględnionych przy realizacji wersji 1.0. Przede wszystkim oczyszczono system z zauważonych błędów. Dodano też kilka funkcji pomocniczych, ułatwiających działanie użytkownika (np. programy DBPRINT, DBREST, funkcja "backup" wywoływana z programu DABINAL). Wyniki doświadczeń zaważyły także na decyzji wydzielenia przedimków i rodzajników jako osobnej

grupy wyrazów rozpoznawanej przez system (o czym wspomnieliśmy już wyżej).

W trakcie eksperymentów szczególną uwagę zwracano na badanie skuteczności działania heurystyk trzech rodzajów:

- heurystyk rozpoznających schemat zdania i kojarzących go z odpowiednim schematem formuły,
- heurystyk kojarzących fragmenty zdania z odpowiednimi elementami przyporządkowanej mu formuły,
- heurystyk dokonujących transformacji do postaci podstawowej różnych form fleksyjnych zwrotów występujących w przykładach.

Eksperymenty wykazały prawidłowe działanie heurystyk z pierwszej i drugiej grupy. Poza wyjątkowymi, szczególnie dobranymi przykładami, system bezbłędnie rozpoznawał schematy zdaniowe znane z wcześniej przedstawionych mu przykładów.

W przypadku, kiedy schemat zdania został rozpoznany prawidłowo, kojarzenie fragmentów zdania z odpowiednimi elementami formuły interpretacyjnej nie nastęrczało w przeprowadzonych eksperymentach żadnej trudności. W żadnym z przebadanych przykładów nie natrafiono na błędne działanie heurystyk kojarzących fragmenty zdania z elementami formuły interpretacyjnej.

Najmniej skuteczne spośród heurystyk nadzorujących działanie systemu okazują się heurystyki sprowadzania różnych form fleksyjnych wyrazów i zwrotów do formy podstawowej. Trudności te były zresztą do przewidzenia, ponieważ wynikają one z samej natury reguł rządzących tworzeniem form fleksyjnych. Są one często niejednoznaczne, od ich obowiązywania istnieją liczne wyjątki, a w dodatku jest ich bardzo dużo. Wskutek tego skuteczność stosowania tych reguł w początkowej fazie uczenia się systemu jest stosunkowo niska, wzrasta jednak w miarę powiększania się zgromadzonych zasobów lek-

svkalnych. Niemniej, wydaje się, że poprawa jakości działania procedur systemu nadzorujących transformacje fleksyjne jest możliwa i powinna być jednym z głównych celów prac nad rozwojem DABINAL-a w przyszłości. W szczególności dotyczy to przekształceń form fleksyjnych całych zwrotów. Procedury dokonujące transformacji formy fleksyjnej zwrotu są w obecnej wersji systemu DABINAL ekstrakcją procedur transformacji dla pojedynczych słów i tylko w minimalnym stopniu uwzględniają budowę gramatyczną rozpatrywanego zwrotu.

Na zakończenie warto zauważyć, że pracom nad systemem DABINAL przyświecają dwa cele. Po pierwsze, dążymy do opracowania praktycznego narzędzia, umożliwiającego szybkie uczenie systemów obsługujących bazy danych "rozumienia" wprowadzanych danych i pytań formułowanych w języku naturalnym. Po drugie, mamy także nadzieję na uzyskanie dzięki DABINAL-owi ciekawego materiału do prac teoretycznych. Sądzymy, że interesujące będzie porównanie sposobu tworzenia i porządkowania wiedzy językowej wypracowanego przez DABINAL ze sposobami stosowanymi w tradycyjnym językoznawstwie. Być może antropocentryzm metodologii językoznawstwa stanowi jedno ze znaczących utrudnień w wykorzystaniu przez systemy komputerowe przygotowanej przez człowieka wiedzy o języku?

Literatura

1. ALLEN J. Natural language understanding. Menlo Park, CA: Benjamin/Cummings 1987 XV, 574 s. (Benjamin/Cummings Series in Computer Science).
2. BATES M., BOBROW R.J. Natural language interfaces: what's here,

- what's coming, and who needs it. W: Artificial intelligence applications for business. Norwood, NJ: Ablex 1984 s. 179-194.
3. BECKER O. Automated language processing. W: Annual Review of Information Science and Technology vol.16 1981 s.113-138.
 4. BOBROW D.G. et al. GUS: a frame driven dialogue system. Artificial Intelligence 1977 nr 2 s.155-173.
 5. CARBONELL J.G., BOGGS W.M., MAULDIN M.L. The XALIBUR project: a natural language interface to expert systems. Proc.8th IJCAI, Karlsruhe 1983 s.653-656.
 6. COMPUTATIONAL models of natural language processing. /Bara B.G., Guida G. (eds.). Amsterdam: North-Holland 1984 327 s. (Fundamental Studies in Computer Science; vol.9).
 7. DOSZKOCS T.E. Natural language processing in information retrieval. J.ASIS 1986 nr 4 s.191-196.
 8. DRAKIN V.I., POPOV E.V., PREOBRAŽENSKIJ A.V. Obščeniye konečnyh polzovatelej s sistemami obrabotki dannyh. Moskva: Radio i svjaz' 1988 288 s.
 9. FORSYTH R., RADA R. Machine learning: applications in expert systems and information retrieval. Chichester: Ellis Horwood and New York; etc.: Halsted Press 1986 277 s. (Ellis Horwood Series in Artificial Intelligence).
 10. GROSZ B.J. et al. TEAM: an experiment in the design of transportable natural-language interfaces. Artificial Intelligence 1987 nr 2 s. 173-243.
 11. HAAS N., HENDRIX G.G. Learning by being told: acquiring knowledge for information management. W: /23/ s. 405-427.
 12. THE HANDBOOK of Artificial Intelligence: vol.1. /Barr A., Feigenbaum E.A. (eds). Stanford, CA: HeurisTech Press 1981 409 s.
 13. HARRIS L.R. User oriented data base query with the ROBOT natural language query system. Int. J. Man-Machine Studies 1977 vol.9 s.697-713.
 14. HENDRIX G.G., WALTER B.A. The intelligent assistant: technical considerations involved in designing Q₂A's natural-language interface. Byte 1987 nr 14 s.251-258.

15. ISKUSSTVENNYJ intellekt: v 3 kn. Kn. 1. Sistemy obščeniĵa i ekspertnye sistemy: spravočnik. Popov E.V. (red.) Moskva: Radio i svjaz' 1990 464 s.
16. ISKUSSTVENNYJ intellekt: v 3 kn. Kn. 2. Modeli i metody: spravočnik. /Pospelov D.A. (red.) Moskva: Radio i svjaz' 1990 304 s.
17. ISKUSSTVENNYJ intellekt: v 3 kn. Kn. 3. Programmnye i apparatnye sredstva: spravočnik. /Zacharov V.N., Choroševskij V.F. (red.) Moskva: Radio i svjaz' 1990 368 s.
18. JACOBS P.S. PHRED: a generator for natural language interfaces. Computational Linguistics 1985 nr 4 s. 219-242.
19. KRAGELOH K.-O., LOCKEMANN P.C. Access to data base systems via natural language. W:/26/s. 49-86.
20. MARTIROSIAN A.A., POGOSIAN E.M.: Obučeniĵe. W: /17/ s. 206-231
21. MICHALSKI R.S. O naturze uczenia się - problemy i kierunki badawcze (I). Informatyka 1988 nr 2 s.4-7.
22. MICHALSKI R.S. O naturze uczenia się - problemy i kierunki badawcze (II). Informatyka 1988 nr 3 s.8-11.
23. MACHINE learning: an artificial intelligence approach. /Michalski R.S., Carbonell J.G., Mitchell T.M. (eds.) Berlin; etc.: Springer 1984 XI, 572 s.
24. MACHINE learning: an artificial intelligence approach. Vol.II /Michalski R.S., Carbonell J.G., Mitchell T.M. (eds.) Los Altos, CA: Morgan Kaufmann Publishers 1986 X, 738 s.
25. MOYNE J.A. Understanding language: man or machine. New York; London: Plenum Press 1985 XVI, 357 s. (Foundations of Computer Science).
26. NATURAL language communication with computers. /Bolc L. (ed.) Berlin; etc.: Springer 1978 VI, 292 s. (Lecture Notes in Control and Information Sciences; 63).
27. OSTROWSKI M., POPOWSKA H., SOLAK J. DABINAL: a natural language learning interface to databases. W: Międzynarodnyj seminar po mašinomu perevodu, Tbilisi 1989: Tezisy dokladov. Moskva: Vsesojuznyj Centr Perevodov 1989 s.230-233.

28. OSTROWSKI M., POPOWSKA H., SOLAK J. Sprawozdanie z realizacji zadania: Opracowanie wersji 1.0 systemu DABINAL uczącego się komunikacji z bazami wiedzy w języku naturalnym. Warszawa: IINTE 1991 47 s. (maszynopis).
29. POPOV E.V. Obsčeniye c EVM na estestvennom jazyke. Moskva: Nauka 1982 360 s.
30. POPOV E.V., PREOBRAZENSKIJ A.B. Osobennosti realizacii EJa-sistem. W: /15/ s.9-32.
31. PREOBRAZENSKIJ A.B. Sostojanije razvitija sovremennyh EJa-sistem. W: /15/ s.32-64.
32. SIMON H.A. Why should machines learn? W: /23/ s.25-37.
33. SPARCK JONES K., BOGURAEV B.K. A natural language front end to data bases with evaluative feedback. W: New applications of data bases. Gardarin G., Genelobe E. (eds.) London: Academic Press 1984 s. 159-182.
34. WAHLSTER W. The role of natural language in advanced knowledge-based systems. W: Artificial Intelligence and Man-Machine Systems: Proc. of an Int. Seminar organized by Deutsche Forschungs- und Versuchsanstalt für Luft- und raumfahrt (DFVLR) Bonn, Germany, May 1986. Berlin; etc.: Springer 1986 s. 62-83. (Lecture Notes in Control and Information Sciences; 80).
35. WALLACE M. Communicating with databases in natural language. Chichester: Ellis Horwood and New York; etc.: Halsted Press 1984 170 s. (Ellis Horwood Series in Artificial Intelligence).
36. WARNER A.J. Natural language processing. W: Annual Review of Information Science and Technology vol.22 1987 s.79-108.
37. WEBBER B.L. Natural language processing: a survey. W: On knowledge base management systems: integrating artificial intelligence and database technologies. /Brodie M.L., Mylopoulos J. (eds.) New York; etc.: Springer 1986 s. 353-363. ((Topics in Information Systems)).

LEARNING NATURAL LANGUAGE BY THE COMPUTER SYSTEM
THE CASE OF "DABINAL" SYSTEM

Summary

The article is devoted to the description of a project the aim of which is to establish the system able to acquire the natural language to the extent necessary for communicating with databases. The computer system which comes into being as the result of realization of the project is called DABINAL (DataBase Interface in Natural Language). The system belongs to the group of mobile systems (transportable ones) i.e. it is not designed for one particular database but should provide possibility of communicating with different databases. All linguistic knowledge necessary for its functioning the system acquires and arranges itself - through out learning. DABINAL is an example of the application of machine learning methods to the construction of natural language database interfaces. DABINAL is designed in the Institute for Scientific, Technical and Economic Information. The first 1.0 version of the system was made in 1991.

The article is composed of two parts. The first part deals with selected issues of artificial intelligence: systems of communicating in natural language with databases and learning by machines. The second part is devoted to the general outline of the system: to the description of its structure, to the examples of learning, to the presentation of the base of linguistic knowledge as well as to the characteristics of the main modules of the system: of the learning module, of the module of analysis and of the module of generalization. To end with there is shown the

way of using the system and the results of the work which has been done to the present.

ОБУЧЕНИЕ ЭВМ ЕСТЕСТВЕННОМУ ЯЗЫКУ

(На примере системы DABINAL)

Р е з ю м е

В статье представлено описание проекта, целью которого является создание системы способной учиться естественному языку в пределах необходимых для общения с базами данных. Автоматизированная система возникающая как результат реализации этого проекта называется DABINAL (DataBase Interface in Natural Language). Система принадлежит к группе мобильных (передвижных) систем, т.е. она не проектируется для одной, конкретной базы данных, но должна предоставлять возможность общения с разными базами. Все языковые знания необходимые для ее действия система накапливает и систематизирует самостоятельно - путем самообучения. DABINAL разрабатывается в Институте научной, технической и экономической информации. В 1991 году была создана версия 1.0 системы.

Статья состоит из двух частей. Первая часть относится к избранным проблемам искусственного интеллекта : к системам общения на естественном языке с базами данных и к обучению машин. Вторая часть посвящена описанию системы DABINAL. Представлены : общая характеристика системы, ее структура, примеры обучения, база языковых знаний, язык посредник и основные модули системы : обучающийся модуль, модуль анализа и обобщающий модуль. Обсужден способ пользования системой и достигнутые результаты работ.

MARIA JANUSZEWSKA

SYSTEMY HIPERTEKSTOWE - RYS HISTORYCZNY

Pierwsze koncepcje hipertekstu i próby ich realizacji. Dalszy rozwój hipertekstu.

Wprowadzenie

Użytkownicy współczesnych systemów komputerowych nie chcą już pełnić roli biernego odbiorcy informacji zgromadzonej i wybranej przez system. Pragną pełnić rolę aktywną i mieć bezpośredni wpływ na tworzenie informacyjnego zbioru systemu oraz wybór i modyfikację strategii wyszukiwawczych.

Pojawia się więc zapotrzebowanie na nowe narzędzie umożliwiające nie tylko szybki, bezpośredni, odpowiednio zróżnicowany dostęp do wybranych dokumentów czy też ich fragmentów, ale również łączenie dokumentów publikowanych z osobistymi notatkami, uwagami czy zdjęciami, a więc tworzenie własnej bazy danych będących osobistą encyklopedią wiedzy dostosowaną do potrzeb danego użytkownika. Tym narzędziem może być właśnie hipertekst.

Większość obecnie istniejących dokumentów charakteryzuje linearność ich struktury, zarówno fizycznej jak i logicznej.

W aspekcie fizycznym dokument może być rozpatrywany jako łańcuch wyrazów tworzących linie strony etc., w aspekcie logicznym wyrazy są elementami nadrzędnych jednostek - zdań, te zaś składają się na akapity, z których składa się pełny tekst dokumentu. Większość systemów komputerowych, m.in. systemy baz danych, cechuje również podobna budowa. Istnieją jednak dokumenty, które mają o wiele bogatszą i bardziej złożoną budowę - tekst nie stanowi w nich jednolitego ciągu słów, bo obok tekstu głównego występuje wiele przypisów, przykładów, odesłań itp. Liczba tego rodzaju materiałów ciągle wzrasta. Właśnie ich budowa stała się podstawą koncepcji hipertekstu.

Koncepcja hipertekstu nie jest nowa, ma już dość długą, bogatą historię, a sam termin hipertekst używany jest w różnych znaczeniach. Mówi się nawet o hipertekstach manualnych, podając jako przykład karty katalogowe powiązane odsyłaczami, a więc tworzące pewnego rodzaju strukturę, czy też encyklopedie zawierające odesłania do innych haseł. Tego rodzaju dokumenty mają już podstawową cechę hipertekstu - jest nią nielinearność.

Pierwsze koncepcje hipertekstu i próby ich realizacji

Po raz pierwszy problemem nielinearności zajął się Vannevar Bush, publikując w 1945 r. artykuł "As We May Think" /6/. Bush uchodzi za twórcę idei hipertekstu, aczkolwiek on sam tej nazwy nie używał. W swojej pracy przedstawił projekt maszyny nazwanej "Memex". "Memex" miał postać biurka wyposażonego w ekran, na którym wyświetlano potrzebne informacje. Urządzenie uzupełniała klawiatura oraz zestaw przycisków i dźwigni służących do sterowania. "Memex" miał zawierać zarówno wielką bibliotekę, jak i osobiste notatki, fotografie itp.; umożliwiał on tworzenie połączeń

między dwoma dowolnymi fragmentami informacji z całej biblioteki. Bush swoją konstrukcję oparł na mikrofilmach i fotokomórkach oraz na mechanicznym urządzeniu, za pomocą którego uzyskiwało się potrzebne dokumenty. Wszystko to przesądziło o długim okresie oczekiwania na potrzebne informacje i nie zachęciło do korzystania z takiego systemu. Do praktycznej realizacji tego nowatorskiego pomysłu zabrakło wówczas wsparcia technicznego, które mogły zapewnić jedynie szybkie komputery. Podstawową cechą systemu "Memex" Busha, różniącą go od innych systemów tego okresu, stanowiła możliwość kojarzenia ze sobą dwóch dowolnych fragmentów bazy. Tę właśnie cechę przejął hipertekst. Koncepcja Busha oparta została na strukturze ludzkiej pamięci:

"Ludzka pamięć operuje skojarzeniami. Jedno uchwycone zdarzenie jest natychmiast w wyniku myślenia kojarzone z następnym, kojarzenie to odbywa się zgodnie ze skomplikowaną siateczką powiązań zawartą w komórkach mózgu" /6/.

Dwadzieścia lat później ideą Busha zainteresował się Douglas Engelbart.

Engelbart stworzył system H-LAM/T (Human Using Language, Artifacts and Methodology in which he is Trained), w którym użytkownik jest jednym z podstawowych elementów systemu, gdyż "Użytkownik i komputer tworzą symbiozę dynamicznie zmieniających się składników, co daje efekt naturalnego rozszerzenia inteligencji użytkownika" /25/. W 1968 H-LAM/T został przekształcony w NLS (oN Line System) /10/, będący narzędziem wspomagającym osoby zajmujące się pracą twórczą. Za pomocą NLS mogły one tworzyć system informacji odpowiadający ich potrzebom, umożliwiając

"umieszczanie w komputerze magazynu wykazów, planów, projektów, programów, dokumentacji, raportów, notatek, bibliografii, odesłań, wykonywanie pracy polegającej na planowaniu, projektowaniu oraz komunikowaniu się z komputerem za pomocą wygodnego urządzenia dodatkowego, jakim jest np.: "mysz" (mysz jest jednym z najbardziej znanych wynalazków Engelbarta).

Bazę danych NLS tworzyły pliki o hierarchicznej strukturze. Każdy plik dzielił się na segmenty (statements) o maksymalnej długości 3000 znaków, a każdy segment miał identyfikator poziomu w obrębie pliku. Segmenty stanowiły odpowiedniki podzbiorów wyodrębnionej informacji, plik zaś odpowiadał zwykle pełnemu dokumentowi. W obrębie plików lub między nimi mogła być tworzona dowolna liczba połączeń.

Taka hierarchiczna struktura pozwalała tworzyć nielinearne połączenia. Przy przeglądaniu plików NLS przewidywał stosowanie filtrów, dzięki którym można było:

- wybrać poziom głębokości informacji,
- obciąć liczbę fragmentów wyświetlanej informacji na każdym poziomie,
- zanotować filtry, które wyświetlały segmenty o określonej zawartości (zob. przykład 1).

NSL umożliwiał wielu użytkownikom równoczesny dostęp do systemu. System ten, ulepszany przez wiele lat, funkcjonuje obecnie pod nazwą NLS/Augument.

Zarówno Bush, jak i Engelbart nie posługiwali się terminem hipertekst. Twórcą tego terminu był Teodor Nelson. W latach sześćdziesiątych stworzył system "Xanadu" /15/.

1. Wprowadzenie.
 - 1.1. Prawo spadkowe.
 - 1.1.1. Przedmiot prawa spadkowego.
 - 1.1.2. Funkcja prawa spadkowego.
 - 1.1.3. Ogólna charakterystyka prawa spadkowego.
 - 1.2. Spadek.
 - 1.2.1. Pojęcie spadku.
 - 1.2.2. Prawa podmiotowe.
 - 1.3. Ogólna charakterystyka dziedziczenia.
2. Dziedziczenie ustawowe.
 - 2.1. Ogólne zasady dziedziczenia.
 - 2.1.1. Dziedziczenie skarbu państwa.
3. Dziedziczenie testamentowe.

a) Pełny dokument, czyli plik, z hierarchicznie uporządkowanymi segmentami (elementy ponumerowane).

1. Wprowadzenie.
 - 1.1. Prawo spadkowe.
 - 1.2. Spadek.
 - 1.3. Ogólna charakterystyka dziedziczenia.
2. Dziedziczenie ustawowe.
 - 2.1. Ogólne zasady dziedziczenia.
3. Dziedziczenie testamentowe.

b) Postać dokumentu ^{a)} po wykonaniu polecenia:
głębokość wyświetlanej hierarchii do 0.0.

1. Wprowadzenie.
 - 1.1. Prawo spadkowe.
 - 1.2. Spadek.
2. Dziedziczenie ustawowe.
 - 2.1. Ogólne zasady dziedziczenia.
3. Dziedziczenie testamentowe.

c) Postać dokumentu a) po wykonaniu polecenia:
ograniczenie segmentów na poziomie 0.0 do dwóch.

Przykład 1.

Ideą Nelsona było utworzenie uniwersalnej biblioteki online, gromadzącej ogólnoswiatowe zasoby dokumentów. Ówczesnie pomysł ten wydawał się szalenie nowatorski, wręcz fantastyczny. Nelson przepowiedział, że stworzenie bibliotek online ukształtuje nowe warunki handlu informacją - a przez to nową erę informacji.

Dalszy rozwój hipertekstu

Już pierwsze próby realizacji systemów hipertekstowych wykazały, jak trudno jest określić podstawowe cechy tego rodzaju systemów. Stworzono wiele systemów, które uchodziły za hipertekstowe, ale brakowało i brakuje nadal jednoznacznej definicji terminu hipertekst.

Próbowano zdefiniować hipertekst poprzez wskazanie, czym on nie jest. Zdaniem Conklina nie są systemami hipertekstowymi, mimo pewnych cech wspólnych:

1. Systemy okien (Window systems) - ze względu na brak bazy danych łączącej punkty węzłowe na ekranie z obiektami w bazie danych.
2. Systemy plików (File systems). Można w nich prowadzić poszukiwania kartotek wywołując nazwę kartoteki, żeby jednak mówić o systemie hipertekstowym łączenie powinno być bardziej wyrażone.

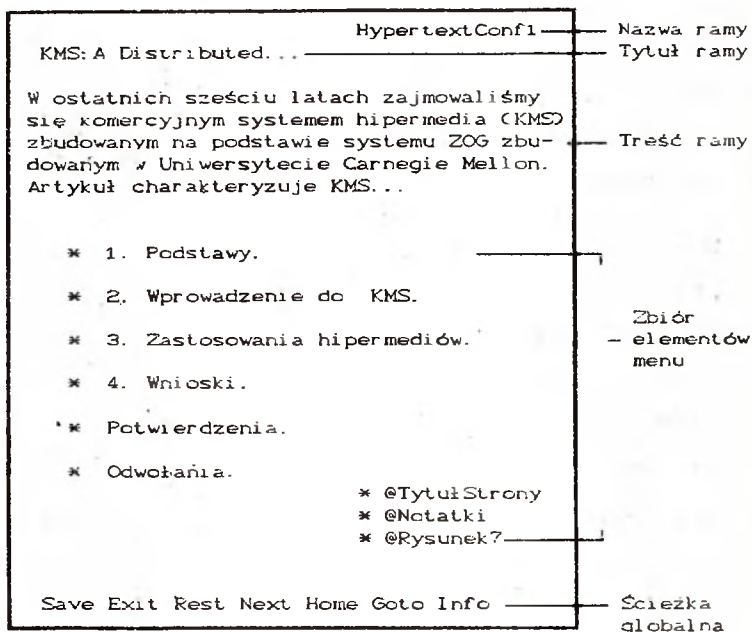
Conklin wskazał też cztery drogi rozwoju systemów hipertekstowych poprzez podanie czterech głównych obszarów zastosowań /B/.

1. Wielkie zbiory dokumentów (Macro Literary Systems), na przykład wielkie biblioteki, w których wszelkie operacje na dokumentach powinny być dokonywane przez komputer. Systemy te są charakterystyczne dla początków hipertekstu. Koncentrowano się wówczas głównie na odpowiedniej organizacji olbrzymiej ilości informacji w prostej i logicznej

przestrzeni, umożliwiającej sprawne nią operowanie. Wynalezienie komputera umożliwiło w jakimś stopniu realizację pomysłu, którego ostatecznym celem miało być połączenie wszystkich bibliotek w trybie online i łatwy dostęp do zawartych w nich informacji (np. NLS Engelbarta).

2. Komputerowe narzędzia (Problem exploration tools) do wspierania wczesnych faz rozwiązywania problemów, w których pojawia się wiele możliwych sposobów rozwiązań. Takim systemem był PIE /14/ - Personal Information Environment, który miał ułatwić programistom rozwiązywanie problemów z różnych punktów widzenia, z odmiennych perspektyw uwarunkowanych zastosowaniami.

3. Wielkie systemy służące do selektywnego przeglądania informacji (Browsing systems), np. system informacji handlowej, kolejowej, itp. Do omawianej grupy należał jeden z pierwszych systemów hipertekstowych ZOG /2, 28/ stworzony w 1972 roku w Carnegie-Mellon University. Składał się on z dużej bazy danych zawierającej segmenty o wielkości ekranu. ZOG umożliwiał jednoczesną pracę wielu użytkowników i miał krótki czas reakcji. W 1989 roku przekształcono ZOG w komercyjny system KMS (knowledge management System) /1/. (patrz Rys. 1). W systemie tym każdy segment, nazwany ramą (frame), ma u góry ekranu jednoliniowy tytuł (jego nazwę własną) . następnie kilka linii tekstu przedstawiających treść lub temat segmentu, a u dołu zbiór elementów menu (nazwanych tutaj selection) oraz linię standardowych komend przejętych z ZOG, nazwanych ścieżką globalną (global path np edit, help, back, next, mark etc). Za pomocą menu można łączyć segmenty poprzez wypisywanie numeru lub nazwy segmentu, a wtedy żądany segment pojawia się na ekranie, zastępując poprzedni. KMS ma strukturę



Rys. 1 KMS - wygląd ramy.

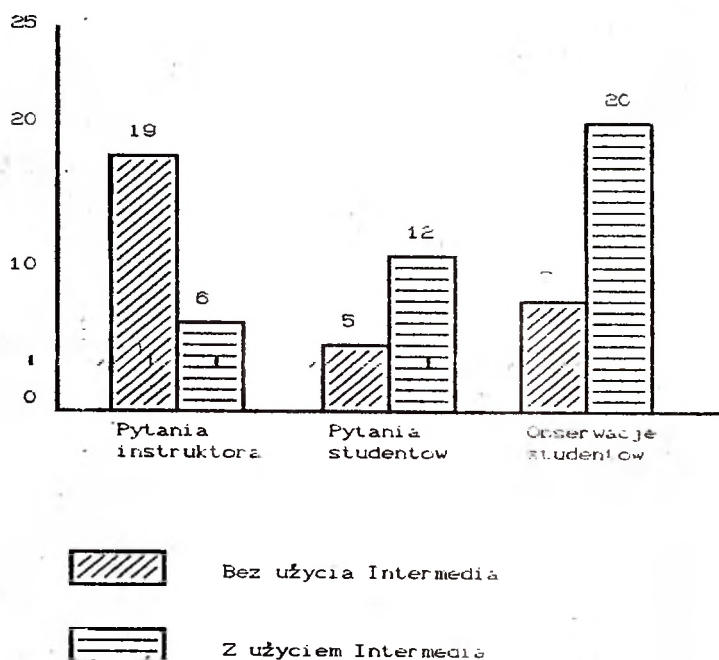
hierarchiczną, aczkolwiek mogą istnieć połączenia między segmentami na różnym poziomie.

4. Doskonalenie technologii hipertekstowej (General hypertext technology). Głównym celem jest tu eksperymentowanie z możliwymi zastosowaniami hipertekstu. Przykładem takiego systemu są Intermedia - system stworzony przez IRIS (Instytut for Research in Information and Scholarship), działający w Brown University. Został stworzony za pomocą systemu operacyjnego 4.2 BSD UNIX i uruchomiony na IBM RT/PC ze stanowiskami Sun, które wspierały Sun's Network File System (NFS). System ten jest najlepszym

ucieleśnieniem idei hipertekstu, umożliwia bowiem tworzenie sieci z różnego rodzaju materiałów, nie tylko tradycyjnych tekstowych, ale również nagranych na video, materiałów dźwiękowych, etc. Za pomocą specjalnych programów intermedia umożliwiają tworzenie połączeń pomiędzy dokumentami. Edytor tekstowy, Inter-text, pozwala tworzyć ładnie sformatowany tekst, umożliwia pisanie artykułów i robienie notatek podczas używania systemu. System ma edytor czasowy (InterVal), który umożliwia zestawianie i wyświetlanie materiałów w ustalonym porządku chronologicznym.

Edytor graficzny (InterDraw) używany w połączeniu z przeglądarką (InterPix) służy do tworzenia i wyświetlania diagramów i innych ilustracji graficznych. Między tego rodzaju dokumentami mogą istnieć połączenia, również kontekstowe. Użytkownik może sam je tworzyć, czemu służy komenda "start link". Podczas tworzenia połączenia przez system użytkownik może wykonywać inne operacje nie powiązane z procesem tworzenia połączenia. Podczas lub też po utworzeniu połączenia może być dodana opcja wyjaśniająca połączenie, tzw. link explainer, która pomaga użytkownikom przy przeszukiwaniu sieci i selekcjonowaniu dokumentów. Poruszanie się po ekranie ułatwia i przyspiesza urządzenie zwane myszą. Wyszukiwanie odbywa się dwustopniowo, najpierw za pomocą globalnej mapy hiperdokumentu umożliwiającej lokalizację odpowiedniego fragmentu, a następnie za pomocą mapy lokalnej, reprezentującej jedno zagadnienie. Istnieje kilka poziomów szczegółowości, według których węzły i połączenia mogą być wyświetlane - od najogólniejszych, pokazujących jedynie całe dokumenty i połączenia między nimi, do najbardziej szczegółowych, wyświetlających połączenia i ich lokalizację w obrębie dokumentów.

Intermedia stanowiły podstawę dla trzech systemów edukacyjnych przeznaczonych dla studentów: "English Literature from 1700 to the Present", "English32" zawierających kurs literatury angielskiej, oraz "Plat Cell Biology", zawierającego informacje o strukturze i funkcji komórek. Systemy te umożliwiają w zasadzie samodzielne odbycie kursu, bez pomocy nauczyciela, który staje się jedynie instruktorem programu. Zauważono, że przy takim systemie nauczania zwiększa się krytycyzm studentów (system zawiera osobiste uwagi i notatki poszczególnych studentów) oraz rozszerza się znacznie zakres zainteresowań studentów (Rys. 2).



Rys. 2. Efekt użycia Intermedia w dyskusji na kursie English 32 /17/

Conklin /8/ charakteryzuje hipertekst opisując modelowe cechy systemu hipertekstowego (cechy te trudno uznać za definicyjne, gdyż pewne systemy, zaliczane do hipertekstowych, wszystkich podanych cech nie posiadają). Według Conklina modelowy system hipertekstowy spełnia następujące warunki:

1. Baza danych jest siecią węzłów (nazywana jest często hiperdokumentem).
2. Między oknami na ekranie a węzłami w bazie danych istnieje jedno-jednoznaczne przyporządkowanie. Każdy węzeł ma nazwę lub tytuł, który jest zawsze wyświetlany w oknie. Tylko niewielka liczba okien jest jednocześnie otwarta (tzn. wyświetlana na ekranie).
3. Okna mogą być przemieszczane, zamykane, można zmieniać ich wielkość, ograniczać do fragmentów (zawierających jedynie część informacji okna (icon), np. składających się jedynie z tytułu i połączeń z innymi ekranami). Zamknięcie okna powoduje jego zniknięcie z ekranu.
4. Wskazując (np. za pomocą myszy) fragment okna, tworzymy nowe okno w tym samym miejscu, w którym znajdowało się poprzednie.
5. Okna mogą zawierać dowolną liczbę fragmentów prowadzących do innych węzłów w bazie danych. Taki fragment okna (linked icon) zawiera krótkie pole tekstowe, które wymienia kontekst węzła, do którego skierowuje. Aby dotrzeć do zawartej w nim informacji należy wskazać je za pomocą myszy, co powoduje odnalezienie przez system wskazanego węzła i otwarcie na ekranie odpowiadającego mu okna.
6. Użytkownik może tworzyć nowe połączenia w bazie danych (nowe węzły), np. dla adnotacji, komentarzy, lub dodawać do już istniejących węzłów nowe powiązania.

7. Baza danych może być przeszukiwana na trzy sposoby:

a) poprzez połączenia otwieramy sukcesywnie nowe okna, zapoznając się z ich treścią.

b) poprzez przeszukiwanie sieci (lub jej części) za pomocą różnych kluczy, np. ciągu liter, cyfr, słów kluczowych.

c) przez przemieszczanie się w hiperdokumencie za pomocą specjalnego graficznego wyszukiwacza, który wyświetla sieć.

Jak widać, termin hipertekst może być więc rozumiany albo jako narzędzie umożliwiające zindywidualizowany dostęp do danych, albo jako odpowiednio zorganizowana struktura bazy danych.

Wydaje się, że jedyną cechą, która łączy wszystkie podejścia do hipertekstu, jest specjalne miejsce użytkownika w systemie. Użytkownik staje się jak gdyby częścią systemu hipertekstowego, gdyż sam go tworzy.

Wszystkie pierwsze próby tworzenia systemów hipertekstowych charakteryzowały się tym, że traktowały hipertekst jako pewną ideę. Więcej było zafascynowania hipertekstem jako nowym narzędziem komunikacji niż konkretnych prób tworzenia takiego narzędzia - na pewno w dużej mierze zdecydował o tym niedostatek techniki. Niemniej próby te otworzyły drogę do dalszych eksperymentów.

W dalszym rozwoju systemów hipertekstowych dwa zdarzenia odegrały tak znaczną rolę, że chyba warto o nich wspomnieć. Pierwsze z nich to agresywna promocja hipertekstowego systemu HyperCard przez firmę Apple, która spowodowała, że hipertekst z mało znanej koncepcji awansował do roli artykułu niezbędnego dla niemal każdego użytkownika komputera (choć sam system był dosyć prymitywny, porównywalny z pierwszymi systemami hipertekstowymi) /26/.

Drugim wydarzeniem, na skalę światową, była pierwsza międzynarodowa konferencja poświęcona w całości problemom hipertekstu - Hypertext'87 (Chapel Hill, North Carolina, USA, listopad 1987). Spotkali się na niej naukowcy z bardzo różnych dziedzin, co pozwoliło rozpatrywać hipertekst z różnych punktów widzenia, np. zastosowania w dydaktyce (nauka języków obcych), filozofii, medycynie /13/ (nie zabrakło również zainteresowania hipertekstem jako produktem handlowym).

Już ten krótki rozwój historyczny hipertekstu pozwala zauważyć, że hipertekst nie jest całkiem nowym wynalazkiem. Stąd wiele osób zadaje pytanie, o co właściwie tu chodzi - przecież to wszystko już było. Wydaje się, że hipertekst jest jednak czymś nowym - jest on przede wszystkim nowym stylem budowania systemów w celu reprezentacji i organizacji informacji.

Literatura

1. AKSCYN R., McCracken L.D., YONDER E. KMS: A Distributed Hypertext System for Sharing Knowledge in Organizations. Hypertext'87 Papers, Chapel Hill, NC, [listopad 13-15,] 1987 s. 1-20.
2. AKSCYN R., McCracken L.D. The ZOG Approach to Database Management. Proceedings of the Trends and Applications Conference: Making Database Work. Gaithersburg, MD, [maj,] 1984.
3. APPLE Computer. HyperCard Script Language Guide: The HyperTalk Language. Addison-Wesley Reading, [marzec] 1988.
4. BOLTER J.D. Hypertext and Creative Writing. Hypertext'87 Papers, Chapel Hill, NC, [listopad] 13-15, 1987 s. 41-50.
5. BROWN P.J. Turning Ideas into Product: the Guide System. Hypertext'87 Papers, Chapel Hill, NC, [listopad] 13-15, 1987 s.33-40.

6. BUSH V. As We May Think. Atlantic Monthly 176, 1 [czerwiec] 1945, s. 101-108.
7. BUSH V. Memex Revisited. W: Science Is Not Enough, V. Bush s. 75-101, William Morrow, 1967.
8. CONKLIN J. Hypertext: A Survey and Introduction, IEEE Computer, 1987, vol. 20 nr 9, s. 17-41.
9. DELISE N. Neptune: A Hypertext System for CAD Applications. Tektronix Computer Research Laboratory, Beaverton, Technical Report: CR-85-50, [styczeń] 1986.
10. ENGELBART D.C. Authorship Provisions in Augment. IEEE 1984 COMPCOM Proceedings, [wiosna] 1984.
11. ENGELBART D.C. A Conceptual Framework for the Augmentation of Man's Intellect. W: Vistas in Information Handling, Washington, D.C.: Spartan Books, 1963. Vol. 1 s. 1-29.
12. FRISSE M.E. Searching for Information in a Hypertext Medical Handbook. Hypertext'87 Papers, Chapel Hill, NC, [listopad] 13-15, 1987 s. 57-66.
13. GARRETT L.N., SMITH K.E., MEYROWITZ N. Intermedia: Issues, Strategies, and Tactics in the Design of a Hypermedia Document System. Computer-Supported Cooperative Work Proceedings, Austin, TX, [grudzień] 3-5, 1986.
14. GOLSTEIN I.P., BOBROW D.G. Description for a Programming Environment. Proceedings of the First Conference on AAAI, sierpień 1980.
15. GREGORY R. Xanadu-Hypertext from the Future. Dr. Dobb's Journal 1983 nr 75 s. 28-35.
16. HALASZ F.G. Reflection on NoteCards: Seven Issues for the Next Generation of Hypermedia Systems. Hypertext'87 Papers, Chapel Hill, NC, [listopad] 13-15, 1987.
17. HYPERTEXT and Pluralism: From Lineal to Non-lineal Thinking. Hypertext'87 Papers, Chapel Hill, NC, [listopad] 13-15, 1987 s. 67-88.
18. JANUSZEWSKA M. Hipertekst - zarys problematyki. Praca magisterska. Warszawa IBIN 1991 maszyn, 83 s.

19. LANDOW G.P. Context32: Using Hypermedia to Teach Literature. Proceedings of the 1987 IBM Academic Information Systems University AEP Conference, 1987.
20. LANDOW G.P. Hypertext in Literary, Education, Criticism and Scholarship. IRIS Publication, Brown University, RI, 1988.
21. MANTEI M., McCracken L.D. Issue Analysis with ZOG, A Highly Interactive Man-machine Interface. Proceedings of the First International Symposium on Policy Analysis and Information Systems, 1979.
22. McCracken L.D., AKSCYN R. The ZOG Approach to Database Management. Carnegie-Mellon University, Pittsburgh, PA, CS-34-113.
23. MEYROWITZ N. Intermedia: The Architecture and Construction of an Object-Oriented Hypertext/Hypermedia System and Applications Framework. Proceedings of the Conference on Object-Oriented Programming Systems, Languages, and Applications, Portland, Oregon, wrzesień 29 - październik 2, 1986.
24. NELSON T. H. Replacing the Printed World: A Complete Literary System. W: Information Proceedings 80. North-Holland Publishing Co., IFIO 1980, s. 1013-1023.
25. NIELSEN J. Trough Hypertext. Communication of the ACM. 1990 vol. 33 nr 3 s.297-310.
26. RAGLAND C. Hypertext, Hypermedia and the Macintosh. MacA.P.L.L.E., sierpień 1987.
27. RASKIN J. The Hype in Hypertext: A Critique. Hypertext'87 Papers, Chapel Hill, NC, listopad 13-15, 1987 s.325-330.
28. ROBERTSON C.K. McCracken L.D., NEWELL A. The ZOG Approach to Man-Machine Communication. Department of Computer Science, Carnegie-Mellon University, Pittsburgh. PA, Technical Report: CMU-CS-79-148, październik 1979.
29. WEYER S.A. As WE May Learn. Multimedia in Education: Interfaces to Knowledge-Education Advisory Council Conference Proceedings, kwiecień 1987.
30. YANKELOWICH N., LANDOW G.P., HEYWOOD P. Designing Hypermedia

Idebases-The Intermedia Experience. IRIS. Brown University, Providence, RI, IRIS Technical Report 87-4. 1987.

HYPERTEXT SYSTEMS - THE HISTORICAL OUTLINE

Summary

The article is devoted to the presentation of the brief historical outline of hypertext systems - of the first attempts by V. Bush and D. Engelbart as well as of the development of these ideas by J. Conklin. The author discusses the basic characteristic features of the hypertext systems.

ГИПЕРТЕКСТОВЫЕ СИСТЕМЫ - ИСТОРИЧЕСКИЙ ОЧЕРК

Резюме

В статье представлен краткий исторический очерк гипертекстовых систем - первые попытки В. Буша и Д.Энгельбарта (V. Bush и D. Engelbarta), а также дальнейшее развитие их предложений Я.Конклином (J. Conklina). Обсуждены основные характерные признаки гипертекстовых систем.

ELŻBIETA ARTOWICZ

Uniwersytet Warszawski

SYSTEM INFORMACJI HUNGAROLOGICZNEJ
WĘGIERSKIEJ BIBLIOTEKI NARODOWEJ

Źródła informacji w systemie. Podsystemy:
informacji bibliograficznej, informacji
biograficznej, informacji o instytucjach.
Wyszukiwanie informacji i obsługa informa-
cyjna użytkowników.

Jednym z zadań Biblioteki Szechenyi'ego, jako węgierskiej biblioteki narodowej, jest gromadzenie, opracowanie i udostępnianie informacji o zagranicznych hungaricach, a więc materiałach wspólnoty węgierskiej, zjawęcej poza obecnymi granicami Węgier^{1/}.

O wadze problemu zadecydowały różne czynniki, a wśród nich:

- rozproszenie wspólnoty węgierskiej i niebezpieczeństwo strat informacji,

- przyspieszenie procesów integracyjnych na świecie i rozwój kontaktów kulturalnych,

- względy praktyczne, na przykład brak zinstytucjonalizowanych kanałów i mechanizmów zapewniających rejestrację hungariców zagranicznych; zagadnienie to spełniała w pewnym, lecz nie wystarczająco

^{1/} Szczegółowej charakterystyce omawianego systemu poświęcono 5 referatów wygłoszonych na Międzynarodowym Kongresie Hungarologicznym w Szeged w sierpniu 1991 r. połączonych z demonstracją jego funkcjonowania. (Zob. Literatura).

jącym zakresie, biblioteka narodowa od momentu jej założenia w 1802 r.

Prace koncepcyjno-projektowe podjęto w ramach współpracy międzybibliotecznej w latach 70-tych. Sformułowano wówczas założenia systemu w wymiarze krajowym i międzynarodowym oraz przyjęto pewne ustalenia definicyjne (np. przyjęto definicję dokumentu "hungaricum", "Węgry"). Warunki do realizacji programu powstały w latach 80-tych, kiedy to rozpoczęto realizację systemu rejestracji danych bibliograficznych oraz przystąpiono do projektowania podsystemów informacji faktograficznej w ramach programów "Informacja o hungaricach" i "Rejestracja hungariców" finansowanych z funduszy resortu kultury. Od 1987 r. prace te przejęła praktycznie Biblioteka Szechenyi'ego, w której powstał Dział Dokumentacji Hungariców. W roku 1990, wykorzystując warunki techniczne dostępne w ramach programu "Infrastruktura informacyjna", przystąpiono do implementacji skomputeryzowanych baz danych.

Zakres zadań systemu wynika z rodzajów informacji w nim przetwarzanych: dokumentacyjnej (bibliograficznej) i faktograficznej (biograficznej oraz informacji o instytucjach). Aktualizacją i udostępnianiem zbiorów danych zajmuje się Dział Dokumentacji Hungariców.

Zakres tematyczny systemu określony jest przez zakres badań hungarologicznych. Jest to ciągle problem otwarty, dyskutowany przez hungarologów. Zakres ten wyznaczają z jednej strony nauki oparte na kryterium etnicznym (językoznawstwo, literaturoznawstwo, etnografia), co prowadzi do rozumienia tej dziedziny jako systemu zamkniętego, z drugiej strony badania hungarologiczne wykraczają poza tradycyjny krąg etniczny, wchodząc w obszar nauk historycznych, historii nauki itp. Punkt widzenia przyjęty

przy gromadzeniu informacji umożliwia respektowanie obu kryteriów w najszerszym ujęciu. Na tej podstawie ustalono definicję dokumentu "hungaricum" z uwzględnieniem trzech odmian informacji: językowej, personalnej i treściowej.

W aspekcie **treściowym** jako hungaricum traktuje się każdy dokument, którego przedmiotem jest język węgierski, Węgry, społeczność węgierska i jej działalność. Prezentacja treści może być wielowarstwowa: historyczna, geograficzna, etnograficzna, ekonomiczna, a w formie tekstowa, ikonograficzna, kartograficzna - bez ograniczeń czasowych.

Przewidziano przetwarzanie informacji z zakresu następujących badań dziedzinowych:

- **historia:** historia Węgier, do 1918 roku w granicach historycznych, informacje dotyczące historii mniejszości węgierskiej od 1918 roku, ruchy migracyjne wspólnoty węgierskiej oraz jej historia na emigracji,
- **literaturoznawstwo węgierskie**, z wyłączeniem oryginalnych dzieł literatury pięknej,
- **językoznawstwo węgierskie i ugrofińskie**,
- **historia kościołów** funkcjonujących na terenach węgierskich lub zamieszkałych przez mniejszość węgierską,
- **historia sztuki węgierskiej**, działalność artystyczna Węgrów na emigracji, zabytki węgierskie, architektura i sztuka ludowa,
- **informacja socjologiczna, socjograficzna, demograficzna i statystyczna** dotycząca wspólnoty węgierskiej i Węgier,
- **prace poświęcone historycznemu terytorium Węgier:** geografia, opisowe nauki przyrodnicze, meteorologia, geologia,
- **system polityczny, społeczny, ekonomiczny, prawny i życie społecznych Węgier;** wydarzenia życia naukowego i artystycznego, ich oddźwięk za granicą.

- wspólnota węgierska jako mniejszość - aspekty polityczne, ekonomiczne, działalność naukowa, artystyczna, kultura języka. Za szczególnie ważną uważa się możliwość wzbogacania wiedzy o możliwościach rozwojowych mniejszości węgierskiej, jej instytucji kulturalnych nie tylko w Europie, lecz także w Australii i Ameryce. Z uwagą śledzi się międzynarodową literaturę dotyczącą nacjonalizmu i mniejszości etnicznych, konferencji, rozmów i działalności organizacyjnej w tej kwestii.

- problemy obszaru Europy środkowowschodniej i ochrony środowiska,
- środkowowschodnio- i południowowschodnio-europejskie instytucje, organizacje o profilu węgierskim, węgierskie uniwersytety i katedry
- ich działalność literacka, artystyczna, wydawnicza;
- biografie i przyczynki biograficzne dotyczące Węgrów.

ŹRÓDŁA INFORMACJI W SYSTEMIE

Głównym źródłem informacji są zbiory Biblioteki Szechenyi'ego wraz z Centralnym Katalogiem Książek. Na podstawie umowy zawartej w 1958 roku z krajami regionu odbywa się wzajemna wymiana opisów bibliograficznych wydawnictw ciągłych, choć jest to źródło coraz uboższe. Materiał ten jest uwzględniany w pierwszej kolejności przy opracowywaniu bibliografii bieżących. Najogólniej biorąc, w systemie wykazuje się materiały źródłowe wyselekcjonowane i opracowywane według różnych kryteriów (węgierskie i zagraniczne). Ich określenie i ocena stały się przedmiotem badań bibliometrycznych w odniesieniu do bibliotek i instytucji zagranicznych, co niezależnie pozwoliło uzyskać obraz zakresu oddziaływania kultury węgierskiej poza granicami kraju.

Oddziaływanie danej kultury można badać z różnych punktów widzenia, odpowiadających różnym kanałom przekazu między kulturami.

Charakterystyka dokumentów hungarologicznych i zbiorów, w których są przechowywane, jest najefektywniejszą metodą z punktu widzenia systemu instytucjonalnego funkcjonującego w oparciu o profesjonalne metody działania.

Z punktu widzenia Biblioteki Szechenyi'ego pierwszorzędnym problemem było ustalenie obszarów pomiaru. Ocena zbiorów w krajach sąsiednich wydawała się z wielu względów pilniejsza i bardziej uzasadniona, jako dotycząca obszaru najsilniejszych związków i oddziaływań kultur węgierskiej. Jednakże w odniesieniu do krajów zachodnioeuropejskich, jako obszaru słabiej rozpoznanego, obiecywała więcej efektów. Z tego względu posługując się metodą ankietową, jako najbardziej dostępną, w pierwszej kolejności podjęto badanie tego rejonu. Jako pierwszy opracowano materiał dotyczący obszaru niemieckojęzycznego (Austria, Niemcy, Szwajcaria), opublikowany w 4 zeszytach. następny krok stanowił region północnoeuropejski (Skandynawia, kraje Beneluxu, Wielka Brytania). Zeszyt sumujący wyniki tego etapu prac są obecnie w druku. Kolejne kraje to: Francja i kraje Europy Południowej. Redagowany jest materiał dotyczący: Hiszpanii, Portugalii, Włoch i Grecji. W trakcie gromadzenia są dane z USA i Kanady.

Przy ustalaniu metod badania zagranicznych zbiorów hungaricó w istotnym problemem jest możliwe dokładne określenie instytucji, będących potencjalnym miejscem gromadzenia hungaricó w. Przeprowadzono tu przede wszystkim przegląd bibliotek danego kraju i ich zakresu przedmiotowego. Badano dane miasto, związki danego regionu z Węgrami w przeszłości. Pozwalało to uzyskać wiarygodny obraz nawiązywania się zbiorów rozważanych bibliotek. Na tej podstawie przeprowadzono wstępną selekcję instytucji i zestawiano ich listę. Jako źródło wykorzystywano międzynarodowe

informatory o systemach bibliotecznych oraz o życiu naukowym (np. World guide to libraries, World of learning) oraz informatory specjalistyczne^{2/} i inne źródła poświęcone przeglądowi systemów bibliotecznych w poszczególnych krajach^{3/}. Ponieważ jedną z podstawowych form działalności Biblioteki Szechenyi'ego jest wymiana międzynarodowa, wyselekcjonowane adresy bibliotek skonfrontowano z ewidencją służby wymiany. Uzyskane informacje uzupełniano na podstawie informatorów i innych źródeł, przygotowano dodatkowe opracowania i przewodniki. Wykorzystano dane z innych zbiorów informacji, np. wyciągi dotyczące średniowiecznych kodeksów, wyniki prac nad nowożytnymi rękopisami prowadzone pod kierunkiem Evy Windisch, dane dotyczące starodruków z katalogu rękopisów Biblioteki Szechenyi'ego i in. W sumie po przebadaniu 304 zbiorów zebrano informacje o 72.108 dokumentach pochodzących sprzed roku 1801.

Zakres i zawartość gromadzonych danych

Gromadzenie danych o zagranicznych zbiorach hungariców w celu ustalenia jednego ze źródeł informacji w systemie, a następnie ich upowszechnianie jest pracą informacyjną. Ich celem było m.in. uzyskanie informacji o przechowywanych w zbiorach materiałach węgierskich, typach dokumentów, ich proporcji w stosunku do całego zbioru i innych aspektach interesujących przy nawiązywaniu kontaktów badawczych. Z tego względu proszono

^{2/} np. Topographie der Osteuropa-, Südosteuropa- und der DDR Sammlungen in der BRD und Berlin/W, München 1978

^{3/} np. Verzeichnis deutscher Informations- und Dokumentationstellen. Wiesbaden 1974

ankietowane biblioteki o dane nt. wielkości i zawartości zbiorów, ich charakterze, systemie katalogowania, pochodzeniu hungariców, wydawnictwach, obsłudze specjalnej i zatrudnionych specjalistach. W przypadku każdej oceny odpowiadała tylko część ankietowanych (około 50%), jednakże w ogólnym podsumowaniu uwzględniono także odpowiedzi cząstkowe lub negatywne, zaznaczając ich wartość.

W trakcie opracowania danych będących źródłem informacji w systemie konieczne było zredagowanie i ustalenie układu materiału. Za najbardziej celowy, z punktu widzenia spodziewanych potrzeb badawczych i łatwości szybkiego przeszukiwania, uznano układ geograficzny - według krajów, miast, a w ich ramach - instytucji w układzie alfabetycznym. Pozornie zadanie to wydaje się proste - pojawił się jednak problem zjednoczonych Niemiec, łączenia i powstawania małych państw, co powodowało problemy przy redagowaniu poszczególnych tomów. Konieczne było ustalenie pewnych logicznych jednostek umożliwiających jednolite opracowanie materiału (np. nazwa i używany język - Szwajcaria, plus nazwa instytucji, Wielka Brytania i Irlandia Północna jako państwo i region kulturalny, Benelux i w jego ramach system instytucji francuskich i flamandzkich). Z punktu widzenia potrzeb wyszukiwania konieczne było opracowanie szeregu indeksów: nazw instytucji wraz z ich wariantami nieoficjalnymi, indeksu przedmiotowego oraz indeksu według typów bibliotek (np. uniwersyteckie, publiczne, kościelne). Kłopoty stwarzała też częsta rozbieżność danych podawanych w ankietach, informatorach i specjalistycznych źródłach, jak również formy odpowiedzi na ankietę.

Stanowiąca element prac projektowych analiza zagranicznych źródeł informacji hungarologicznej pozwoliła sformułować pewne ogólne wnioski przydatne w badaniu oddziaływania innych kultur.

Biblioteki można oceniać jako ważny kanał spontanicznej komunikacji międzykulturowej oraz jako przejaw zinstytucjonalizowanej współpracy z danym krajem. Cecha spontaniczności zależy od zainteresowania danym krajem, stworzeniem kontaktów badawczych i warunków dla nich - przykładem może być tradycja badań ugrofinistycznych na Uniwersytecie w Getyndze i osobisty wkład J. Sojnovicsa, stworzenie ośrodka badań zagadnień Europy Środkowoschodniej w Monachium, obecny rozwój badań hungarologicznych i aktywizacja nowych i tradycyjnych ośrodków - w Paryżu, Hamburgu, Rzymie, Bolonii, Bloomington.

Historycznie najdawniejszą warstwę materiałów hungarologicznych tworzą rękopisy, mapy i inne druki o charakterze źródłowym powstałe jako wynik działalności indywidualnych grup, osób lub instytucji. Do grupy tej dołączają się dokumenty związane z działaniem w danym regionie społeczności węgierskiej, w tym będące np. spuścizną po wybitnych osobistościach tej wspólnoty (np. Erdei Grünwald Béla w Londynie).

Z badań wynika, że oprócz czynników funkcjonalnych główną inspiracją rozwoju badań jest działalność pojedynczych osób oraz dobrze sformułowane programy działania.

Zakres oddziaływania kulturalnego Węgier w Europie Zachodniej maleje wraz ze wzrostem odległości geograficznej. Przy analizie obszaru badań i zbiorów należy też uwzględniać czynniki historyczno-kulturowe i społeczno-ekonomiczne. Z analizy wynika, że poza obszarem austriackim i niektórymi regionami niemieckimi, zbiory francuskie, szwajcarskie i północnoeuropejskie zawierają znacznie mniej hungariców, a spośród krajów południowoeuropejskich najbogatszymi zbiorami dysponują Włochy.

Badania przyniosły także szereg cennych informacji szczegółowych - np. odnaleziono listy Kossutha, mapy, dokumenty poświęcone obyczajom węgierskiej społeczności Niemców, materiały ikonograficzne z zakresu architektury i sztuki węgierskiej (ok. 3500).

Z punktu widzenia kompletności zgromadzony materiał nie osiągnął wymaganego poziomu, co po części można tłumaczyć względami materialnymi, po części jednak jest to spowodowane przyczynami zawodowymi, na które składają się:

- uzasadnione celami badawczymi i dydaktycznymi instytucji programy gromadzenia zbiorów,
- poziom obsługi informacyjnej niezbędnej do właściwego gromadzenia. Obsługa ta należy do obowiązków i kompetencji zainteresowanego kraju, a więc w tym wypadku Węgier.

Na podstawie gromadzonych krajowych i zagranicznych źródeł informacji Biblioteka Szeczeni'ego rozpoczęła implementację trzech podsystemów informacji hungarologicznej stanowiących docelowo element centralnego zintegrowanego systemu informacji. Zadania związane z budową każdego z podsystemów są realizowane na podstawie wspólnego zbioru materiałów zgodnie z zasadą jednokrotnego opracowania i wielokrotnego wykorzystania.

PODSYSTEM INFORMACJI BIBLIOGRAFICZNEJ

Opracowanie bibliografii hungariców zagranicznych rozpoczęto w Bibliotece Szeczeni'ego w roku 1971. W ich wyniku opublikowano 2 bieżące bibliografie:

- "Zagraniczne publikacje/wydawnictwa w języku węgierskim (1972-76): wykaz zagranicznych czasopism w języku węgierskim" - pod red. Marii Orvos, (Külföldi magyar nyelvű kiadvánvok),
- "Przegląd literatury w zakresie hungariców (1971-76): przegląd zagranicznych czasopism w zakresie hungariców - pod red. Jánosa

Szentmibályiego i Judit Szerb (Hungarika irodalmi szemle"). Bibliografie te stały się podstawą stworzenia zintegrowanej bibliograficznej bazy danych do roku 1989. Opublikowano je ponadto jako niezależne wydawnictwa dla Centralnego Katalogu książek.

Zaprezentowanie zagranicznych hungariców książkowych w formie bibliografii bieżącej planowano w Bibliotece Szechenyi'ego już na początku lat 70-tych, a jej początkiem stało się opublikowanie pracy "Zagraniczne książki węgiersko-języczne 1961-70". Najpierw przygotowano rękopis bibliografii książek opublikowanych w Rumunii, następnie w 1981 r. jako pierwszy element serii ukazały się na mikrokartach materiały dotyczące Czechosłowacji, po czym opublikowano materiały rumuńskie.

W latach 70-tych specjaliści Biblioteki Szechenyi'ego rozpoczęli także prace przygotowawcze o charakterze metodologicznym dotyczące: profilu wydawnictw i ich układu, zakresu źródeł informacji, gromadzenia i przechowywania informacji. Przy podejmowaniu zadania dostarczenia w bibliografii pełnego przeglądu produkcji hungarologicznej należało bowiem liczyć się z wieloma trudnościami, na co zwracał uwagę już w 1936 roku József Fitz we wstępie do "Bibliotekarstwa węgierskiego", np.:

- na podstawie jakich kryteriów ustalić górną granicę gromadzenia (języcznych, terytorialnych, przedmiotowych),
- jakie są praktyczne możliwości przechowywania informacji i jej wartość,
- w jaki sposób można zapewnić dopływ niezbędnej informacji.

Cele prowadzonych prac bibliograficznych są dwojakie:

- z jednej strony rejestracja bibliograficzna hungariców zagranicznych,

- z drugiej strony stworzenie systemu informacji online zapewniającego ciągłą obsługę użytkowników.

W ramach prac projektowych zdefiniowano oba rodzaje usług, ponieważ różnice między nimi mają charakter zarówno treściowy, jak i metodologiczny. Różnice treściowe wynikają stąd, że zadaniem SIW jest w pierwszej kolejności dostarczanie informacji o najszerzej rozumianych badaniach hungarologicznych bez względu na rodzaj i formę źródła informacji (artykuły, doniesienia w kronikach czasopism, pryczyńki, anonsy), natomiast bibliografia nakłada pewne ograniczenia formalne i treściowe, pomija wydawnictwa ulotne, zwraca uwagę na materiały węgierskojęzyczne, choć nie węgierskoprzemiotowe, a także niewęgierskojęzyczne, lecz pisane przez węgierskich autorów. Na tych zasadach opiera się struktura zintegrowanej bazy danych, której część zasobów jest wspólna, część służy tylko do obsługi informacyjnej, a część tylko do przygotowania bibliografii. Metodologicznie wspólną zasadą przyjętą w bazie dla obu form obsługi jest jednolity opis bibliograficzny zgodny z międzynarodowymi standardami. Rekordy bazy danych są aktualizowane stosownie do potrzeb obu rodzajów usług - np. o nowe opisy bibliograficzne lub wyrażenia języka słów kluczowych, hasła przedmiotowe itp.

Kryteria gromadzenia danych bibliograficznych zostały określone na podstawie trzech punktów widzenia:

a) typy hungaricóv zagranicznych: w języku węgierskim, autor Węgier i Węgry jako przedmiot dokumentu,

b) typy dokumentów (samoistne, niesamoistne):

- samoistne (książki, czasopisma, wideokasety; kompletność tej części zbioru ma zapewnić m.in. wykorzystanie Centralnego Katalogu Bibliotek, jego obecna objętość wynosi rocznie 1000-2000 tytułów.

- niesamoistne (artykuły, studia); osiągnięcie kompletności jest tu prawie niemożliwe, ponieważ oznaczałoby to konieczność pełnego referowania zawartości węgierskich wydawnictw ciągłych, a także wyczerpujący przegląd zagranicznej produkcji czasopism oraz prasy - jako programu maximum - ponad 10 000 tytułów. Dostępne środki pozwalają przetwarzać rocznie 4-5 tys. tytułów. Konieczne było ustalenie listy rangowej. Opracowana na podstawie dostępnego na Węgrzech wykazu lista czasopism hungarologicznych obejmuje około 2000 wydawnictw ciągłych. Są to w pierwszej kolejności węgierskie i zagraniczne czasopisma naukowe, niekiedy roczniki, węgierskojęzyczne czasopisma literackie i kulturalne, religijne, tygodniki i parę dzienników. Obecnie możliwe jest opracowanie w systemie ponad 300 tytułów czasopism: 100 tytułów węgierskojęzycznych i 218 obcojęzycznych. Proporcje w liczbie pozycji wybieranych z analizowanych źródeł są dokładnie odwrotne. Analiza czasopism jest bardzo dokładna, chodzi bowiem o upowszechnianie jak najszerszym kręgom materiałów mniej znanych.

W celu wzbogacenia źródeł informacji pozyskuje się także odpowiednio zaadaptowaną informację z innych systemów - np. prawniczą, polityczną, ekonomiczną, socjograficzną. W tym względzie niezwykle cenna będzie w przyszłości współpraca z zagranicznymi placówkami hungarologicznymi.

c) ustalenie priorytetów badawczych w poszczególnych dziedzinach, ale także opracowanie bibliograficzne literatury pięknej.

Poza skomputeryzowanym systemem informacji pozostaje problem oceny nowo ukazujących się czasopism węgierskich i obcojęzycznych, a także uzupełnianie listy źródłowych czasopism bazy danych. Przegląd czasopism oznacza, że prace bibliograficzne są tu kluczowym problemem w zapewnieniu jak największej kompletności zbiorów

bibliografii narodowej. Zadanie to zostało włączone także do zadań Węgierskiej Akademii Nauk.

Zgodnie z planami zarejestrowane dokumenty będą ukazywać się w formie rocznych bibliografii traktowanych jako część bibliografii narodowej. Hungarica zagraniczne obejmą 2 bibliografie:

- "Magyar nemzeti bibliográfia. Külföldön megjelent hungarikumok. Könyvek [Węgierska bibliografia narodowa. Hungarica ukazujące się za granicą. Książki] (jak również inne wydawnictwa samostne i dotyczące ich doniesienia);

- "Magyar nemzeti bibliográfia. Külföldön megjelent hungarikumok. Cikkek (tanulmányok) [Węgierska bibliografia narodowa. Hungarica ukazujące się za granicą. Artykuły, studia].

Skomputeryzowany system umożliwi w przyszłości tworzenie wydawnictw skumulowanych przy stosunkowo niewielkich nakładach. Dla potrzeb kumulacji materiału konieczna jest selekcja według różnych punktów widzenia (kraje, regiony, język i ich kombinacje) - możliwe będzie np. kumulacyjne wydanie materiałów czeskich i rumuńskich (1961-70). Obecnie baza zawiera 7000 charakterystyk wyszukiwawczych dokumentów. Zakładany przyrost rekordów w bazie wynosi: minimum 4500 rekordów rocznie - maksimum 10 000 rekordów.

PODSYSTEM INFORMACJI BIOGRAFICZNEJ - HUNGARIKA NÉVKATSZER
(Index Biograficus Hungaricus)

Celem bazy jest rejestracja informacji o osobistościach węgierskich, które w XX wieku wyemigrowały z kraju. Rejestracja biografii odbywa się z jednej strony według kryteriów leksykonu biograficznego, a z drugiej według wymagań indeksu. Istnieje dość bogata tradycja węgierskiej leksykografii biograficznej, np. Dávid Czwittinger (1711), József Szinyei (1901), Pál Gulyás (1939-1944), leksykon węgierskiej literatury emigracyjnej pod

redakcją Csaby Nagyó opracowany w ramach współpracy Instytutu Literaturoznawstwa Węgierskiej Akademii Nauk i Muzeum Literatury S. Petőfiiego.

Rozbudowa bazy umożliwiła poszerzenie rejestru o osoby dysponujące znacznym dorobkiem, a nie uwzględnione w opublikowanych leksykonach, oraz według różnych aspektów (dzieła, dziedzina, specjalizacja, region, okres itp.) przy wykorzystaniu międzynarodowych źródeł biograficznych, opublikowanych w Niemczech, Francji i USA.

Baza danych biograficznych będzie mogła służyć za podstawę równoczesnego przygotowania wydawnictw oraz bieżącej obsługi informacyjnej. "Hungarika Néukatszer", jest bowiem równocześnie indeksem biograficznym i skomputeryzowanym systemem informacji. W jego budowie wykorzystano kilka przykładów międzynarodowych indeksów biograficznych, jednakże do jego stworzenia posłużyły przede wszystkim prace ewidencyjne prowadzone w ramach realizacji podstawowych funkcji Biblioteki Szechenyi'ego. Z tego punktu widzenia w pierwszej kolejności wykorzystano rejestr wybitnych osobistości węgierskich opracowywany w Dziale Gromadzenia oraz w Dziale Katalogowania.

Z punktu widzenia rozwoju zasobów bazy danych przy jej zakładaniu konieczne było przyjęcie dwu punktów widzenia: wykorzystanie opublikowanych (lub nie) zbiorów danych, umożliwiających szybką rozbudowę bazy i uwzględnienie prac zrealizowanych wcześniej oraz możliwie wyczerpująca aktualizacja danych na podstawie bieżących publikacji prasowych oraz ankietowania. Wykorzystano następujące typy źródeł:

- rejestry Biblioteki Szechenyi'ego,
- specjalistyczne i ogólne leksykony biograficzne, dotyczące węgierskiej literatury emigracyjnej i jej twórców,

- monografie, podręczniki historii literatury opublikowane w kraju i za granicą,
- antologie współczesnej literatury węgierskiej powstałej w Słowacji, w krajach Europy Zachodniej, w Rumunii oraz na Ukrainie Zakarpackiej,
- prace zbiorowe,
- katalogi wystaw,
- książki adresowe, roczniki, druki ulotne, bieżące artykuły prasowe, zaproszenia na wystawy.

W kontekście prac prowadzonych w tych działach jednym z pierwszych zadań do rozwiązania było zdefiniowanie zakresu osób uwzględnianych w rejestrach hungariców, a więc zdefiniowanie "kryteriów przynależności do wspólnoty węgierskiej", jak również dalsze określenie cech umożliwiających identyfikację takich osób włączając zakres ich działalności i region geograficzny. Ustalenie tych cech zadecydowało o strukturze bazy danych. Zakres rejestrowanych danych i sposób ich prezentacji potraktowano z uwzględnieniem tych punktów widzenia.

Zakres opisu biograficznego

W charakterystyce wyszukiwawczej osób uwzględnia się następujące dane:

- nazwisko i imię (wraz z ewentualnymi wersjami),
- data i miejsce urodzenia/śmierci,
- poziom i miejsce studiów,
- dane o udaniu się za granicę (kiedy, dokąd),
- dalsze zmiany miejsca pobytu (w jakich krajach żył i działał),
- zajęcie, zawód, zakres działalności.

Dzięki skomputeryzowanemu przetwarzaniu materiału może być uzupełniany i selekcyjonowany według różnych punktów widzenia oraz dostępny według różnych kryteriów, np.:

- nazwisko,
- nazwa geograficzna (miejsce urodzenia, nauki/studiów, miejsca pobytu), itd.
- data (urodzenia, śmierci, podróży/przemieszczania się),
- zajęcie, dziedzina działalności (ze wskazaniem tzw. funkcji hun-garica, a więc pełnionej w środowiskach i instytucjach węgierskich),
- inne specjalne punkty widzenia, np. status (przynależność do wspólnoty węgierskiej), materiały ikonograficzne.

Zakres bazy i granice czasowe

Zasadniczym problemem w definiowaniu zakresu bazy danych było udzielenie odpowiedzi na pytanie: Kogo uważa się za Węgra? a więc jakie są tzw. kryteria przynależności narodowej. Kryteria takie zaproponował, przedyskutował, a następnie zatwierdził w 1976 roku Komitet Historii Bibliotek Węgierskiej Akademii Nauk. W trakcie opracowywania tych zasad zwrócono uwagę, iż pewne punkty widzenia należy potraktować ze szczególną wnikliwością, np. problem przynależności do wspólnoty węgierskiej. Przyjęto tu następujące kryteria:

- Węgrem jest ten, kto urodził się na Węgrzech,
- znaczną część nauki/studiów odbył na Węgrzech,
- publikował w języku węgierskim, jest więc węgierskojęzyczny,
- uważa się za Węgra.

W publikacjach prasowych, w wypowiedziach publicznych na ogół nie rozróżnia się stopni przynależności do wspólnoty węgierskiej, rozumiejąc tę relację bardzo szeroko. Z punktu widzenia potrzeb badawczych NEVKATSZER dąży do wypracowania możliwie obiektywnego i stopniowalnego wzorca przynależności do wspólnoty, umożliwiającego późniejsze grupowanie materiału.

Podstawową warstwę zbioru danych tworzą osoby pierwszej generacji, przy których występuje zbieżność wszystkich kryteriów (urodzenie się na Węgrzech, nauka na Węgrzech, posługiwanie się językiem węgierskim, uważanie się za Węgra).

Drugą warstwę stanowią osoby pierwszej generacji, które opuścili Węgry przed 18 rokiem życia, studiów wyższych nie odbywały na Węgrzech. Tu decydujące jest posługiwanie się językiem węgierskim oraz uważanie się za Węgra.

Trzecia i kolejne warstwy stanowią osoby kolejnych generacji.

Metoda przetwarzania danych polega na tym, iż każda grupa jest opatrywana odpowiednim kodem, co umożliwia odrębne lub wspólne przetwarzanie bądź selekcjonowanie danych.

Kolejnym problemem metodologicznym wymagającym rozstrzygnięcia na etapie prac projektowych było ustalenie "kogo uważa się za osobę wybitną/osobistość, zasługującą na wymienienie w którymkolwiek kręgu działalności i w jakiś sposób przyczyniającą się do rozwoju wspólnoty węgierskiej". Przyjęto, że dokładna ewidencja osób z tego zakresu jest istotna w takim samym stopniu, jak ewidencja osób posiadających znaczny dorobek, ponieważ uzyskanie tego typu danych dla potrzeb badawczych oznacza bardzo często duży wysiłek. Przy ustalaniu progu wartości dorobku wypracowano szczegółowe zasady, które można określić następująco: w rejestrze NEVKATSZER uwzględniane są z jednej strony osoby publikujące, z drugiej zaś specjaliści działający w sferze intelektualnej, biznesmeni, artyści, twórcy, a także prowadzący działalność w kręgach wspólnoty węgierskiej na jej rzecz - nauczyciele, księża, przywódcy związkowi, kierownicy grup artystycznych.

W trakcie prac projektowych oceniano, że jakkolwiek procesy emigracyjne miały miejsce w różnych okresach historycznych, z różnych powodów i w różnych proporcjach, to okresem najsłabiej

udokumentowanym są fale emigracji i przesiedlenia w XX wieku. Katalog zawiera dane o osobach urodzonych po 1880 roku, przy czym szczególną wagę przywiązuje się do historii najnowszej oraz do zebrania materiałów dotyczących okresu po 1945 roku. Przy opracowaniu materiały dotyczące czasów wcześniejszych opatruje się specjalnymi kodami.

Problemy wykorzystania

Przy praktycznym stosowaniu zasad sporządzania charakterystyk wyszukiwawczych poszczególnych osób okazało się, że sformułowane na etapie projektowania systemu zasady wymagają weryfikacji.

Wiele kontrowersji budzą zasady ustalania tożsamości narodowej. Nie zakładano na przykład śledzenia losów emigracji węgierskich grup etnicznych, ponieważ jest to przedmiotem innych badań. W wielu wypadkach trudno było ustalić przynależność - np. Sasów siedmiogrodzkich, Szwabów naddunajskich lub osób urodzonych na terenie monarchii.

Pojawiały się problemy z przyjęciem za podstawową jedną z kilku wersji nazwiska (węgierskiej lub międzynarodowej) - np. VESARELY lub VÁSÁRHELYI, AMERIGO TOT lub TÓTH IMRE, ponieważ pod nazwiskiem węgierskim rejestruje się wszystkie osoby urodzone na Węgrzech z zaznaczeniem innych form danego nazwiska;

W wykorzystanych źródłach często pojawiały się sprzeczne dane; rejestruje się najpełniejsze i najbardziej wiarygodne z zaznaczeniem rozbieżności w odpowiednim polu charakterystyki danej osoby: modyfikacji danych dokonuje się tylko na podstawie konkretnych źródeł;

Problematyczny jest często czas wyjazdu z Węgier i ustalenie przebiegu kariery za granicą; często konieczne są rozstrzygnięcia arbitralne - np. uzyskanie dyplomu studiów na Węgrzech i za

granicą, data ostatniej publikacji na Węgrzech, metryka dzieła wydanego za granicą.

W wypadku osób urodzonych w częściach kraju odłączonych od Węgier problemy sprawia ustalenie kraju urodzin. W tym wypadku istotna jest data urodzenia. W braku daty trudno rozstrzygnąć czy dana osoba urodziła się na Węgrzech, w Czechosłowacji czy w ZSRR. Analogiczne problemy powstają przy określaniu miejsca uzyskania dyplomu.

Nowe nazwiska wprowadzane są do bazy na podstawie udokumentowanych źródeł, które często są jednak fragmentaryczne. Niekiedy trudno ustalić tożsamość osoby wymienianej w jednym źródle, np. jako KISS PÉTER - poeta żyjący w USA, a w innym jako KISS PÉTER - dziennikarz żyjący w USA; Trudności z ustalaniem tożsamości spowodowane są także rozbieżnościami w pisowni nazwiska, np. KOVÁCS LÁSZLÓ - FIZYK - USA, KOVÁTS LÁSZLÓ - FIZYK - USA. Zgromadzone dane wymagają ciągłej aktualizacji ze względu np. na zmiany zawodów, śmierć, przesiedlenie się do innego kraju, repatriację. Wymienione przykłady ilustrują problemy najbardziej typowe, nie wyczerpują jednak pełnego ich zakresu.

Budowana baza danych biograficznych ma służyć różnym celom, przede wszystkim ma wypełnić lukę treściową w zbiorach biograficznych. Funkcjonuje jako część skomputeryzowanego systemu informacji Biblioteki Szechenyi'ego. Obecnie zawiera 16000 pozycji, jest systematycznie aktualizowana. Na jej podstawie publikowany jest Index Biographicus Hungaricus, możliwe jest wyszukiwanie online. Perspektywicznie planuje się rozpowszechnianie informacji na dyskach optycznych. Baza ta jest uzupełnieniem archiwum materiałów źródłowych.

PODSYSTEM INFORMACJI O INSTYTUCJACH

Obecnie nie jest jeszcze skomputeryzowany. Ma służyć celom analogicznym jak informacja biograficzna - gromadzeniu danych bieżących i historycznych o węgierskich instytucjach za granicą (towarzystwach i stowarzyszeniach, kościołach, szkołach, wydawnictwach, muzeach itp.), a zwłaszcza: a) o instytucjach stworzonych przez Węgrów, b) o zagranicznych zbiorach bibliotecznych zawierających znaczące dla badań hungarologicznych zbiory hungariców, a także przebadanie krajowych zbiorów bibliotecznych (poza zbiorami biblioteki narodowej) z punktu widzenia hungariców zagranicznych. Ocena wyników badań jest zadaniem przyszłościowym.

Jako źródło informacji wykorzystywane są materiały opracowywane w ramach gromadzenia danych biograficznych. Opracowano zestaw cech identyfikacyjnych instytucji: nazwa oficjalna i jej odmiany, adres, telefon, telex, rok założenia, stan osobowy, nazwisko kierownika, okres funkcjonowania, cel, charakter, działalność, publikacje, filie, związki z innymi instytucjami, zasoby - źródła finansowania.

Dotychczas opracowany zbiór zawiera dane o około 2500 instytucjach. Pierwszym rezultatem prac nad podsystemem informacji o instytucjach związanych z badaniami hungarologicznymi jest wydawnictwo seryjne opublikowane w 4 tomach w 1990 roku pt. "Wykaz bibliotek zagranicznych przechowujących hungarica". Zadanie opracowania takiego rejestru podejmowała w latach 70-tych Biblioteka Szechenyi'ego, a następnie w latach 80-tych na pierwszym zjeździe krajowych i zagranicznych bibliotekarzy węgierskich sprecyzowano cel stworzenia takiego rejestru.

WYSZUKIWANIE INFORMACJI I OBSŁUGA INFORMACYJNA UŻYTKOWNIKÓW

W pracach projektowych nad systemem założono jednokrotne, wieloaspektowe opracowanie źródłowe informacji i jej wielokrotne wykorzystanie w podsystemach (bazach danych). Implikuje to równoległe użycie dla celów indeksowania i wyszukiwania kilku języków informacyjno-wyszukiwawczych: języka słów kluczowych UKD, języka haseł przedmiotowych oraz kodów dla nazw geograficznych i instytucji. Przy wyszukiwaniu według słów kluczowych możliwe jest tworzenie łańcuchów synonimów wyszukiwawczych łącznie z użyciem eufemizmów. Kody stosowane są przy wyszukiwaniu według cech geograficznych. Autorzy systemu podkreślają problemy powstające zarówno przy indeksowaniu, jak i wyszukiwaniu według klucza geograficznego; na przykład przy kodowaniu nazw państw i miast np. na terenie Czechosłowacji, dawne miasta węgierskie trzeba było opatrywać kodami Węgier. Charakterystyki wyszukiwawcze dokumentów w bazie bibliograficznej opatrzone są kodem kraju publikacji i języka oraz adnotacją. Problemem dla operatorów systemu i użytkowników jest zmieniająca się historycznie i współcześnie informacja o regionach w krajach sąsiadujących, np.:

- Serbsko-Chorwacko-Słoweńskie Królestwo, które następnie otrzymało nazwę Nagybecskerek, przetłumaczoną oficjalnie na Veliki Beckerek, w 1930 roku zmienioną na Petrograd, a w socjalistycznej Jugosławii w 1947 roku na Zrenjanin; używanie węgierskich nazw miejscowości stało się na długi czas tabu politycznym;
- Karpatalja - Verlovina-Karpat-Ukrajna-Karpatontuli Ukrajna. Konieczne jest równoległe używanie parametrów geograficznych i czasowych.

Obsługę informacyjną użytkowników prowadzi służba Działu Dokumentacji Biblioteki Szechenyi'ego "Hungarika Informacio" [Informacja o hungaricach], korzystając zarówno ze środków tradycyjnych, jak i nowoczesnych.

Skomputeryzowana baza danych bibliograficznych jest przeznaczona przede wszystkim do wyszukiwania online. Baza bibliograficzna zawiera obecnie ok. 7000 rekordów. Za kilka lat przy wspomnianym zakładanym powiększaniu się zbiorów bazy o minimum 4500 do maksimum 10000 jednostek rocznie niezbędne będzie rozróżnienie wyszukiwania bieżącego i retrospektywnego.

Planowana jest rozbudowa sieci informatycznej, co umożliwi dostęp do bazy użytkownikom zewnętrznym za pośrednictwem centralnego systemu Biblioteki Szechenyi'ego NEKTÁR. Obecnie wyszukiwanie online jest możliwe za pośrednictwem sieci IIF (Információ Infrastruktúra Fejlesztése Program - [Program Rozwoju Struktury Informacyjnej]).

Lokalne wykorzystanie bazy jest obecnie możliwe w formie kart informacyjnych i indeksu (subskrypcja), mikrofisz i zbiorów na dyskietkach, z których można tworzyć własne bazy.

PERSPEKTYWY I KIERUNKI ROZWOJU

Obecnie system informacji funkcjonuje na mikrokomputerach kompatybilnych z IBM przy użyciu pakietu MICRO ISIS, zaadaptowanego przez pracowników Biblioteki Szechenyi'ego. Budowane systemy i bazy danych będą połączone w ciągu 2-3 lat z centralnym systemem biblioteki NEKTÁR i udostępniane za jego pośrednictwem. Celem perspektywicznym jest utworzenie europejskiej sieci informacji hungarologicznej. Pierwsze kroki podjęto na europejskiej konferencji ośrodków hungarologicznych w Paryżu w 1990 roku.

Literatura

1. BERKE Barnabásné: NEKTÁR. Nemzeti könyvtári átfogó információ rendszer. In: Könyvtáros 46. évf. 9 sz. (1990), pp. 527-535.
2. COLLOQUE EUROPEEN des centres de hungarologie organisé par le Centre Interuniversitaire d'Études Hongroises en collaboration avec l'Institut Hongrois de Paris. Paris 10-12 Octobre. Recommendations.
3. FARAGÓ Lászlóne, GÁL Julia, KOVÁCS Ilona: A külföldi hungarika gyűjtemények felmérése (A hungarika anyagot őrző külföldi könyvtárak címjegyzéke c. sorozat munkálatai). Szeged 1991, s. 10.
4. HALASZNE Nagy Marta, NÉMETH Maria, KOVÁCS Ilona: Hungarika névkatszer, Szeged 1991, s. 12.
5. KOVÁCS Ilona: Hungarika dokumentációs műhely - koncepció és távlatok a külföldi hungarika információk feltárasában. Szeged 1991, s. 12.
6. LUIF Miklós: A Zürichi Zentralbibliothek magyar vonatkozású forrasai. Magyar Könyvtárosok II. Tudományos találkozója. Bp. 1985, aug. 26-27.
7. MAKKAI Béla, N. RÁCZ Aranka, KOVÁCS Ilona: Hungarika információ és hungarológiai kutatás. Szeged 1991, s. 9.
8. ORVOS Maria, SZERB Judit, KOVÁCS Ilona: A hungarika bibliográfia rendszere. Szeged 1991, s. 11.
9. OSZK Irattár 521/1976. A könyvtari dokumentumok magyar vonatkozásai (Hungarikum meghatározás), 1977 március.
10. WINDISCH Éva, V.: Az ujkori hungarikum kéziratkataszter /UUK/ munkálatairól. /In:/ Magyar Könyvszemle 105. évf. 1 sz. (1989), p. 61-70.

10.02.1992 r.

SYSTEM OF HUNGAROLOGICAL INFORMATION OF THE HUNGARIAN
NATIONAL LIBRARY

Summary

There is presented the process of the establishment of the system of hungarological information, its sources of information, the scope and the content of the collected data. There are discussed in detail the separate elements of the system: the subsystem of bibliographic information, the subsystem of biographic information, the subsystem of information on institutions and the perspectives of development.

СИСТЕМА ВЕНГРОВЕДЧЕСКОЙ ИНФОРМАЦИИ ВЕНГЕРСКОЙ НАЦИОНАЛЬНОЙ
БИБЛИОТЕКИ (HUNGARICA).

Р е з ю м е

В статье представлен процесс создания венгроведческой информационной системы "Hungarica", использованные источники информации, объем и содержание накапливаемых данных. Подробно охарактеризованы отдельные элементы системы: библиографическая подсистема, биографическая подсистема и подсистема информации об учреждениях. Указаны перспективы ее развития.

JADWIGA SADOWSKA

Biblioteka Narodowa

WYBRANE REGUŁY METODYKI OPRACOWANIA PRZEDMIOTOWEGO

W OCENIE UŻYTKOWNIKÓW KATALOGÓW*

Wyniki badań ankietowych wybranych zasad metodyki opracowania przedmiotowego zawartych w "Instrukcji tematowania i katalogu przedmiotowego", przeprowadzonych w katalogu przedmiotowym Biblioteki Uniwersytetu Łódzkiego i katalogu przedmiotowym Książnicy Miejskiej w Toruniu oraz wśród studentów Instytutu Bibliotekoznawstwa i Informacji Naukowej.

Tworząc katalogi biblioteczne zwykle odwołujemy się do jakichś norm i zasad metodycznych, które ułatwiają zachowanie ich jednolitości. W przypadku polskich katalogów przedmiotowych w bibliotekach uniwersalnych taką pomocą jest podręcznik Adama Łysakowskiego^{1/}. W druku jest "Instrukcja tematowania i katalogu przedmiotowego"^{2/}, która w założeniu miała być kontynuacją i uzupełnieniem wspomnia-

* Materiał ten był przedstawiony w nieco innej postaci na ogólnopolskiej konferencji nt. Opracowania przedmiotowego dokumentów oraz problemów dydaktycznych opracowania rzeczowego w szkołach wyższych, Gdańsk, 7-9.05.1991

1/ A. Łysakowski: Katalog przedmiotowy. Podręcznik. Warszawa: PZWS 1946

2/ Instrukcja tematowania i katalogu przedmiotowego. Oprac. J. Sadowska. Warszawa: SBP; BN 1991, 100 s.

nego podręcznika. Ponadto każda biblioteka wypracowuje pewne charakterystyczne dla siebie zasady, najczęściej wynikające ze specyfiki gromadzonych zbiorów.

Tworząc katalogi staramy się też brać pod uwagę dobro użytkownika. Obiektywnie trzeba jednak stwierdzić, że bardzo mało wiemy o oczekiwaniach użytkownika wobec katalogów bibliotecznych, o tym w jakim stopniu proponowane przez bibliotekarzy rozwiązania są akceptowane, czy użytkownik jest zadowolony z poszukiwań, co jest powodem niepowodzenia w poszukiwaniach. Nie prowadząc odpowiednich badań pracownicy katalogów rzeczowych mogą odwołać się tylko do intuicji, to jest pewnego rodzaju przeświadczenia, że przyjęte rozwiązania są najbardziej korzystne dla użytkowników katalogów.

W ostatnich latach w Zakładzie Bibliotekoznawstwa Instytutu Książki i Czytelnictwa Biblioteki Narodowej podjęto takie badania^{3/}, a ich częścią była m.in. ankieta (tu podajemy tylko wyniki) przeprowadzona wśród użytkowników katalogu przedmiotowego w Bibliotece Uniwersytetu Łódzkiego (52 osoby), katalogu przedmiotowego Książnicy Miejskiej w Toruniu (57 osób) oraz wśród studentów Instytutu Bibliotekoznawstwa i Informacji Naukowej UW (29 osób). Ankiety przeprowadzono więc wśród 138 osób - rzeczywistych i potencjalnych użytkowników katalogów przedmiotowych. Zawierała ona 22 pytania sprawdzające akceptację dla:

- a) form rozwiniętych i skrótów stosowanych jako tematy, głównie korporatywne,

^{3/} Wyniki pierwszego etapu badań zob. T. Turowska, K. Malicka: Wykorzystanie katalogów rzeczowych w naukowych bibliotekach uniwersalnych "Przegląd Biblioteczny" 1990 z. 3-4 (w druku).

- b) form polskich i obcych stosowanych głównie jako tematy osobowe,
- c) form prostych i zinwertowanych stosowanych jako tematy toponimiczne,
- d) haseł prostych i rozwiniętych wyrażających przedmiot,
- e) haseł adekwatnych i uogólnionych w stosunku do przedstawionej treści.

Nie traktujemy otrzymanych wyników jako ostatecznych, ani też nie wyciągamy z nich daleko idących wniosków, natomiast stwierdzamy pewne fakty, które powinny podlegać dalszej weryfikacji.

Akceptacja dla form rozwiniętych i skróconych

Pytania o formy rozwinięte i skrócone dotyczyły trzech przypadków tematów korporatywnych i jednego ogólnego. Niżej podajemy ogólnie procentowe wskaźniki obrazujące akceptację dla każdej formy (Suma nie równa się 100%, ponieważ można było udzielić więcej niż jedną odpowiedź):

1. Organizacja Narodów Zjednoczonych	76,8
ONZ	47,1
2. Polska Akademia Nauk	85,5
PAN	37,7
3. Organizacja Jedności Afrykańskiej	68,8
OJA	14,5
4. Niezidentyfikowane obiekty latające	56,5
UFO	44,9

Przedstawione przykłady wskazują, że nawet tak znane skróty jak ONZ i PAN mają niższe notowania niż formy rozwinięte. Być może powodem takiego wyboru są negatywne doświadczenia użytkowników wyniesione z poszukiwań w katalogach, w których skróty bywają szeregowane niejednolicie.

"Instrukcja" zaleca stosowanie form rozwiniętych, przyjmując skróty tylko wyjątkowo dla nazw obcych, których forma skrócona jest

bardziej znana i częściej używana niż forma rozwinięta oryginalna lub jej polski odpowiednik, np. UNICEF, UNESCO, UNRRA.

Akceptacja dla form polskich i obcych

Pytania o formy polskie i obce dotyczyły nazw osobowych świętych, panujących i starożytnych oraz jednego określnika w formie łacińskiej.

1. Marek, święty	81,9
Marcus, sanctus	9,4
2. Owidiusz	83,3
Publius Ovidius Naso	14,5
3. Elżbieta I, królowa Anglii	81,8
Elizabeth I, Queen of England	4,3
4. Biblia - Stary Testament	55,8
Biblia - Testamentum Vetus	0,7

Otrzymałe odpowiedzi wskazują, że użytkownicy zdecydowanie akceptują formy polskie. Takie też formy zaproponowano w "Instrukcji", wprowadzając formy polskie zamiast łacińskich dla nazw osobowych starożytnych oraz dla części Biblii. Pozostałe typy nazw już wcześniej funkcjonowały w metodyce opracowania przedmiotowego w postaci spolszczonej.

Akceptacja dla form prostych i zinwertowanych

Pytania o formy proste i zinwertowane dotyczyły tematów toponimicznych. Pierwotnie w "Instrukcji" zakładano przyjęcie formy prostej jako bliższej językowi naturalnemu. Taką też formę stosują niektóre katalogi (np. katalog Biblioteki Uniwersyteckiej w Warszawie). Ostatecznie zdecydowano się na formę zinwertowaną, zakładając, że z taką formą styka się użytkownik w encyklopediach ogólnych i słownikach fachowych. Wątpliwości, które pojawiły się w trakcie opracowywania "Instrukcji" zostały w pewnej mierze potwierdzone przez badaną grupę. Żadna z form nie została zaakcep-

towana w sposób zdecydowany, aczkolwiek ostatecznie kilkunastoprocentową przewagę zyskały formy zinwertowane.

1. Kanał Sueski	53,6
Sueski Kanał	50,0
Suez - Kanał	10,9
2. Puszcza Białowieska	54,3
Białowieska Puszcza	52,9
Białowieża - Puszcza	11,6
3. Morze Bałtyckie	44,2
Bałtyckie Morze	39,8
Bałtyk	27,5

Sposób wyrażania przedmiotu

W "Instrukcji" przedstawiono dwa sposoby wyrażania przedmiotu: sposób bezpośredni, to jest za pomocą samego tematu i pośredni za pomocą tematu i określnika. Zalecano korzystanie ze sposobu pierwszego, proponując ograniczanie w miarę możliwości sposobu drugiego. Oto jak wyglądały odpowiedzi:

1. Mieszkania - spółdzielczość	28,3
Spółdzielczość mieszkaniowa	73,2
2. Studenci - teatr - Polska	24,6
Teatr studencki - Polska	75,4
Teatr studencki polski	23,2
3. Kultura - historia - Polska	30,4
Polska - kultura - historia	35,5
Kultura polska - historia	50,0

Przykłady te w zasadzie potwierdzają przyjętą w "Instrukcji" regułę, aczkolwiek dziwić może niska akceptacja dla tematu Teatr studencki polski; być może takie wyrażenie nie jest odbierane jako samodzielnie funkcjonujące w języku.

Zasada bezpośredniego wyrażania przedmiotu, jeśli istnieje odpowiednia nazwa, została rozciągnięta również na grupę tematów jednostkowych, w tym na tak zwane tematy rzeczowe jednostkowe, do któ-

rych zostały zaliczone przedmioty typu: języki programowania, systemy informacyjne i komputerowe, urządzenia i środki transportu, maszyny, ciała niebieskie itp. Okazało się, że w stosunku do nazw tych przedmiotów użytkownicy nie zachowują się jednakowo. Oto odpowiedzi:

1. BASIC, język programowania	54,3
Języki programowania - BASIC	63,0
2. Jowisz, telewizor - naprawa	12,3
Telewizor - Jowisz - naprawa	38,4
3. Fiat 126p, samochód - obsługa	55,8
Samochód osobowy - Fiat 126p - obsługa	44,9

Jak widać tego typu przedmioty postrzegane są przez nazwę klasową. Odstępstwem jest tu wyższa akceptacja dla tematu Fiat 126p, co jest prawdopodobnie związane z piśmiennictwem - są książki poświęcone poszczególnym markom samochodów, podczas gdy raczej rzadko spotyka się książki poświęcone poszczególnym typom telewizorów, pralek czy maszyn do pisania.

Podobnie zachowali się użytkownicy wobec tematów etnicznych. Do tej pory w metodyce opracowania przedmiotowego (z wyjątkiem kp Biblioteki Gdańskiej PAN) nazwy grup etnicznych były wyrażane za pośrednictwem terytorium z określnikiem ludność. Nie dotyczyło to tych grup etnicznych, które nie były związane z określonym terytorium np. Cyganie, Żydzi. Ponieważ reguła ta nie była jednolita, tzn. niejednakowo traktowano przedmioty tego samego typu, wydało nam się uzasadnione wprowadzenie tematów etnicznych na określenie różnych grup ludności. Pytania dotyczyły tematu Polacy i Ukraińcy. A oto jak wyglądały odpowiedzi:

1. Litwa - ludność polska	87,7
Polska - Ludność - Litwa	11,6
Polacy - Litwa	13,8

2. Ukraina - ludność - etnografia	62,3
Obyczaje - Ukraina	44,2
Ukraińcy - etnografia	21,7
Ludność - etnografia - Ukraina	6,5

Niska akceptacja dla tych tematów jest zaskakująca, tym bardziej że w pytaniach celowo zostały użyte nazwy etniczne (Obyczaje Ukraińców; Polacy na Litwie).

Wyszczególnianie i uogólnianie

Zagadnienie to należy do trudniejszych w metodyce opracowania przedmiotowego i wiąże się m.in. z adekwatnością haseł w stosunku do treści dokumentów. Otrzymane odpowiedzi potwierdzają większą akceptację użytkowników dla haseł adekwatnych, ale wskazują też, że znaczna część użytkowników wybiera równocześnie hasła uogólnione.

1. Treść dokumentu: uprawa czarnej porzeczki

Krzewy jagodowe - uprawa	21,0
Porzeczka - uprawa	27,5
Porzeczka czarna - uprawa	41,3
Czarna porzeczka - uprawa	42,0

2. Treść dokumentu: działalność ONZ

Organizacje międzynarodowe	10,9
ONZ	47,1
Organizacja Narodów Zjednoczonych	76,8

3. Treść dokumentu: działalność OJA

Organizacje międzynarodowe	12,3
Afryka - organizacje	48,5
OJA	14,5
Organizacja Jedności Afrykańskiej	68,8

4. Treść dokumentu: życie i działalność królowej Elżbiety I

Anglia - historia - 16 w.	44,2
Elżbieta I, królowa Anglii	81,8
Elizabeth I, Queen of England	4,3

5. Treść dokumentu: życie i twórczość Owidiusza	
Literatura łacińska - historia - 1 w.	23,2
Publius Ovidius Naso	14,5
Owidiusz	83,3
6. Treść dokumentu: programowanie w języku BASIC	
Języki programowania	13,8
Języki programowania - BASIC	63,8
BASIC, język programowania	54,3
7. Treść dokumentu: naprawa telewizorów kolorowych Jowisz	
Telewizor - naprawa	29,9
Telewizor kolorowy - naprawa	41,3
Telewizor - Jowisz - naprawa	38,4
Jowisz, telewizor - naprawa	12,3

Na podstawie przedstawionych przykładów widać, że użytkownicy nie są zainteresowani wyszczególnianiem bezwzględnym. Sytuację tę potwierdza przykład, w którym hasło uogólnione zostało wskazane częściej niż adekwatne:

8. Treść dokumentu: choroby kory mózgowej	
Mózg - choroby	58,0
Kora mózgowa - choroby	52,9

Zjawisko to jest trudne do wytłumaczenia zwłaszcza w sytuacji, gdy użytkownik "nie błądzi" po katalogu, lecz ma do wyboru formy hasła adekwatnych w stosunku do treści i uogólnionych. Prawdopodobnie istnieje tu związek między znajomością poszczególnych nazw, ich samodzielnością znaczeniową w ocenie użytkowników a wyborem hasła adekwatnego lub uogólnionego. Zagadnienie to wymagałoby dokładniejszego zbadania, ponieważ należy ono do podstawowych w metodyce opracowania przedmiotowego i ma bezpośredni wpływ na kształt opisów przedmiotowych, a przede wszystkim na opisy wielokrotne.

Podsumowując możemy sformułować następujące wnioski:

- bardziej akceptowane są formy polskie niż obce,
- bardziej akceptowane są formy rozwinięte niż skrócone,

- c) w nazwach geograficznych niewielką przewagę mają formy zinwertowane nad prostymi,
- d) w przypadku nazw ogólnych wyższą akceptację mają hasła proste niż rozwinięte,
- e) ogólnie preferowane są bardziej hasła adekwatne niż uogólnione,
- f) nazwy urzędzeń, środków transportu, języków programowania postrzegane są za pośrednictwem nazwy klasowej,
- g) niska jest akceptacja dla tematów etnicznych,
- h) na ogół nie ma zbyt wielkich rozbieżności w odpowiedziach udzielanych przez respondentów pochodzących z różnych ośrodków, co wskazuje, że przedstawione hasła są mniej więcej jednakowo oceniane,
- i) kierunek rozwiązań metodycznych przyjętych w "Instrukcji" można uznać za właściwy, choć wymagający pewnych modyfikacji w odniesieniu do tematów jednostkowych rzeczowych i etnicznych.

16.05.1991 r.

SOME RULES OF METHODOLOGY OF THE SUBJECT CATALOGUING IN USERS' ESTIMATION

Summary

The are discussed the results of the inquiry about some methodological rules of subject cataloguing presented in "Instrukcja tematowania i katalogu przedmiotowego" (Directions for appointment of subject headings and for cataloguing) by J. Sadowska. The inquiry was carried out among users of the subject catalogue of the Library of Łódź University and of the subject catalogue of

the Public Municipal Library in Toruń, as well as among students of the Institute of Librarianship and Information Science of the Warsaw University. 138 actual and potential users of subject catalogues were put through examination. The main conclusion emerging from the inquiry is that the methodological approach adopted in the "Directions" is appropriate, though some modifications regarding the ethnic and individual subject headings are needed.

НЕКОТОРЫЕ ПРАВИЛА МЕТОДИКИ ПРЕДМЕТИЗАЦИИ В ОЦЕНКЕ ПОТРЕБИТЕЛЕЙ

Р е з ю м е

В статье обсуждены результаты анкетного исследования выбранных правил методики предметизации содержания в "Instrukcji tematowania i katalogu przedmiotowego" Я. Садовской. Исследование было проведено среди потребителей предметного каталога Библиотеки Лунзкого университета, потребителей предметного каталога Городской массовой библиотеки в Торуне, а также среди студентов Института библиотекосведения и научной информации Варшавского университета. Исследования были проведены среди 138 потребителей предметных каталогов. Главным выводом истекшим из проведенного исследования является констатация, что нахождение принятых методических решений можно признать соответствующим, хотя оно требует некоторых модификаций.

JACEK ŚLADEWSKI

Biblioteka Narodowa

PROBLEMY METODYKI TEMATOWANIA W BIBLIOTEKACH UNIWERSALNYCH

Porównanie sposobów opracowania przedmiotowego piśmiennictwa w trzech uniwersalnych bibliotekach naukowych: Bibliotece Uniwersyteckiej w Warszawie, Bibliotece Gdańskiej PAN oraz Bibliotece Narodowej.

Celem pracy jest przedstawienie różnic w opracowaniu przedmiotowym piśmiennictwa w trzech uniwersalnych bibliotekach naukowych: Bibliotece Uniwersyteckiej (BUW) w Warszawie, Bibliotece Gdańskiej Polskiej Akademii Nauk (BG PAN) i Bibliotece Narodowej (BN). Podstawą analizy były opisy przedmiotowe zamieszczone w "Przewodniku Bibliograficznym" z roku 1990 (nr 4 i 9) - korzysta z nich z niewielkimi zmianami katalog przedmiotowy BN. W przypadku BUW i BG PAN podstawą był ogląd opisów przedmiotowych w katalogach. Opisy dotyczyły tych samych dokumentów.

Chodziło o uzyskanie odpowiedzi na następujące pytania: na ile różnią się charakterystyki wyszukiwawcze w tych bibliotekach, a na ile są zbieżne; czy są różnice w nazywaniu przedmiotu (tematy) i wyrażaniu aspektów (określniki); czy charakterystyki te różnią się kompletnością (liczba haseł przedmiotowych). Próbujemy też

"Zagadnienia Informacji Naukowej" 1992 nr 1(59)

znaleźć uzasadnienie dla takiego lub innego wyrażenia treści dokumentu, aczkolwiek nie oceniamy tu adekwatności opisu przedmiotowego w stosunku do treści.

Analiza taka wydaje się interesująca ze względu na to, że badane biblioteki odwołują się w zasadzie do wspólnej metodyki^{1/}. Są to reguły zawarte w podręczniku Adama Łysakowskiego^{2/}, przy czym BG PAN i BN modyfikują bądź uściślają jego wskazania. Jednakże nawet wierne trzymanie się zasad podręcznika nie usuwa konieczności dokonywania pewnych interpretacji i podejmowania samodzielnych decyzji. Powoduje to nieuchronnie narastanie różnic w metodyce opracowania przedmiotowego w poszczególnych bibliotekach. Jednocześnie każda z bibliotek, o których tu mowa, dysponuje własnym słownikiem języka haseł przedmiotowych - jhp - (w formie kartotekowej lub edycyjnej).

Wszystko to sprawia, że charakterystyki przedmiotowe tych samych dokumentów, sporządzane przez różne biblioteki, różnią się znacznie od siebie. Jedną z prób przezwyciężenia przynajmniej części z tych różnic jest przygotowanie nowej instrukcji tematowania i katalogu przedmiotowego^{3/}.

Przedstawiony tu materiał ilustruje tylko niektóre problemy metodyczne. Pełniejsza analiza porównawcza wymagałaby bardziej pogłębionych badań. W artykule pokazano jedynie wybrane przykłady różnic w opracowaniu tych samych dokumentów. Można uznać, że

^{1/} Czarnecka J.: Katalog przedmiotowy w dużych bibliotekach uniwersalnych (na podstawie ankiety). W: Opracowanie rzeczowe zbiorów w dużych bibliotekach uniwersalnych. Warszawa 1989.

^{2/} Łysakowski A.: Katalog przedmiotowy. Podręcznik. Warszawa 1946.

^{3/} Instrukcja tematowania i katalogu przedmiotowego. Oprac. J. Sadowska

dobór materiału jest reprezentatywny, gdyż do porównań przyjęto wszystkie opisy z dwóch zeszytów "Przewodnika Bibliograficznego", z wyjątkiem działu "Literatura piękna", gdzie nie należy się spodziewać większych różnic w opisie przedmiotowym.

Pierwszą rzeczą, która zwraca uwagę podczas przeglądania zebranego materiału, jest niejednakowa liczba haseł przedmiotowych użytych do opisu tego samego dokumentu. Będzie jeszcze o tym mowa dalej. Istnieją też liczne różnice w sposobie wyrażania przedmiotu głównego. Poniższy przykład dobrze ilustruje oba te zjawiska.

Zbigniew Skorny: "Mechanizmy regulacyjne ludzkiego działania"

BUW: Czynność

Psychologia społeczna - monografie

Osobowość - psychologia

BG PAN: Działanie - psychologia

BN: Motywacja

Zachowanie

Za główny przedmiot dokumentu w dwóch przypadkach uznano zjawisko, które otrzymało dwie różne nazwy: "Czynność", "Działanie". Jednakże w trzecim przypadku za przedmiot uznane zostało "Zachowanie", co świadczy już nie o różnicach językowych, ale o odmiennym wyniku analizy przedmiotowej, o odmiennej ocenie treści dokumentu (w słowniku jhp BN^{4/} znajduje się bowiem temat "Czynności"). Można podać jeszcze inny przykład takich rozbieżności:

^{4/}Słownik języka haseł przedmiotowych Biblioteki Narodowej. Oprac. J. Trzcicka i E. Stępnikowa. Warszawa 1989.

Andrzej Góralski: "Twórcze rozwiązywanie zadań"

BUW: Matematyka - zadania i ćwiczenia

BG PAN: Heurystyka

BN: Heurystyka

Również i tu indeksujący różnią się w ocenie tego, jaki jest przedmiot publikacji. Celem tego artykułu nie jest jednak, jak już wspomniano, dyskusja na temat adekwatności poszczególnych rozwiązań. Pewne sugestie lub oceny są możliwe tylko wówczas, gdy różnice w opracowaniu przedmiotowym mają charakter wyłącznie (lub przede wszystkim) językowy, a zatem nie wynikają z odmiennego rozumienia tekstu, lecz z odmiennego nazwania jego treści.

Wśród tych różnic zwraca uwagę dowolność użycia liczby pojedynczej lub mnogiej. Mamy więc, w zależności od biblioteki, tematy: "Grupa społeczna" lub "Grupy społeczne", "Czasownik rosyjski" i "Czasowniki rosyjskie". Nie zmienia to oczywiście znaczenia tematu, ale w jakiś sposób utrudnia stworzenie wspólnej bazy danych dla wielu bibliotek i korzystanie z niej. Niestety, nie ma, i chyba nie może być, ścisłych reguł stosowania w jhp liczby pojedynczej lub mnogiej. Można jedynie postulować wprowadzenie większej liczby odsyłaczy od form odrzuconych w danym katalogu, a przyjętych w innych.

Inne różnice w formie tematów także nie muszą zmieniać ich zakresów. "Wojna 1939-45" (BUW), "Wojna 1939-45 światowa" (BG PAN) i "Wojna 1939-1945 r." (BN) stanowią wyrażenia synonimiczne, podobnie jak "Kieleckie" (BUW) i "Kielce, okręg" (BN). Za synonimy można też uznać takie pary tematów, jak: "Katechizm katolicki" (BUW i BN) i "Katechizm rzymsko-katolicki" (BG PAN), "Topologia" (BG PAN i BN) i "Topologia ogólna" (BUW).

Zdarza się jednak, że ten sam przedmiot otrzymuje w różnych bibliotekach różne nazwy, o wyraźnie innych odcieniach znaczeniowych. Obok podanego już przykładu ("Czynność" i "Działanie") w katalogach występują np.: "Radiestezja" (BUW) i "Różdzkarstwo" (BN), "Metempsychoza" (BUW) i "Reinkarnacja" (BG PAN i BN).

Innego typu różnice ilustruje następujący przykład:

Aleksander Krzemieniecki: "Tabor kolejowy"

BUW: Kolejy żelazne - tabor kolejowy

BG PAN: Kolejnictwo - tabor - Polska - XX w.

BN: Tabor kolejowy - podręcznik szkolny

Jak widać, przedmiot może być wyrażony za pomocą samego tematu lub tematu z określnikiem. Przy okazji warto zwrócić uwagę na zjawisko redundancji występujące w haśle przedmiotowym BUW. Powoduje ono niepotrzebne wydłużenie zapisu. Oto kilka przykładów, w których określnik zawiera te elementy treści, które wystąpiły już w temacie: .

Przestępstwo - zwalczanie i zapobieganie przestępczości

Prawo autorskie - ochrona praw autorskich

Radio - słuchowiska radiowe

Środowisko naturalne - badania środowiska

Przykłady te pochodzą z katalogu BUW. J. Sadowska podaje, że tego typu określniki, zwane przez nią opisowymi, są częste w metodyce radzieckiej^{5/}. Określnik opisowy nie musi zresztą odznaczać się nadmiarowością, jak choćby w haśle z katalogu BG PAN: "Mickiewicz Adam - wpływ na literaturę czeską".

Barczo przejrzystym przykładem różnic w stosowaniu określników jest opracowanie książki Stanisława Piotra Kaczmarskiego i Aliny Wójcik "Z angielskim w podróży":

^{5/}Sadowska J.: Metodyka opracowania przedmiotowego w polskich bibliotekach uniwersalnych. "Przegląd Biblioteczny" 1989 nr 3.

BUW: Język angielski - podręczniki

BG PAN: Język angielski - frazeologia

BN: Język angielski - rozmówki

W tej sprawie dokładnych wskazówek udziela wspomniana "Instrukcja..." i do niej należy się odwołać. Ze względu na specyficzną formę dokumentu zaleca się tu stosowanie określnika "rozmówki".

Różnice, o których była mowa do tej pory, dotyczą albo formy tematu, albo określnika, albo formy całego hasła przedmiotowego. Są to w dużej mierze różnice, które dają się usunąć, choć niejednokrotnie wymagałoby to dużych melioracji w katalogach. Pora teraz przyjrzeć się rozbieżnościom znacznie poważniejszym, obejmującym strukturę całego opisu przedmiotowego i wynikającą z odmienności analizy rzeczowej, przeprowadzanej w różnych bibliotekach. Już pierwszy z omawianych w tekście przykładów należy do tej kategorii. Oto inny:

Jiddu Krishnamurti: "Wolność od znanego"

BUW: Życie - filozofia

Wolność - generalia

Introspekcja

Ascetyka

BG PAN: Filozofia życia

BN: Krishnamurti Jiddu

Medytacja

Religia - zagadnienia

Przyjmując, że każda z bibliotek dysponuje zbliżonym, pod względem ilościowym, zasobem leksykalnym słownika jhp, trudno wytłumaczyć aż tak duże zróżnicowanie opisów przedmiotowych. Należy przypuszczać, że w grę wchodzi tu różnice w zrozumieniu tekstu, a także dość arbitralne decyzje co do poziomu szczegółowości, na jakim tekst zo-

stanie opisany. W słowniku BG PAN znajdujemy przecież hasła: "Religia - filozofia" oraz "Filozofia a religia" - żadne z nich jednak nie znalazło zastosowania przy opracowywaniu omawianego dokumentu. Z kolei w opisie sporządzonym przez BN nie uznano za stosowne wskazanie na aspekt etyczny rozważań Krishnamurtiego, ani na ich związek z teorią poznania. Rzecz jasna, tematującego często ograniczają i ukierunkowują w pracy liczne przykłady dawniejszych opisów podobnych tekstów, w grę wchodzi też szereg spisanych i niepisanych decyzji. Innymi słowy - indeksujący nie jest powodowany jedynie ogólną metodyką katalogu i najlepszą wolą zrozumienia tekstu. Powinien jednak pamiętać, że ciąży na nim obowiązek stałej rewizji zastanych rozwiązań i odpowiedzialnego, samodzielnego postępowania.

Rozważmy inny przykład. Książka Rogera Trigga "Rozum a zaangażowanie" została potraktowana następująco:

BUW: Relatywizm

Etyka - generalia

Życie - filozofia

BG PAN: Filozofia a religia

Wittgenstein Ludwig

BN: Nauka a religia

Religia - filozofia

Wydaje się, że słusznie postanowiono w BG PAN uwzględnić w opisie temat osobowy. W języku polskim nie ma zbyt wielu materiałów dotyczących Wittgensteina, a więc każdy większy tekst wart jest odnotowania. Ponadto temat osobowy dostarcza potencjalnemu czytelnikowi ważnej informacji na temat kontekstu, w jakim toczą się rozważania autora, sytuuje książkę w pewnym otoczeniu. Opisy przedmiotowe BUW

i BN, pozbawione tego rodzaju wskazówki, grzęzną w ogólnikowości. Warto może w tym miejscu wspomnieć o pożytecznej praktyce zaopatrywania tematu osobowego w dopowiedzenie w postaci dat. Umieszczone w nawiasie okrągłym daty narodzin i śmierci mogą być cenną wskazówką w sytuacji, gdy samo nazwisko mówi niewiele lub zgoła nic. Dopowiedzenia tego typu stosuje konsekwentnie BUW. Jest to zgodne ze wskazaniami "Instrukcji..."

Na zróżnicowanie charakterystyk przedmiotowych tego samego dokumentu wpływa jeszcze jedna okoliczność. Otóż brak jest, jak dotąd, jednolitych reguł tworzenia opisu wielokrotnego^{6/}. Widać to np. w sposobie opracowania pamiętników. W BN ustalono zasadę, że pamiętniki otrzymują obowiązkowo temat osobowy - nazwisko ich autora. Zasady tej nie uznaje BG. BN stara się też zawsze wyrazić w postaci tematu - przedmiot wspomnień, a więc rodzaj działalności uprawianej przez autora, względnie czas i miejsce akcji. Robi to też na ogół BUW. Sytuacja ta rodzi rozbieżności jaskrawo widoczne w poniższym przykładzie:

Paweł Jasienica: "Pamiętnik"

BUW: Literaci polscy - pamiętniki - 20 w.

Jasienica Paweł (1909-1970)

BG PAN: Pamiętniki polskie - XX w.

BN: Jasienica Paweł

Pamiętniki polskie - 20 w.

Wilno, okręg - historia - 1918-1939 r. - pamiętniki

Rewolucja 1917 r. październikowa i wojna domowa - pamiętniki

^{6/} Stępnikowa E.: Opis wielokrotny w języku haseł przedmiotowych. Ref. wygłoszony na Ogólnopolskiej konferencji nt. Opracowania przedmiotowego dokumentów oraz problemów dydaktycznych opracowania rzeczowego w szkołach wyższych. Gdańsk, 7-9 maja 1991 r.

Reguły opisu wielokrotnego wpływają znacząco na sposób przeprowadzenia analizy przedmiotowej. Jeśli wiadomo, że w niektórych przypadkach nie wystarczy poprzestać na temacie formalnym (typu "Pamiętniki"), wówczas podejście tematującego według takiej reguły będzie całkiem inne niż kogoś, kto owej reguły nie uznaje. Bywa też tak, że jakaś reguła opisu wielokrotnego ma charakter otwarty, a więc stanowi na przykład, iż obok tematu formalnego może, ale nie musi pojawić się hasło dotyczące treści dokumentu. Dotyczy to m.in. tematów "Literatura dziecięca" i "Literatura młodzieżowa", według reguł przyjmowanych w BN. Trzeba wówczas najpierw zdecydować, czy treść jest warta wyrażenia, a dopiero potem przystąpić (lub nie) do właściwej analizy.

Trzy opisy książki Wojciecha Iwańczaka "Ludzie miecza, ludzie modlitwy i ludzie pracy" wykazują liczne rozbieżności zarówno formalne, jak i treściowe:

BUW: Historiozofia - Czechosłowacja

Czechosłowacja - historia średn. - generalia

BG PAN: Społeczeństwo - historiografia czeska - XIV-XV w.

BN: Husytyzm - zagadnienia

Literatura czeska - historia - 14 -15 w.

Struktura społeczna - historia - Czechy - 14-15 w.

Zauważmy, że w opisach tych nie tylko nie występuje żaden wspólny temat, ale nie ma też w nich choćby dwóch haseł, które można by uznać za synonimiczne. Wspólnym elementem trzech opisów pozostaje tylko geograficzny i czasowy zasięg charakteryzowanego tekstu, przy czym w katalogu BUW jest on określony mocno nieprecyzyjnie. Łatwo też dostrzec charakterystyczne, widoczne już wcześniej zjawisko: opisy BG PAN są najkrótsze, wyraźnie różnią się pod tym względem od pozostałych.

Odnosi się to zarówno do liczby haseł przedmiotowych, jak i bezwzględnej długości opisu, wyrażonej w liczbie wszystkich jednostek leksykalnych.

Opis dłuższy to z pewnością opis bardziej precyzyjny, o większej wartości informacyjnej, jeśli pominąć zdarzające się tu i ówdzie zjawisko redundacji. Katalogi wykorzystujące opisy wielokrotne są sprawniejszym narzędziem informacyjno-wyszukiawczym. Z drugiej strony zbyt wielka liczba haseł utrudnia ich wykorzystanie, toteż na ogół ogranicza się ją do pięciu - sześciu. W większości przypadków adekwatny opis wymaga więcej niż jednego tylko hasła. Tematując możemy bowiem popełnić dwa rodzaje błędów: nadmiernie uogólnić treść dokumentu lub zbyt szczegółowo wyliczać jej składniki. Na ten pierwszy błąd, chyba poważniejszy, narażeni są ci, którzy najkrótsze opisy uważają za najlepsze.

Czasem jednak lapidarność bywa zaletą. Książkę Teodora Szymanowskiego "Powrót skazanych do społeczeństwa" opisywano następująco:

BUW: Przesłępstwo - zwalczanie i zapobieganie przestępczości
Prawo karne wykonawcze

BG PAN: Więźniowie - resocjalizacja - Polska - XX w.

BN: Adaptacja społeczna - zagadnienia
Kara pozbawienia wolności - socjologia
Opieka społeczna - zagadnienia
Pedagogika resocjalizacyjna - zagadnienia

Osobliwością metodyki BG PAN jest możliwość tworzenia tematów w formie nazw grup ludności. W powyższym przykładzie temat tego typu został użyty bardzo trafnie. Trudno natomiast oprzeć się wrażeniu, że mnogość haseł przedmiotowych w opisie BN wynika z niemożności

wyrażenia treści wprost i z konieczności "obejścia" przedmiotu kilkoma hasłami. Świadczy o tym m.in. powtarzający się określnik "zagadnienia". Ujednolicając słownictwo opisu przedmiotowego, trzeba odpowiedzieć sobie na pytanie o formę tematu i zdecydować, czy lepiej używać formy "Polacy" (BG PAN) czy "Polska - ludność" (BN), czy pisać "Adwokaci" czy też "Adwokatura - zawód", czy wreszcie zaakceptować temat "Więźniowie" czy "Kara pozbawienia wolności" z odpowiednimi określnikami.

Zwróćmy na koniec uwagę na pożytek płynący z szybkiego i elastycznego reagowania na nowe zjawiska w piśmiennictwie. Chodzi o tworzenie nowych tematów jednostkowych i chronologicznych. Oto jak opracowano "Porozumienia Okrągłego Stołu: Warszawa, 6 lutego - 5 kwietnia 1989 r."

BUW: Polska - historia - od 1944 (źródła)

BG PAN: Polityka - Polska - XX w. - źródła

Gospodarka - reforma - Polska - XX w. źródła

BN: Konferencja Okrągłego Stołu, Warszawa 1989 r.

Z całą pewnością wprowadzenie przez BN nowego tematu chronologicznego jest słuszne. Literatura na temat tego wydarzenia historycznego, już dziś pokaźna, będzie w dalszym ciągu narastać. Opisy sporządzone w BUW i BG PAN są stanowczo zbyt ogólne, a więc nieadekwatne. Użytkownikowi katalogu pozostaje wówczas wnioskowanie o treści dokumentu na podstawie tytułu.

Podsumowując, trzeba stwierdzić, że zebrany materiał ujawnił liczne i daleko idące różnice w opracowaniu przedmiotowym. Tylko w dwóch przypadkach na sześćdziesiąt zbadanych wszystkie trzy opisy przedmiotowe były jednakowe (jeśli pominąć formę gramatyczną określnika). Zgoda dotyczyła przy tym dokumentów najłatwiej-

szych w opracowaniu: podręcznika języka angielskiego i atlasu anatomicznego. W pozostałych przypadkach różnice dotyczyły zarówno słownictwa, jak i gramatyki języka haseł przedmiotowych. Ujawniały się one już na poziomie pojedynczego hasła. Na poziomie całego opisu zróżnicowanie spowodowane było po większej części odmienne przeprowadzaną analizą przedmiotową, na którą z kolei mają często wpływ reguły opisu wielokrotnego, różniące się od siebie w poszczególnych bibliotekach. Jednakże owe reguły nie tłumaczą aż tak dużego zróżnicowania charakterystyk tego samego dokumentu.

W każdej z bibliotek spotyka się interesujące rozwiązania, które warto byłoby upowszechnić. Mogłyby one funkcjonować w ramach jednolitej, powszechnie obowiązującej metodyki tematowania i katalogu przedmiotowego. Stworzenie takiej metodyki, a w konsekwencji usunięcie większości różnic w sporządzanych przez biblioteki opisach, jest dalekosiędnym celem, przed którym stoją teoretycy i praktycy opracowania rzeczowego w Polsce. Bez jego realizacji trudno sobie wyobrazić stworzenie ogólnego systemu wymiany danych na temat treści dokumentów, systemu opartego na języku naturalnym.

24.09.1991 r.

METHODOLOGICAL PROBLEMS OF APPOINTMENT OF SUBJECT
HEADINGS IN UNIVERSAL LIBRARIES

Summary

There are presented the differences in subject cataloguing of literature in three universal libraries: in the University

Library in Warsaw (BUWD), in the Gdańsk Library of the Polish Academy of Sciences (BG PAN) and in the National Library. The basis of the analysis were subject entries published in the "Bibliographic Guide" which are used by the National Library as well as the review of subject entries in the University Library in Warsaw and in the Gdańsk Library. All entries from two issues of "Przewodnik Bibliograficzny" were adopted for comparison except for the division "Belles-lettres". The collected material showed numerous and far-reaching differences in subject cataloguing - only in two of sixty analysed cases the all three entries were identical. In other cases differences concerned both vocabulary and grammar of the language of subject headings and they appeared on the level of individual entry.

ПРОБЛЕМЫ МЕТОДИКИ ПРЕДМЕТИЗАЦИИ В УНИВЕРСАЛЬНЫХ БИБЛИОТЕКАХ

Р е з ю м е

Представлены различия в методике предметизации в трёх универсальных библиотеках: в Библиотеке варшавского университета (BUW), в Главской библиотеке Польской академии наук (BG PAN), а также в Национальной библиотеке (BN). Основой анализа являлись предметные описания пометочные в национальной библиографии "Przewodniki Bibliograficzne", а также обзор предметных описаний в Библиотеке Варшавского университета и Библиотеке Польской академии наук. Для сравнения использовались все описания содержания в двух выпусках "Przewodnika Bibliograficznego" за исключением раздела "Художественная литература". Собранный материал выявил многочисленные различия в методике предметизации - лишь в двух случаях из шестидесяти обследованных три проанализированные описания были совпадены. В остальных случаях различия касались так же, как и грамматики языка предметных рубрик.

JADWIGA WOŹNIAK

Instytut Bibliotekoznawstwa
i Informacji Naukowej UW

KARTOTEKA HASEŁ WZORCOWYCH DLA
PUBLICZNIE DOSTĘPNEGO KATALOGU ONLINE
SIECI BIBLIOTEK UNIwersYTETU WARSZAWSKIEGO

Komunikat o stanie prac

Stan zaawansowania prac nad kontrolowaną kartoteką haseł wzorcowych (bibliograficznych i przedmiotowych) i plany na 1992 rok.

Około półtora roku temu rozpoczęto w Bibliotece Uniwersytetu Warszawskiego prace, których celem długofalowym jest utworzenie publicznie dostępnego katalogu online sieci bibliotek UW. Projektowany katalog będzie oferował użytkownikom pewną liczbę kluczy wyszukiwawczych odwzorowujących zarówno cechy bibliograficzne, jak i treściowe dokumentów. Każdej jednostce katalogowej będą przyporządkowane dwa typy haseł - hasło bibliograficzne i hasło przedmiotowe. Podjęto decyzję, iż hasła^y zostaną zorganizowane i będą kon-

^y W tekście używam terminu hasło w znaczeniu "ujednolicony klucz wyszukiwawczy"; inaczej mówiąc, hasła są to klucze wyszukiwawcze o kontrolowanej formie i semantyce. Niekiedy, gdy nie powoduje to niejasności, używam też terminu hasło jako skróconej formy wyrażenia hasło przedmiotowe.

trolowane kartoteką haseł wzorcowych (KHW). Wśród haseł bibliograficznych wyodrębniono: hasła osobowe, hasła korporatywne, hasła typu tytuły ujednoczone (w przypadku anonimów klasycznych) i nazwy serii. Ponadto w KHW znajdzie się klasa wyrażen języka haseł przedmiotowych, wśród których pewną część będą stanowiły tematy osobowe, korporatywne i rzeczowe jednostkowe. Oczywista jest więc potrzeba zapewnienia nie tylko związków między na przykład bibliograficznymi hasłami osobowymi i tematami przedmiotowymi osobowymi, ale i pełnej ich przekładalności. Związki haseł bibliograficznych z hasłami charakterystyki przedmiotowej są tym elementem, który odróżnia kartoteki wzorcowe jako formę organizacji i kontroli leksyki języka informacyjno-wyszukiwawczego od innych form stosowanych w zautomatyzowanych systemach katalogowych.

Z punktu widzenia struktury logicznej KHW jest jednym zbiorem jednostek leksykalnych, w którym odwzorowane są i nazwane relacje zachodzące między różnymi, tj. należącymi do różnych wcześniej wymienionych kategorii, hasłami ujednoczonymi oraz między hasłami i formami odrzuconymi (nie przyjętymi dla języka danego systemu informacyjno-wyszukiwawczego). Użytkownik formułując instrukcję wyszukiwawczą będzie korzystał z zasobu leksykalnego zarejestrowanego w kartotece oraz ze znajdujących się tam również metainformacji o znaczeniach wyrażen i sposobach ich wykorzystania w procesie wyszukiwania informacji.

Pomocą w rozwiązywaniu problemów dotyczących formy i struktury bibliograficznych haseł osobowych, korporatywnych czy tytułów ujednoczonych służył projekt polskiej normy "Hasło opisu bibliograficznego" oraz wytyczne IFLA (na przykład, "Guidelines for authority and reference entries / recommended by the Working Group

on an International Authority System; approved by the Standing Committees of the IFLA Section on Cataloguing and the IFLA Section on Information Technology.- London: IFLA International Programme for UBC, 1984).

Trudniejsze zadanie przypadło osobom zajmującym się ujednolicaniem haseł dotyczących opisu przedmiotowego, gdyż ta problematyka nie była do tej pory przedmiotem prac normalizacyjnych w aspekcie tworzenia kartotek wzorcowych^{2/}. W krajach takich jak na przykład USA, Kanada czy Francja biblioteki rozwiązują ten problem sięgając do kartotek autorytatywnych, czyli narodowych kartotek wzorcowych, i po dokonaniu niezbędnych a dozwolonych modyfikacji przejmując je na swój użytek. W Polsce nie ma kartotek autorytatywnych, do tworzenia których upoważnione są i zobowiązane biblioteki narodowe. Praktycznie nie ma też dostępu do kartotek autorytatywnych innych państw. Program opracowania kartoteki wzorcowej haseł przedmiotowych dla sieci bibliotek Uniwersytetu Warszawskiego i powiązania jej z kartotekami bibliograficznych haseł wzorcowych w jedną, logiczną kartotekę sterującą procesami wyszukiwania w katalogu online jest więc przedsięwzięciem pionierskim w skali polskiego bibliotekarstwa; może stać się pierwszym i co za tym idzie modelowym rozwiązaniem dla pewnej grupy bibliotek sieci krajowej.

^{2/} Przyjęto się używanie terminu kartoteka haseł przedmiotowych (przez analogię do "kartoteki haseł osobowych", "kartoteki haseł korporatywnych" etc.) na określenie kartoteki haseł wzorcowych służących do odwzorowywania cech treściowych katalogowanych dokumentów, chociaż termin ten nie wydaje się poprawny, gdyż sugeruje, iż do kartoteki wprowadza się jedynie hasła przedmiotowe, podczas gdy przede wszystkim są wprowadzane jednostki leksykalne języka haseł przedmiotowych, tzn. tematy i określniki, zaś hasła przedmiotowe (czyli zdania JHP) w znacznie mniejszej i ograniczonej merytorycznie liczbie przypadków.

Kartoteka wzorcowa języka haseł przedmiotowych rozpatrywana w izolacji od systemu informacyjno-wyszukiwawczego jest jednym z typów słowników kontrolujących zasób leksykalny JHP. W skali całego systemu - jest jednym z elementów podsystemu wyszukiwania. Tam, gdzie jest to możliwe i uzasadnione, hasła są powiązane z hasłami ujednoliconymi z innych kartotek wzorcowych, tworząc wirtualny ciąg form związanych. Z punktu widzenia użytkownika wywołanie którejkolwiek formy uruchamia konsultowane z nim wyszukiwanie z zastosowaniem form pozostałych. Charakterystyki wyszukiwawcze dokumentów zakodowane w języku haseł przedmiotowych mogą być podstawą przeszukiwania zbioru za pomocą instrukcji wyrażonych w JHP, ale także po odpowiednim ich posortowaniu będą także mogły być wykorzystane do szukania za pomocą słów kluczowych.

Zamierza się wprowadzić istotne zmiany dotyczące odwzorowania relacji paradygmatycznych w kartotece haseł wzorcowych JHP. Będą nazwane i przedstawione następujące rodzaje relacji: relacja ekwiwalencji wyszukiwawczej (łącząca odrzucone formy haseł z formami przyjętymi dla systemu), relacje hierarchiczne (generyczna i w uzasadnionych przypadkach mereologiczna) oraz wybrane relacje kojarzeniowe. Pod wieloma względami organizacja i prezentacja leksyki będzie bliższa rozwiązaniom i postaci tezaurusa niż tradycyjnym słownikom tematów i określników.

Zaplanowano także prace nad określnikami. Polegać one będą z jednej strony na stworzeniu jawnej listy określników i wykazaniu ich w porządku abecedowym kartoteki wraz z informacjami dotyczącymi powiązań (relacji) i stosowania, z drugiej strony - na opracowaniu pomocy metodycznych dla osób zajmujących się przedmiotowaniem, czyli wykazów określników typowych dla różnych kategorii i podkategorii tematów.

Do końca 1991 roku wykonano następujące zadania:

* sformułowano zasady tworzenia kartotek wzorcowych dla publicznie dostępnego katalogu online sieci bibliotek UW oraz format definiujący strukturę i zawartość informacyjną rekordów w tych kartotekach;

* rozpoczęto prace nad koncepcją i kartotekami określników oraz ujednoliconych haseł bibliograficznych osobowych i korporatywnych.

W 1992 roku planuje się kontynuowanie rozpoczętych prac. Wiele wskazuje na to, że w proces realizacji projektu opracowanego w BUW włączą się żywo nim zainteresowane, biblioteki szkół wyższych, zwłaszcza Biblioteka Uniwersytetu Jagiellońskiego, Biblioteka Akademii Górniczo-Hutniczej oraz Biblioteka Uniwersytetu Gdańskiego. Gotowość współpracy choćby trzech nowych zespołów będzie mieć znaczący wpływ na tempo realizacji projektu. Do tej pory zajmowały się nim bowiem dwie osoby zatrudnione w Sekcji Środków Lingwistycznych Oddziału Zastosowań Informatyki BUW pracujące pod kierunkiem pani mgr Teresy Głowackiej oraz niewielka liczba osób angażowanych dorywczo dzięki środkom finansowym otrzymanym z Ministerstwa Edukacji Narodowej w formie tzw. funduszu badawczego i dofinansowania ogólnotechnicznego.

10.02.1992 r.

THE INDEX OF THE PATTERN HEADINGS FOR THE WARSAW UNIVERSITY
LIBRARIES' PUBLICLY AVAILABLE ONLINE CATALOGUE
the state of the art

Summary

The Warsaw University Library started the works on the establishment of the publicly available online catalogue of the library network at the University. Headings used in the discussed catalogue will be supervised through out the index of pattern headings. There is presented the state of the advancement of works by the end of 1991 and the plans for 1992.

КАРТОТЕКА ОБРАЗЦОВЫХ ЗАПОЛНОК ДЛЯ ОБЩЕДОСТУПНОГО ДИАЛОГОВОГО КА-
ТАЛОГА СЕТИ БИБЛИОТЕК ВАРШАВСКОГО УНИВЕРСИТЕТА

Сообщение о состоянии работ

Р е з ю м е

Библиотека Варшавского университета начала работы по созданию общедоступного диалогового каталога сети Библиотек Университета. Заголовки этого каталога будут проверяться на основе картотеки образцовых (библиографических и предметных) записок. Обсуждено состояние работ на конец 1991 года, а также планы на 1992 год.

HENRYK SZARSKI

Biblioteka Główna i OINT
Politechniki Wrocławskiej

NOWA METODA DOSTĘPU DO LITERATURY NAUKOWEJ - ADONIS

Zasady działania programu ADONIS. Wyniki eksperymentu. Zamierzenia na przyszłość.

Kłopoty finansowe bibliotek w Polsce powodują systematyczne ograniczanie zakupów literatury, w tym przede wszystkim czasopism naukowych. Utrudnia to, w sposób oczywisty, pracownikom naukowym dostęp do publikowanych w nich na bieżąco informacji o wynikach najnowszych badań. Podobne problemy występują również w krajach znacznie zamozniejszych. Co prawda, tamtejsze biblioteki dysponują większym budżetem, lecz jest on niewystarczający na gromadzenie całej liczącej się literatury naukowej, przede wszystkim w dziedzinach rozwijających się tak szybko, jak biologia, inżynieria materiałowa, informatyka, elektronika. Ogromne zapotrzebowanie na dostęp do literatury źródłowej skłania do poszukiwania nowych i oryginalnych rozwiązań technicznych i organizacyjnych pozwalających sprostać wymaganiom użytkowników.

Article Delivery Over Network Information System - ADONIS jest to program umożliwiający praktyczne przetestowanie nowych metod publikowania i rozpowszechniania czasopism naukowych. Wyko-

rzystanie dysków optycznych (CD-ROM - Compact disc - read only memory) winno poprawić szybkość i jakość dostarczanych dokumentów przy mniejszych kosztach niż stosowane dotychczas kopiowanie na kserografach. Zaoszczędzone w ten sposób pieniądze trafią w części do wydawców czasopism dla wyrównania strat wynikających z ewentualnego zmniejszenia się zapotrzebowania na drukowane wersje czasopism.

Koszty eksperymentu ponoszą międzynarodowi wydawcy, jak: Blackwell Scientific Publications, Butterworth Scientific, Churchill Livingstone Medical Journals, Elsevier Science Publishers, C.V. Mosby, Mungsgaard International Publishers, Pergamon Journals, Springer Verlag, Georg Thieme Verlag, John Wiley oraz 13 ośrodków w Europie, USA, Meksyku, Australii i Japonii, zajmujących się na dużą skalę dostarczaniem kopii dokumentów źródłowych, w tym przede wszystkim artykułów z czasopism naukowych.

Wśród ośrodków europejskich znajdują się:

- Biblioteka Brytyjska, Boston Spa,
- Centrum Dokumentacji Naukowej i Technicznej (CDST), Paryż,
- Biblioteka Medyczna, Kolonia,
- Królewska Akademia Nauk (KNAW), Amsterdam,
- Instytut Naukowo-Technologiczny (ICYT), Madryt,
- Królewski Instytut Medycyny, Sztokholm.

Na potrzeby ADONIS wybrano 219 czasopism biomedycznych i biochemicznych wydanych w latach 1987-1988. Kierowano się przy tym dużym zapotrzebowaniem na artykuły z tego zakresu oraz faktem, iż zamówienia na kopie dotyczą na ogół artykułów pochodzących z czasopism z ostatnich trzech lat od momentu wydania.

Zasady działania

Zasady działania programu ADONIS są następujące.
Co tydzień zawartość uwzględnionych w ADONIS czasopism indeksowana

jest (tylko dane bibliograficzne bez indeksu rzeczowego) przez Excerpta Medica w Amsterdamie. Indeksowane są nie tylko wszystkie artykuły, ale także streszczenia, listy do redakcji oraz stałe rubryki redakcyjne. Nie podlegają indeksowaniu ogłoszenia, reklamy i spisy treści.

Po zaindeksowaniu każda pozycja otrzymuje opisany poniżej numer identyfikujący ją jednoznacznie w systemie ADONIS.

Następnie indeksy, w postaci czytelnej maszynowo (kod ASCII), przesyłane są do firmy SCANMEDIA LTD w Anglii i wprowadzane do komputera. Jednocześnie do komputera wprowadza się pełną treść wszystkich zaindeksowanych artykułów. Jest to robione za pomocą specjalnych urządzeń, tzw. scannerów, umożliwiających odczytywanie, kodowanie i przesyłanie do pamięci maszyny całych stron tekstu łącznie z zamieszczonymi tam rysunkami i zdjęciami.

W dalszej kolejności całość danych (teksty artykułów i indeksy) przepisywana jest na dysk optyczny (tzw. dysk wzorcowy) w firmie Philips and Dupont Optical Company w Hannoverze. W tej samej firmie powstaje odpowiednia liczba kopii wykonanych na podstawie dysku wzorcowego. Dyski te przesyłane są co tydzień do bibliotek i ośrodków uczestniczących w eksperymencie. Tam indeks zbiorczy uzupełnia się danymi z ostatniego dysku. Tak przygotowane dyski są już gotowe do realizacji zamówień napływających od użytkowników.

Dyski dostępne są w bibliotekach już po 4 tygodniach od otrzymania czasopisma w Amsterdamie, tzn. wcześniej niż różnego rodzaju systemy informacyjne (np. DIALOG, ESA-IRS, STN) przekażą informację o treści artykułów i spowodują tym samym napływ zamówień na kopie konkretnych artykułów.

Każdy artykuł indeksowany w ADONIS otrzymuje specjalny numer identyfikacyjny. Składa się on z numeru ISSN (8 znaków) rozszerzone-

go o dwie ostatnie cyfry roku publikacji, kolejnego numeru dla każdego artykułu (5 znaków), poczynając od "1" na początku roku, oraz jednoznakowego numeru kontrolnego. Łączna długość numeru ADONIS wynosi zatem 16 znaków, co pozwala na rozróżnienie 99 999 artykułów i innych materiałów publikowanych każdego roku w danym czasopiśmie.

Do wyszukiwań wykorzystywać można dowolne elementy opisu bibliograficznego (np. autor, tytuł), a także numer artykułu stosowany w ADONIS. Można również oznaczać pozycje wybrane do drukowania. Wyszukiwanie trwa średnio 8 sekund (po załadowaniu dysku), a drukowanie pierwszej strony - 20 sekund. Następne strony drukowane są z szybkością 8 stron na minutę.

Wyniki eksperymentu

W trakcie dwóch lat zaindeksowano ogółem 199 440 artykułów, streszczeń, listów do redakcji itp. z 219 czasopism wydanych w latach 1987-1988. Całość zapisana została na 84 dyskach CD-ROM. Dyski te były wykorzystywane przez 13 ośrodków z 10 krajów w celu dostarczenia kopii dokumentów źródłowych. Łącznie wydrukowano ponad 50 tys. zamówionych pozycji. Zrealizowane zostały dwa główne cele eksperymentu. Wykazano mianowicie, że koszty dostarczania dokumentów za pomocą CD-ROM są niższe niż przy konwencjonalnych metodach wykonywania kopii (przy utrzymaniu ich wysokiej jakości) oraz potwierdzona została większa, w porównaniu z innymi rozwiązaniami, przydatność numeru identyfikacyjnego ADONIS na potrzeby wyszukiwania dokumentów.

W ostatnich miesiącach trwania eksperymentu, w wyniku usprawnienia programów wyszukiwawczych, uzyskano oszczędności około 50%. Składają się na to przede wszystkim oszczędności czasu pracy ludzkiej (przy jej jednoczesnym ułatwieniu), miejsca w magazynach (1,1 m półki na ustawienie 84 dysków wobec około 40 m dla przechowywania czasopism oryginalnych), a także wyeliminowanie kosztów oprawy czasopism. Ważne jest również, że dyski mogą być przechowywane na miejscu kopiowania, podczas gdy czasopisma muszą być na ogół dostarczane z mniej lub bardziej oddalonego magazynu. Zaproponowany w ADONIS 16-znakowy numer identyfikacyjny okazał się bardziej przydatny do wyszukiwania dokumentów w systemach online. Jednocześnie można oczekiwać, że wydawcy czasopism wprowadzą numer ADONIS dla każdego artykułu już na etapie drukowania czasopism.

Zamierzenia na przyszłość

Rezultaty eksperymentu ADONIS uzasadniają uruchomienie od początku 1991 r. normalnych usług opartych na warunkach handlowych. Wprowadzono przy tym wiele ulepszeń technicznych i programowych, w tym m.in.:

- zwiększenie pojemności dysków, tak aby pozwalały pomieścić liczbę stron tekstu znacznie większą niż dotychczas (5-7 tys. stron);
- poprawę jakości obrazów wyświetlanych na monitorach wykorzystywanych na potrzeby ADONIS;
- wprowadzenie udoskonalonych czytników wielokrotnionych (tzw. jukebox), pozwalających na odczyt dużej liczby dysków (opracowano już czytniki obsługujące 240 dysków) przy jednoczesnym zmniejszeniu do 4 sekund średniego czasu dostępu do informacji;
- udoskonalenie programowania dla zwiększenia szybkości wyszukiwań oraz wydruku.

Wiele instytucji wystąpiło o udział w pracach nad dalszą eksploatacją ADONIS. Spośród licznych ofert wybrano następujące:

Excerpta Medica (Amsterdam) - indeksowanie czasopism,
Satz Rechnen Zentrum (Berlin) - wprowadzanie danych do komputera,
Nimbus Information System (Monmouth) - produkcja dysków,
LASEC (Berlin) - prace nad oprogramowaniem.

Tematyka czasopism dotychczas uwzględnionych w ADONIS dotyczyła przede wszystkim medycyny (47%), biologii (19%), farmakologii (16%) i chemii (11%) - wg klasyfikacji stosowanej w Ulrich International Periodicals Directory. Selekcji czasopism dokonano głównie na podstawie analizy prenumerat składanych przez najważniejsze firmy farmaceutyczne. O ile wybór zakresu tematycznego okazał się trafny, o tyle liczba 219 tytułów czasopism indeksowanych w ADONIS była stanowczo za mała, na co zwracało uwagę wielu użytkowników, szczególnie tych, którzy wywodzili się z dużych ośrodków i laboratoriów badawczych. Z tego powodu zestaw czasopism został powiększony prawie dwukrotnie, obejmując ponad 400 tytułów.

Ośrodki zainteresowane dyskami ADONIS będą mogły je otrzymać wraz z oprogramowaniem do wyszukiwań na podstawie odpowiedniej licencji. Opłaty za roczną prenumeratę dysków (w roku 1991 - ok. 12 000 \$), a także opłaty za rozpowszechnienie kopii będą mniejsze, jeżeli biblioteka posiada czasopisma również w postaci drukowanej^{1/}. Dane o rozpowszechnianiu kopii z poszczegól-

^{1/} Aktualnie żadna z bibliotek w Polsce nie prowadzi usług z wykorzystaniem systemu ADONIS. Zainteresowani mogą uzyskać dodatkowe informacje, pisząc na adres: ADONIS, PO Box 839, 1000AV Amsterdam, Holandia.

nych czasopism rejestrowane będą automatycznie, a opłaty uzgadniane dla każdego czasopisma z jego wydawcą. Dyski trafiać będą do bibliotek średnio w dwa tygodnie po ukazaniu się drukowanej wersji czasopisma.

Literatura

1. BRADBURG D. ADONIS-The View of the Users. IFLA Journal 1988 nr 14 s. 132-136
2. BRAID A. Document delivery - the dawn of a new era. IATUL Quarterly 1989 nr 3 s.207-213
3. FRIEND F.J. ADONIS: a happy medium or falling between two stools? IATUL Quarterly 1990 nr 4 s.239-243
4. STERN B.T. Publishing on CD-ROM in mixed mode. W: Proceedings 10th International Online Meeting 1986 (10) s. 23-31
5. STERN B.T., Compier H.C. ADONIS - Document delivery in the CD-ROM age. Interlending and Document Supply 1990 nr 3 s.79-87.
6. THENZEN J. Von der Bibliothek zur Discothek? ADONIS nimmt den Zweiten Anlauf. ABI-Technik 1987 (7) s.81-89

20.05.1991 r.s

ADONIS - A NEW METHOD OF MAKING SCIENTIFIC LITERATURE ACCESSIBLE

Summary

ADONIS is an experiment consisting in the use of CD-ROM technology for acquisition, transfer and copying contents of scientific journals. Libraries which have taken part in the programme received CD-ROM containing informations from biomedical and biomedical journals. Realization of the order submitted by user consisted in retrieval of the relevant article from the disk and in printing it off by means of the laser printer. The two years exploitation of ADONIS confirmed the possibility of obtaining a considerable financial savings with preserving the high quality of services provided for users.

ADONIS - НОВЫЙ МЕТОД ДОСТУПА К НАУЧНОЙ ЛИТЕРАТУРЕ

Резюме

ADONIS это эксперимент состоящий в использовании техники CD-ROM для накопления, передачи и копирования содержания научных журналов. Участвующие в программе библиотеки получали оптические диски содержащие информации из биомедицинских и биологических журналов. Реализация заявки предоставленной потребителем состоит в поиске соответственной статьи на диске CD-ROM и в ее отпечатке на бумагу с помощью лазерного печатающего устройства. Двухлетняя эксплуатация системы ADONIS подтвердила возможность существенной экономизации финансовых ресурсов при сохранении высокого качества услуг предоставляемых потребителям.

WIESŁAW BABIK

Ośrodek Informacji Naukowej PAN

EUROPEAN ACADEMIC RESEARCH NETWORK (EARN)

Geneza i struktura sieci EARN. Zarządzanie. Dostęp do sieci w Polsce. Oferowane usługi.

Sieć EARN jest w części darem firmy IBM dla środowiska naukowego krajów Europy Zachodniej, w tym Polski. Stowarzyszenie Polski z EARN w maju 1990 roku otwarło przed polskimi naukowcami jakościowo nowy etap współpracy i uczestnictwa w różnych programach badawczych.

Geneza i struktura sieci

EARN jest siecią komputerową utworzoną na przełomie 1983/1984 roku, łączącą komputery zainstalowane w uniwersytetach i ośrodkach badawczych z zakresu wszystkich dyscyplin z 26 krajów Europy Zachodniej, Środkowego Wschodu i Afryki. Sieć ta stanowi połączenie z powstałą 3 lata wcześniej w USA siecią BITNET (skrót od: Because It's Time Network) działającą w Japonii, Meksyku i USA i holenderską siecią NETNORTH w taki sposób, że zespół EARN/BITNET i NETNORTH tworzy jedną sieć łączącą obecnie około 2600 komputerów, tworząc tym samym jedną z najbardziej znaczących sieci komputerowych na świecie. Za pośrednictwem tej sieci za pomocą słuz (gateways) istnieje możliwość dostępu do prawie wszystkich znanych sieci badawczych:

"Zagadnienia Informacji Naukowej" 1992 nr 1 (59)

1. międzynarodowych, np. CSNET (Computer Science Network), EUNET (European UNIX Network); HELPNET (High Energy Physics Network);
2. regionalnych, np. NORDUNET w krajach skandynawskich; UUCP (sieć maszyn UNIX); ARPANET;
3. krajowych, np. OFN w Niemczech, JANET w Wielkiej Brytanii.

EARN nie różni się koncepcyjnie i organizacyjnie od amerykańskiej sieci BITNET i dlatego jest traktowana jako jej wyodrębniona część europejska. Daje to pełne możliwości współpracy obu zintegrowanych sieci. Stąd często używa się nazwy EARN/BITNET.

Koncepcja EARN została oparta na technologii telekomunikacyjnej firmy IBM nazywanej "System Network Architecture" (SNA). EARN działa od początku 1984 roku. W maju 1990 roku w skład sieci EARN/BITNET wchodziło 2913 węzłów. Oprócz krajów Europy Zachodniej i Jugosławii do sieci włączono następujące kraje: Algieria, Cypr, Egipt, Indie, Izrael, Jordania, Maroko, Syria, Tunezja i Wybrzeże Kości Słoniowej.

Na poziomie technicznym omawiana sieć ma charakter sieci zdecentralizowanej. Oznacza to, że odpowiedzialność za spójność i poprawność informacji w sieci ponoszą wszyscy członkowie posiadający węzły sieci.

Zarządzanie siecią

EARN jest siecią komputerową umożliwiającą dostęp do wszystkich uczelni i niekomercyjnych instytucji badawczych na terenie Europy Zachodniej, Afryki i Środkowego Wschodu. Jej zadaniem jest umożliwienie wymiany informacji pomiędzy członkami sieci. Sieć nie może być jednak wykorzystywana do wymiany informacji o charakterze komercyjnym, politycznym, tajnym i religijnym. Transmitowane dane nie podlegają specjalnej ochronie.

EARN działa jako zarejestrowane we Francji stowarzyszenie (EARN Association) rozumiane jako organizacja jego członków. Stowarzyszenie to jest zarządzane przez Radę Dyrektorów (Board of Directors), odbywającą swoje robocze posiedzenia dwukrotnie w roku. Sprawami bieżącymi Stowarzyszenia zajmuje się wybrany przez Radę Komitet Wykonawczy (Executive Committee), na czele którego stoi Prezydent Rady Dyrektorów (President of the Board). Siedzibą Biura Komitetu (EARN Office) jest miejscowość Orsay pod Paryżem. O przyjęciu nowego państwa do Stowarzyszenia decyduje Rada Dyrektorów.

Dostęp do EARN w Polsce

Główny węzeł krajowy sieci EARN w Polsce o nazwie PLEARN został umieszczony w budynku Centrum Informatycznego Uniwersytetu Warszawskiego w Warszawie przy ulicy Krakowskie Przedmieście 26/28. Węzeł ten jest połączony linią dzierżawioną z głównym węzłem duńskim znajdującym się w Lyngby k/Kopenhagi.

Aby umożliwić dostęp do EARN jak najszerszemu gronu polskich naukowców obecnie wykorzystuje się w Polsce sieć KASK (skrót od: Krajowa Akademska Sieć Komputerowa), w oparciu o którą uruchomiono węzły w:

- * Politechnice Wrocławskiej;
- * Politechnice Łódzkiej;
- * IPI PAN w Warszawie;
- * Uniwersytecie Mikołaja Kopernika w Toruniu;
- * CYFRONECIE w Krakowie.^y

^yZob. też bieżący numer "Zagadnień..." s. 202.

Usługi oferowane przez sieć EARN

EARN jest przede wszystkim narzędziem łączności pomiędzy użytkownikami umożliwiającym następujące usługi:

- wysyłanie i odbiór komunikatów i listów (tzw. poczta elektroniczna);
- przesyłanie zbiorów różnych typów, tj. danych, programów i dokumentów;
- wymiana informacji w trybie online, w tym telekonferencje;
- zdalne wprowadzanie i realizowanie zadań (jobs);
- dostęp do zdalnych aplikacji baz danych i bibliotek.

Dwa pierwsze rodzaje usług dają możliwość m.in. abonowania magazynów i periodyków sieciowych, uczestniczenia w grupach dyskusyjnych (tzw. "users discussion groups") oraz uczestniczenia w zespołach problemowych.

Niestety do tej pory nie wszystkie wymienione rodzaje usług i tryby pracy w sieci są dostępne dla polskich użytkowników. Aktualnie w Polsce można korzystać tylko z poczty elektronicznej i dokonywać transferu zbiorów. Praca w trybie online oraz zdalne wprowadzanie i realizowanie zadań są obecnie niedostępne.

Literatura

1. PODRĘCZNIK sieci EARN pod systemem VM/SP. Uprac. w Centrum Informatycznym Uniwersytetu Warszawskiego przez D. Burzyńską na podstawie wydania francuskiego: Dominique Pinse: Manuel d' Utilisation sous systeme VM/SP. Warszawa 1990
2. RYKACZEWSKA-WIOROGÓRSKA B. Sieć komputerowa EARN w Polsce. Biuletyn EARN-Polska 1990 Nr 1 s. 4-7.
3. RYKACZEWSKA-WIOROGÓRSKA B. Sieć komputerowa EARN w Polsce. Aktualne Problemy Informacji i Dokumentacji 1990 nr 4 s.5-8.

4. ZIENKIEWICZ A. Akademicka Sieć Komputerowa. Biuletyn EARN-Polska 1991 Nr 2 s. 5-9.

6.01.1992

EUROPEAN ACADEMIC RESEARCH NETWORK (EARN)

Summary

There are discussed: the origins and structure of the EARN network, the means of management, the access to the network in Poland and the provided services from which only electronic mail and the transfer of files is available for the Polish users.

ЕВРОПЕЙСКАЯ АКАДЕМИЧЕСКАЯ ИССЛЕДОВАТЕЛЬСКАЯ СЕТЬ (EARN)

Р е з ю м е

Охарактеризованы происхождение и структура сети EARN, способы ее управления, доступ к EARN в Польше, а также предоставляемые ею услуги. Из всех видов обслуживания EARN в Польше доступны только электронная почта и передача файлов.

RECENZJE I OMÓWIENIA

PROPOZYCJA NOWEJ KLASYFIKACJI DLA MATERIAŁÓW KARTOGRAFICZNYCH *

Problematyka stosowania odpowiedniego języka informacyjno-wyszukiwawczego (jłw) do charakterystyki treściowej i wyszukiwania informacji (materiałów kartograficznych) w zbiorach kartograficznych jest dyskutowana od lat. Plonem tych dyskusji jest między innymi wielość i różnorodność stosowanych obecnie systemów klasyfikacyjnych. Nie przystają one jednak do stale rosnących i coraz bardziej różnicujących się potrzeb informacyjnych użytkowników zbiorów kartograficznych. Praktyczne bowiem posługiwanie się tymi klasyfikacjami powoduje w wielu wypadkach możliwe do uniknięcia lub co najmniej możliwe do zmniejszenia straty często bardzo ważnych dla użytkownika informacji o tych materiałach, a przez to zmniejsza stopień zaspokojenia jego potrzeb informacyjnych. Systemy te, zbudowane w oparciu o tradycyjne podstawy metodologiczne, opierają się głównie na strukturze samych zbiorów kartograficznych i tylko w niewielkim stopniu uwzględniają ich cechy - dy-

*Wiesław BABIK: Wykorzystanie fasetowej organizacji języka informacyjno-wyszukiwawczego do budowy klasyfikacji materiałów kartograficznych. Praca doktorska wykonana pod kierunkiem doc.dr.hab. E. Ścibora, obroniona na Uniwersytecie Warszawskim, Wydz. Neofilologii, 17 marca 1992 r. Recenzenci: prof.dr hab. O.A. Wojtasiwicz, prof. dr hab. B. Świderski, Warszawa 1992, 354 s. maszyn.

stynktywne, które mogą stanowić elementy relewantne dla użytkowników tego rodzaju zbiorów.

Kompleksowy charakter informacji kartograficznej oraz sposobów jej prezentacji w materiałach kartograficznych wskazuje na potrzebę takiej organizacji pola semantycznego jiw, która umożliwiałaby kompleksowość ujęć materiałów kartograficznych zarówno w aspekcie treści, jak i formy, a przez to pozwalałaby na wieloaspektowe indeksowanie i wyszukiwanie informacji. Możliwość taką wydaje się stwarzać fasetowa organizacja jiw.

Tematem rozprawy jest próba wykorzystania teorii fasetowej organizacji jiw do budowy klasyfikacji fasetowej dla materiałów kartograficznych. Zagadnienie to dotyczy zarówno sposobu odzworowywania fragmentu wiedzy o rzeczywistości wbudowywanej w jiw typu klasyfikacji, jak i samej metodyki budowy klasyfikacji fasetowej.

Źródłem prezentowanej koncepcji są najnowsze tendencje w konstrukcji jiw znajdujące odbicie w literaturze z zakresu ogólnej teorii i metodyki budowy klasyfikacji, a zwłaszcza koncepcji fasetowej organizacji jiw, które chociaż pojawiły się w latach trzydziestych naszego stulecia, to jednak nadal uchodzą za nowatorskie oraz teoria klasyfikacji fasetowej opracowana przez S.R. Ranganathana, a rozwijana obecnie przez Centrum Dokumentacji Badań i Szkolenia (Documentation Research and Training Centre - DRTC, Bangalore - Indie) i Zespół Badań Klasyfikacji (Classification Research Group -CRG, Londyn - Wielka Brytania).

Klasyfikacja materiałów kartograficznych należy w Polsce do zagadnień mało opracowanych i chociaż tuż po II wojnie światowej problematyka ta była bardzo żywo dyskutowana, nie znalazła do tej pory właściwego odzwierciedlenia ani w literaturze przedmiotu, ani w działaniach praktycznych.

Przedmiotem rozprawy jest projektowanie jiw do opisu i wyszukiwania materiałów kartograficznych. Język ten winien umożliwiać minimalizację strat informacji, powstających na etapach indeksowania i wyszukiwania materiałów kartograficznych. Zaproponowany język to klasyfikacja fasetowa, a więc jiw o wielowymiarowej strukturze (fasetowej organizacji) pola semantycznego.

Proponowany schemat klasyfikacji stanowi egzemplifikację sposobu rozwiązania sformułowanego w pracy problemu badawczego. Ponieważ większość materiałów kartograficznych stanowią mapy, projektowana klasyfikacja dotyczy przede wszystkim map; uwzględnia jednak również i inne rodzaje materiałów kartograficznych.

Budowę klasyfikacji oparto na podziale zbioru materiałów kartograficznych. Wymagało to wcześniejszego określenia - i to w sposób bardzo dokładny - zakresu tego zbioru, a także szeregu uściśleń terminologicznych, zarówno w teorii jiw, jak i w stosunku do elementów dzielonego zbioru.

Na zakres i strukturę pola semantycznego budowanej klasyfikacji składają się:

- cechy formalne materiałów kartograficznych;
- suma treści pytań informacyjnych;
- struktura informacyjna obszaru rzeczywistości, który ma odwzorowywać klasyfikacja;
- istniejące i potencjalne potrzeby informacyjne użytkowników materiałów kartograficznych.

Przy budowie schematu klasyfikacji fasetowej dla materiałów kartograficznych posłużono się metodyką opisaną w literaturze przedmiotu. Do gromadzenia leksyki wykorzystano dotychczas szeroko stosowaną metodę dedukcyjną, polegającą na wykorzystywaniu już istniejących źródeł słownictwa oraz metodę indukcyjną, polegającą na indeksowaniu reprezentatywnego zbioru dokumentów. Jak wykazała praktyka, stosowanie tylko jednej z wymienionych metod jest niewystarczające. Dlatego wykorzystywano obie metody jednocześnie. Tego typu procedura gromadzenia leksyki, nazywana w teorii jiw metodą mieszaną, jest powszechnie stosowana przy budowie systemów leksykalnych jiw.

Przy konstrukcji systemu leksykalnego budowanego jiw - zgodnie z dotychczasową praktyką w zakresie tworzenia klasyfikacji fasetowych - oparto się na materiale empirycznym, przyjmując za źródła leksyki słownictwo języka naturalnego użyte w opisach materiałów kartograficznych zamieszczonych w bibliografiach materiałów opublikowanych w Polsce w ciągu ostatnich 10 lat, tytułach samych materiałów kartograficznych, oraz w ich legendach. Wykorzystano również słownictwo wybranych katalogów bibliotecznych oraz drukowanych katalogów polskich zbiorów kartograficznych, a także tablice istniejących klasyfikacji materiałów kartograficznych oraz terminologię kartograficzną stosowaną w piśmiennictwie z zakresu kartografii i nauk o Ziemi. Za podstawowe źródła leksyki przyjęto "Bibliografię Geografii Polskiej" i bibliografię publika-

cji Państwowego Przedsiębiorstwa Wydawnictw Kartograficznych (PPWK) zamieszczoną w tomie 8 "Polskiego Przeglądu Kartograficznego".

Leksykę gromadzono stosując metodę swobodnego indeksowania opisów bibliograficznych. Słownictwo opisów bibliograficznych uzupełniono słowami kluczowymi z zapytań użytkowników skierowanych do Centralnej Informacji Kartograficznej w Warszawie oraz powstałych tam zestawień tematycznych.

Zbrane w ten sposób terminy wprowadzono do specjalnie zaprojektowanej terminologicznej bazy danych. Ze wszystkich opisów bibliograficznych materiałów źródłowych otrzymano zbiór około 2000 elementarnych jednostek leksykalnych. Zbiór ten uznano za reprezentatywny w momencie, gdy dalsze gromadzenie słownictwa powodowało niewielki przyrost liczby nowych słów, co i tak nie spowoduje istotnych zmian w strukturze semantycznej projektowanej klasyfikacji.

Konstrukcję planu znaczeniowego (planu treści) oraz konstrukcję planu wyrażania (notacji) projektowanego języka przeprowadzono dokonując opisu jego systemu leksykalnego oraz realizacji tekstowej projektowanej klasyfikacji, przy czym założono wtórność systemu notacyjnego względem organizacji leksyki jiw. W systemie leksykalnym wyróżniono części składowe nazwane kategoriami semantycznymi. Składają się na nie wyrażenia powiązane wspólnym elementem semantycznym, między którymi istnieją określone relacje semantyczne, m.in. porządkujące hierarchicznie ten zbiór i czyniące go spójnym.

Przy projektowaniu systemu klasyfikacyjnego bardzo ważnym zadaniem było nie tylko ustalenie jego zakresu tematycznego, lecz przede wszystkim budowa struktury systemu leksykalnego.

Zakres tematyczny miał zasadniczy wpływ na wybór podstawowego słownictwa systemu, struktura zaś na rozmieszczenie poszczególnych terminów w tablicach i określanie relacji między tymi terminami.

Organizację systemu leksykalnego przeprowadzono stosując metodę kategoryzacji semantycznej (fasetyzację) zgromadzonej leksyki, która stanowi podstawowy sposób zapewnienia wielowymiarowości pola semantycznego jiw. Kategoryzacja elementarnych jednostek leksykalnych jiw jest drugim - obok wyróżniania relacji paradygmatycznych między jednostkami leksykalnymi jiw - środkiem organizacji jego systemu leksykalnego. Jest to podział jednostek leksykalnych na stosunkowo nieliczne klasy, których wszystkie elementy pełnią identyczną funkcję w jiw.

Podział jednostek leksykalnych jiw na pewne kategorie nie tylko ułatwia organizację zbioru informacyjnego, lecz przede wszystkim stanowi efektywny środek poprawnej interpretacji znaczenia wyrażen jiw, co ma duży wpływ na optymalizację procesu wyszukiwania informacji.

Projektując strukturę systemu leksykalnego klasyfikacji przeprowadzono następujące czynności:

- dokonano wyboru ogólnych kategorii semantycznych wyznaczających strukturę systemu klasyfikacji;
- dokonano wyboru pól znaczeniowych pokrywających ustalony zakres systemu leksykalnego;
- dokonano wyboru relacji systematyzujących leksykę w obrębie poszczególnych pól semantycznych.

Spośród wielu relacji, istniejących pomiędzy jednostkami leksykalnymi, do organizacji systemu leksykalnego wykorzystano jedynie

relacje hierarchiczne: generyczne, mereologiczne i hierarchii tematycznej.

Organizacja systemu leksykalnego według wymienionych relacji paradygmatycznych doprowadziła do utworzenia drzew klasyfikacyjnych odwzorowujących układ jednostek leksykalnych w poszczególnych polach semantycznych.

Leksyka wyróżnionych kategorii, podkategorii i faset dotyczy:

1. rodzaju materiałów kartograficznych;
2. treści materiałów kartograficznych;
3. przeznaczenia tych materiałów;
4. skali i odwzorowania;
5. terytorium;
6. daty sytuacji oraz daty wydania.

Przeprowadzona kategoryzacja słownictwa stosowanego do opisu materiałów kartograficznych pozwoliła wyodrębnić kategorie semantyczne, konstytuujące równocześnie porządek poziomy, czyli rozmieszczenie jednostek leksykalnych w tekstach jiw, oraz porządek pionowy określający porządek jednostek leksykalnych w tablicach klasyfikacyjnych.

Do uporządkowania leksyki przyjąłem 6 następujących kategorii semantycznych, zaliczonych do przyjętych w klasyfikacji dwóch wymiarów pola semantycznego jiw:

- wymiar cech formalnych (FORMA):
 - ELEMENTY FORMALNO-WYDAWNICZE (E),^y
 - PREZENTACJA (P),

^yW nawiasach umieszczono przyjęte w klasyfikacji kody literowe służące do oznaczania kategorii.

- wymiar cech treściowych (TREŚĆ):

- TEMAT (S),
- PRZESTRZEŃ (A),
- CZAS (T)

oraz jako kategorię wyodrębnioną - kategorię WYKORZYSTANIE (U).

Klasyfikacja w wersji zaprezentowanej w pracy ma charakter otwarty. Zawiera 1079 elementarnych jednostek leksykalnych wraz z nazwami własnymi oraz 6 nazw kategorii, 22 nazwy podkategorii i 54 nazwy faset. Jako nazwy kategorii semantycznych przyjęto nazwy pewnych szerokich znaczeń wyróżniających w planie treści jiw obszerne pola semantyczne. Wyznaczają je zwykle takie elementy znaczenia, które są obecne w definicjach znaczenia wielu zgromadzonych terminów. Są to więc powtarzające się elementy konotacji wyrażeń jiw. Za nazwy poszczególnych faset przyjęto nazwy kryteriów podziału leksyki na najniższym poziomie.

Szeregi klasyfikacyjne utworzone w poszczególnych fasetach mają bardzo zróżnicowaną liczebność. Występują w nich zarówno szeregi dwuelementowe, jak i rozwinięte szeregi kilkunasto- a nawet kilkudziesięcioelementowe, na ogół podzielone na sektory obejmujące kilka lub kilkanaście jednostek leksykalnych zgrupowanych na podstawie różnych zasad podziału.

Konstrukcja rozprawy odzwierciedla tok przeprowadzonego rozumowania. Praca składa się z trzech części.

W pierwszej części dokonano określenia zakresu dzielonego zbioru materiałów kartograficznych, specyfikacji jego elementów, a także wyróżniono cechy dystynktywne tych elementów. Wymagało to szczegółowej analizy i charakterystyki materiałów kartograficznych w celu uchwylenia ich elementów relewantnych, istotnych z punktu

widzenia odwzorowywania rzeczywistości oraz potrzeb użytkowników. Na podstawie przeprowadzonej charakterystyki stwierdzono niezbędną wielowymiarową strukturę pola semantycznego jiw dla materiałów kartograficznych, bowiem taka struktura może umożliwić pełniejsze odwzorowywanie wyróżnionych elementów relewantnych w tekstach jiw, to jest w charakterystykach wyszukiwawczych dokumentów oraz instrukcjach wyszukiwawczych zapytań użytkowników. Opis materiałów kartograficznych dokonany w części I rozprawy pozwolił wyróżnić cechy dystynktywne materiałów kartograficznych, skonfrontować je z potrzebami użytkowników i uznać je za cechy relewantne w procesie wyszukiwania materiałów kartograficznych. Wyróżnione elementy stanowiły podstawę uzasadnienia przyjęcia wielowymiarowej struktury pola semantycznego projektowanej klasyfikacji.

W drugiej części rozprawy skonfrontowano możliwości obecnie najszerszej stosowanych systemów klasyfikacji w zakresie stopnia i sposobów odwzorowywania wyróżnionych elementów relewantnych. Na podstawie literatury dokonano specyfikacji oraz analizy sposobów odwzorowywania elementów relewantnych w wybranych klasyfikacjach ogólnych i klasyfikacjach przeznaczonych wyłącznie dla materiałów kartograficznych, stosowanych w bibliotekach polskich i zagranicznych. Przedstawiono również projekty nowych rozwiązań w dziedzinie klasyfikacji materiałów kartograficznych. Przy charakterystyce systemów stosowanych w polskich bibliotekach oparto się na wynikach rekonesansu, przeprowadzonego na potrzeby rozprawy w wytypowanych bibliotekach naukowych. Okazało się, że w stosowanych klasyfikacjach ogólnych i specjalistycznych wiele elementów relewantnych wyróżnionych w części I rozprawy w ogóle nie jest

odwzorowywana. Elementy te nie zostały nawet uwzględnione w najnowszych projektach klasyfikacji materiałów kartograficznych. Straty informacji powstające przy wyszukiwaniu informacji na podstawie istniejących klasyfikacji mogłyby zostać znacznie zmniejszone w klasyfikacji o wielowymiarowym polu semantycznym.

Część trzecia rozprawy została poświęcona omówieniu własnej propozycji klasyfikacji fasetowej dla materiałów kartograficznych. Na podstawie literatury przedmiotu przedstawiono najpierw elementy teorii fasetowej organizacji jiw, a następnie opisano przyjętą metodykę budowy klasyfikacji oraz zaprezentowano zrzęb (schemat) zbudowanej klasyfikacji. Zaprojektowaną klasyfikację zaopatrzone w projekty systemu notacyjnego (mieszana notacja hierarchiczna), gramatyki oraz instrukcji indeksowania dokumentów i zapytań użytkowników.

Zbudowana klasyfikacja jest projektem specjalistycznego jiw typu klasyfikacji fasetowej, służącego do indeksowania i wyszukiwania materiałów kartograficznych w systemie informacji dokumentacyjnej z możliwością wykorzystywania jej w komputerowych bazach danych o materiałach kartograficznych.

Siła semantyczna zaproponowanej klasyfikacji jest większa niż obecnie stosowanych schematów klasyfikacji materiałów kartograficznych. Tradycyjne klasyfikacje, z reguły o budowie hierarchicznej, explicite podają cechy dystynktywne materiałów kartograficznych. Są to zatem systemy wyliczające. Włączają one każdy dokument tylko do jednej klasy semantycznej. Zbudowana klasyfikacja - chociaż praktyczne posługiwanie się nią wymaga od użytkownika zarówno dobrej znajomości jej struktury, jak i dużej wiedzy o materiałach kartograficznych - umożliwia bardziej selektywne

wyszukiwanie materiałów kartograficznych. Pozwala ona bowiem na włączanie ich do tyłu klas, ile cech dystynktywnych mają przedmioty lub sytuacje opisywane i odzwierciedlane w ich charakterystykach wyszukiwawczych. Umożliwia więc wieloaspektowe indeksowanie i wyszukiwanie materiałów kartograficznych.

Swojego projektu klasyfikacji nie traktuję jako schematu zamkniętego. Zarówno liczba wymiarów, jak i liczba wyróżnionych kategorii, podkategorii i faset może zostać zmieniona. Schemat ten może zostać wykorzystany lub adaptowany do potrzeb konkretnych zbiorów kartograficznych.

Wiesław Babik

UWAGI DO "INSTRUKCJI TEMATOWANIA I KATALOGU PRZEDMIOTOWEGO"*

Ukazała się "Instrukcja tematowania^{1/} i katalogu przedmiotowego" opracowana przez Jadwigę Sadowską z Biblioteki Narodowej.

* Tekst niniejszy jest rozszerzoną wersją recenzji "Instrukcji tematowania i katalogu przedmiotowego". Oprac. Jadwiga Sadowska. Warszawa: Stowarzyszenie Bibliotekarzy Polskich; Bibl. Narod. 1991, 100 s. Recenzja ta została złożona w redakcji czasopisma "Przegląd Biblioteczny".

^{1/} Terminy tematowania i przedmiotowania są traktowane w specjalistycznym języku bibliotekarzy jako synonimy oznaczające indeksowanie w języku haseł przedmiotowych. W tekście częściej używam tylko jednego z nich - przedmiotowanie - jako tego, którego znaczenie strukturalne wskazuje na istotę opracowania przedmiotowego - odzworowywanie informacji o przedmiocie dokumentu.

Jak pisze autorka we Wstępie, "Instrukcja (...)" wywodzi się i ma ambicje kontynuować tradycję zapoczątkowaną przez podręcznik "Katalog przedmiotowy"^{2/} Adama Łysakowskiego; z założenia ma być jego weryfikacją i uzupełnieniem. Środowisko bibliotekarskie na pewno przyjęło tę publikację z dużym zainteresowaniem i wdzięcznością jako jedną z nielicznych pomocy metodycznych dla osób zajmujących się opracowaniem przedmiotowym. Ale choć w tytule pracy zostało użyte sformułowanie "Instrukcja", to mamy tu do czynienia z publikacją o bardziej złożonej naturze. Obok przepisów ustalających reguły postępowania, wskazówek i rad, "Instrukcja (...)" dokonuje rewizji pewnych zasad teoretycznych języka haseł przedmiotowych. Wprawdzie zmiany te nie były celem, dla którego praca powstała, niemniej warto im się nieco bliżej przyjrzeć i przy tej okazji zastanowić się choćby nad niektórymi problemami systemowymi JHP.

Ważniejsze wyrażone w "Instrukcji (...)" zmiany w systemie języka haseł przedmiotowych dotyczą przede wszystkim kategoryzacji tematów i określników. Zaczniemy od tematów. Wyodrębniono dwa ich podziały: według kryterium formy, a właściwie rodzaju nazwy, za pomocą której wyrażane jest znaczenie (tematy ogólne i jednostkowe) oraz według kryterium funkcji i zakresu stosowania (tematy ogólne i tematy jednostkowe: osobowe, etniczne, chronologiczne, geograficzne i toponimiczne, korporatywne, rzeczowe jednostkowe). Te dwa podziały zdają się tworzyć jakby dwa wymiary jednej typologii, choć informacja taka explicite nie jest zawarta w "Instrukcji (...)". Inaczej mówiąc, każdy temat wyodrębniony według kryte-

^{2/} Adam Łysakowski: Katalog przedmiotowy. Podręcznik. Warszawa: PZWS 1946.

rium funkcji i zakresu stosowania jest wyrażany przez jakąś nazwę - ogólną lub jednostkową. Tematy ogólne są rzeczywiście bardzo trudne do scharakteryzowania z punktu widzenia ich zakresu semantycznego. Autorka zdecydowała się jednak wymienić pewne kategorie tematów ogólnych, takie jak na przykład nazwy: chorób, gatunków świata roślinnego i zwierzęcego oraz ich przedstawiciele, kamieni szlachetnych, przedstawicielei wyznań religijnych, kierunków i prądów filozoficznych, literackich, kulturalnych, społecznych itd. Dlaczego uznano, iż warto wskazać właśnie wymienione wyżej kategorie jako te. "które powinny być odzwierciedlone w postaci odpowiednich do ich nazw tematów ogólnych"? Odpowiedź na to pytanie zapewne wynika z jakiś względów praktycznych, gdyż teoretycznego uzasadnienia nie daje się znaleźć.

W części poświęconej tematów ogólnym samodzielnym podrozdziałem zostały wyróżnione tylko tematy językowo-etniczne^{3/} charakteryzujące się tym, że "w nazwie ich zawarta jest informacja o przynależności językowej, etnicznej lub państwowej". Wyrażenia określające cechę językową, etniczną lub państwową możemy dodawać: a) do tematów z dziedziny literatury, np.: LITERATURA AMERYKAŃSKA (...); b) do tematów z dziedziny językoznawstwa, gdzie oznacza cechę językową, np.: SŁOWNIK^{4/} ROSYJSKO-POLSKI (...-); c) do tematów z dziedziny sztuki, gdzie oznacza cechę etniczną lub państwową, np.: RZEŻBA FRANCUSKA (...); d) do tematów z dziedziny kultury, gdzie oznacza cechę etniczną lub państwową, np.: PIEŚŃ UKRAIŃSKA .

^{3/} Termin ten w znaczeniu używanym w "Instrukcji (...)" wprowadziła do literatury Jadwiga Ćwiekowa w publikacji Opracowanie przedmiotowe piśmiennictwa. (Problemy teoretyczne i praktyczne). Warszawa: OIN PAN, 1974.

^{4/} Czy temat SŁOWNIK w tym znaczeniu jest naprawdę tematem z dziedziny językoznawstwa?

Zauważyliśmy, iż cecha przynależności językowej, etnicznej czy państwowej zdominowała inne cechy wyrażane przez tematy, w kontekście których się pojawia. Znalazło to wyraz w samej nazwie tej grupy tematów. Nie jest jakby ważne, że chodzi o LITERATURE AMERYKAŃSKĄ, RZEŻBĘ FRANCUSKĄ, PIEŚŃ UKRAIŃSKĄ, ale że jest to AMERYKAŃSKIE, FRANCUSKIE, UKRAIŃSKIE. Nie jest błędem wyodrębnianie takich tematów, jeśli są one potrzebne w katalogu przedmiotowym, nazywanie ich jednak tematami językowo-etnicznymi wydaje się jakby nieuzasadnione, gdyż nie nazywają one obiektów językowo-etnicznych a jedynie językowo-etniczne nacechowanie (punkt widzenia, aspekt) jakiś innych przedmiotów.

Także określniki zostały podzielone przez autorkę "Instrukcji (...)" według dwóch różnych zasad: formy (określniki ogólne i jednostkowe) oraz funkcji i zakresu stosowania (określniki treściowe - klasowe, przedmiotowe, lokalizujące - i określniki formalne). Podobnie jak w przypadku tematów można powiedzieć, iż każdy określnik wyodrębniony według kryterium funkcji i zakresu może być wyrażony za pomocą albo nazwy jednostkowej, albo nazwy ogólnej. Ponownie nasuwa się też pytanie, jakimi względami praktycznymi była umotywowana taka segmentacja klasy określników tj. wyodrębnienie wśród określników treściowych podklas określników klasowych i przedmiotowych. Z teoretycznego punktu widzenia uzasadniona wydaje się jedynie opozycja na poziomie określnik treściowy $\frac{\cdot}{\cdot}$ określnik formalny, gdzie określnik treściowy (zwany także rzeczowym) jest wyrażeniem odwzorowującym cechy przedmiotowe^{5/} informacji zawartych w dokumentach, podczas gdy

^{5/} Do cech tych należą także dające się w miarę precyzyjnie określić informacje o przestrzennej i czasowej lokalizacji treści dokumentu konstytuujące grupy określników geograficznych i chronologicznych, zwanych łącznie lokalizującymi.

określniki formalne mówią o naturze i postaci dokumentów.

Z określenia znaczeń terminów określnik rzeczowy i określnik formalny wynika między innymi, iż te dwie kategorie semantyczne określników w pewnych ich podzbiorach mogą się pokrywać. Inaczej mówiąc ta sama forma językowa, ten sam ciąg znaków języka haseł przedmiotowych może pełnić funkcję albo określnika rzeczowego, albo określnika formalnego. Na przykład w haśle MATEMATYKA-PODRĘCZNIK określnik podręcznik jest określnikiem formy, ale w haśle MATEMATYKA-PODRĘCZNIK-METODY staje się określnikiem rzeczowym.

Analiza zagadnienia określników treściowych daleko wykracza poza zakres tego tekstu, którego jednym z celów jest jedynie zasygnalizowanie pewnych problemów teoretycznych języka haseł przedmiotowych zainspirowanych między innymi lekturą "Instrukcji (...)". Dlatego chciałabym skoncentrować uwagę na mniej skomplikowanych problemach określników formalnych i na ich przykładzie pokazać, że na system JHP może (a może trzeba) popatrzeć z innych, odmiennych od tradycyjnych punktów widzenia.

Każdy dokument, czyli każda utrwalona informacja, ma jakąś formę, ale nie zawsze jest ona brana pod uwagę podczas przedmiotowania. Wweksponowanie formy ma miejsce wówczas, gdy przedstawia ona istotną wartość dla selekcji informacji w systemie lub gdy treść dokumentu jest na tle ogólna lub nieprecyzyjna, że nie sposób nazwać przedmioty dokumentu, ale przyjęta przez autora forma przedstawienia tej treści może w miarę precyzyjnie poprzez skojarzenie je wskazywać. (Drugi z wymienionych przypadków wyrażony jest przez tzw. temat formalne).

Dla polskiej tradycji języka haseł przedmiotowych, zwłaszcza jego realizacji w postaci katalogu przedmiotowego, charakterystycz-

ne było używanie terminu określnik formalny w znaczeniu określnik "odzwierciedlający formę dokumentu: piśmienniczą (...) i rzadziej formę wydawniczą (...)"^{6/}. Taka interpretacja znaczenia, i co za tym idzie zakresu stosowania określników formalnych, nie wydaje się spełniać warunku poprawności, jeśli porówna się ją z listami określników, co do których bibliotekarze zgadzają się, iż są to określniki formalne. Na przykład, jeśli przyjąć że określniki formalne odwzorowują formę piśmienniczą dokumentów, to zakres ich stosowania z definicji winien być ograniczony do przedmiotowania dokumentów piśmienniczych, czyli takich, w których informacja jest utrwalona w subkodzie jakiegoś języka naturalnego. Do dokumentów piśmienniczych nie można zaliczyć na przykład albumów; dla których naturalna jest przewaga reprodukcji obrazów (czyli sygnałów utrwalonych za pomocą znaków graficznych) nad tekstem w języku naturalnym. Tymczasem praktyka wykazuje, że określnik formalny album nie tylko jest używany, ale wręcz niezbędny do przedmiotowania pewnych typów dokumentów. Takich przykładów można wskazać więcej. Oczywista jest więc potrzeba zweryfikowania dotychczasowych poglądów na naturę i zakres stosowania określników formalnych.

Na podstawie analizy zakresu stosowania określników formalnych w różnych systemach informacyjno-wyszukiwawczych zidentyfikowałam pięć głównych grup cech dokumentów odwzorowywanych przez te określniki. Proponowane zestawienie nie wyznacza jednak ani klasyfikacji, ani nawet typologii tych określników; jest jedynie wyliczeniem ich potencjalnych atrybutów takim, że jednemu określni-

^{6/}Zob. Instrukcja tematowania i katalogu przedmiotowego. Oprac. J. Sadowska... s. 66.

kowi mogą w tym samym czasie przysługiwać cechy z różnych klas. Inaczej mówiąc, znaczenia określników formy mogą się składać z elementów semantycznych przynależnych do więcej niż jednej z wymienionych niżej grup.

Do klasy **określników formalnych**^{7/} zaliczane są terminy odwzorowujące:

(a) formę i/lub sposób ujęcia, wyrażenia oraz prezentacji informacji zawartych w dokumentach;

Podklasę w tak wyodrębnionej klasie cech tworzyłyby tzw. cechy piśmiennicze odwzorowujące postać, w jakiej jest wyrażona treść dokumentu piśmienniczego. Typowymi określnikami byłyby tu terminy nazywające rodzaje piśmiennicze, na przykład: **encyklopedia, podręcznik, norma, patent, monografia, recenzja, słownik, bibliografia** oraz terminy nazywające rodzaje i gatunki literackie, na przykład: **dramat, powieść, opowiadanie, reportaż.**

(b) formę wydawniczą dokumentu, tj. formę jego opublikowania, określoną z punktu widzenia samoistności wydawniczej, na przykład: **artykuł, odbitka, nadbitka** oraz części wydawniczych dokumentu, na przykład: **czasopismo, seria;**

(c) rodzaj utrwalonego sygnału;

Kryterium typu sygnału dokonuje w uniwersum dokumentów następującego uporządkowania: sygnał dźwiękowy - dokumenty audialne (słuchowe), takie jak: **taśmy magnetofonowe, płyty gramofonowe, płyty kompaktowe** itd.; sygnał dotykowy - dokumenty dotykowe, na przykład **druki brajlowskie**; sygnał optyczny - dokumenty wizual-

^{7/} Ponieważ propozycje te odbiegają od rozwiązań tradycyjnych, dla uniknięcia możliwych nieporozumień wynikających z nazywania różnych obiektów tym samym terminem, może lepiej byłoby określniki tej nowej formuły nazwać określnikami formy a nie formalnymi.

ne (oglądowe), na przykład: **fotografie, przezrocza, filmy, eksponaty muzealne**; wśród dokumentów wizualnych dobrze jest wyodrębnić podklasę dokumentów graficznych, czyli takich, w których informacja została utrwalona za pomocą znaków graficznych, na przykład: **rysunki, grafika, mapy, wykresy**; jeśli znaki graficzne są subkodem języka naturalnego, to taki dokument nazywa się dokumentem piśmienniczym. Z punktu widzenia rodzaju sygnału nie jest istotne, czy dany dokument jest drukiem, dokumentem powielonym czy rękopisem;

(d) postać fizyczną (materialną) dokumentu, na przykład: **zwój, kodeks**;

(e) przeznaczenie czytelnicze;

Określenie przeznaczenia czytelniczego na ogół nie występuje samodzielnie, lecz w postaci frazy tworzącej część bardziej złożonego semantycznie określnika. Na przykład: **filmy dla dzieci, podręcznik akademicki, literatura młodzieżowa**.

Zgodnie z gramatyką języka haseł przedmiotowych określniki formy zajmują ostatnią pozycję w haśle przedmiotowym. Ponieważ pewne określniki mogą pełnić bez zmiany postaci zewnętrznej zarówno funkcję określników rzeczowych, jak i określników formy, należałoby się zastanowić, czy w wypadkach, gdy dany określnik występuje w funkcji określnika rzeczowego może "zamknąć" hasło przedmiotowe, wtedy gdy mogą powstawać wątpliwości co do interpretacji funkcji tego określnika i, co za tym idzie, znaczenia hasła. Inaczej mówiąc, być może przewidywałby się pewne zabezpieczenia gramatyczne nakazujące w takich przypadkach dodawanie po owym określniku innego, umownego określnika ujednoznaczniającego zdanie. Na przykład, w systemie RAMEAU funkcję taką w odniesieniu do pewnej grupy określników pełni wyrażenie "historia i krytyka" (np. Hugo,

Victor, 1802-1885 - adaptacje muzyczne; Hugo, Victor, 1802-1885 - adaptacje muzyczne - historia i krytyka).

Kwestia formy dokumentu wydaje się mieć dużą wagę praktyczną, skoro w "Instrukcji (...)", obok fragmentu poświęconego określnikom formalnym, znalazły miejsce rozdziały "Forma dokumentu jako przedmiot" i "Tematowanie różnych form dokumentów". Autorka scharakteryzowała sposoby indeksowania takich dokumentów jak: albumy, antologie, atlasy, czasopisma, encykliki i bulle papieskie, encyklopedia (ogólne i dziedzinowe), książki pamiątkowe, listy, materiały konferencyjne, modlitewniki, pamiętniki, podręczniki, przemówienia, przewodniki turystyczne, recenzje, reportaże, rozmówki, słowniki, śpiewniki i wypisy^{8/}. Wydaje się jednak, że nie dość wyraźnie zostały odróżnione problemy indeksowania formy jako przedmiotu treści i formy jako jednej z wcześniej scharakteryzowanych cech pozat treściowych. Na przykład, w fragmencie dotyczącym czasopism jest mowa o nich zarówno jako o przedmiotach dokumentu, co znajduje wykładnię w postaci tematu (tzw. formalnego), jak i o pewnej formie wydawniczej, co znajduje wykładnię w postaci tematu (formalnego - w przypadku czasopism ogólnych) lub określnika formalnego (w przypadku wydawnictwa o określonym profilu tematycznym).

Porównując "Instrukcję (. .)" ze wspomnianym już "Katalogiem przedmiotowym" i pracami Jadwigi Ćwiekowej^{10/}, odnoszę wrażenie,

^{8/} Dla porównania, w podrozdziale dotyczącym określników formalnych zostały bliżej scharakteryzowane tylko następujące nazwy form: album, ikonografia, indeks, spis, rejestr, skrowidz, encyklopedia, słownik, publicystyka, szkice, recenzja.

^{9/} Jadwiga Ćwiekowa: Opracowanie przedmiotowe piśmiennictwa. (Problemy teoretyczne i praktyczne). Warszawa: OIN PAN, 1974; Opracowanie tematyczne piśmiennictwa. Język haseł przedmiotowych. Katalog przedmiotowy. Warszawa; PWN, 1988.

iż autorka opracowywała tę publikację jakby pod naciskiem dwóch przeciwstawnych sił - chęci oderwania się od niepoprawnych czy zdezaktualizowanych elementów tradycji katalogu przedmiotowego i nacisku by jednak od tej tradycji zbyt daleko się nie oddalić. Rozdarcie to jest zrozumiałe, jeśli weźmie się pod uwagę okoliczności, w jakich powstawała praca i cele, jakim miała służyć. Może właśnie z powodu tych dwoistości jest ona dobrym punktem wyjścia do podjęcia, obok prac metodycznych, skierowanych ku praktyce, także prób rewizji niektórych podstaw teoretycznych języka haseł przedmiotowych, zwłaszcza że wbrew wcześniejszym spekulacjom język ten (choć już w znacznie zmienionej postaci) nie należy do przeszłości opracowania rzeczowego. Wnikliwej i krytycznej analizy wymagają wreszcie zagadnienia przedmiotowania, takie jak: analiza informacyjno-logiczna dokumentów dla systemu przedmiotowego, sposoby wyrażania przedmiotów, aspektów, własności i części przedmiotu, poziomy wyszczególniania informacji w opisie.

Praca Jadwigi Sadowskiej w istotnym stopniu wypełnia lukę w piśmiennictwie polskim z zakresu metodyki języka haseł przedmiotowych i opracowania przedmiotowego. Dlatego bez względu na różne uwagi, które będzie miał każdy jej czytelnik, warta jest upowszechnienia i przede wszystkim tego, żeby się z nią wnikliwie i rzetelnie zapoznać.

Jadwiga Woźniak

9.03.1992 r.

NARZĘDZIA ORGANIZACJI WIEDZY A INTERFEJS UŻYTKOWNIKA
- MATERIAŁY PIERWSZEJ MIĘDZYNARODOWEJ KONFERENCJI
MIĘDZYNARODOWEGO TOWARZYSTWA ORGANIZACJI WIEDZY

W dniu 22 lipca 1989 r. grupa członków Niemieckiego Towarzystwa Klasyfikacji (Deutsche Gesellschaft für Klassifikation) - wywodzących się z RFN i spoza tego państwa - założyła Międzynarodowe Towarzystwo Organizacji Wiedzy (International Society for Knowledge Organization - ISKO). Sceneria tego wydarzenia była dość niecodzienna - nowe międzynarodowe towarzystwo naukowe zostało założone bowiem w restauracji Intercity na centralnym dworcu kolejowym we Frankfurcie nad Menem. Do kilkusobowego grona inicjatorów należała m.in. dr Ingetraut Dahlberg, poprzednio wieloletnia przewodnicząca Komitetu Badań Klasyfikacyjnych Międzynarodowej Federacji Dokumentacji (FID Committee on Classification Research - FID/CR).

ISKO pragnie skupiać specjalistów ze wszystkich krajów świata w zakresie konceptualnej organizacji wiedzy (ang. conceptual organization of knowledge), dokonywanej za pomocą klasyfikacji, indeksowania, systematyzowania terminologii, analizy pojęć itp. Celem towarzystwa - zgodnie z jego statutem - jest "popieranie badań, rozwoju i zastosowań wszystkich metod służących do organizacji wiedzy - ogólnie lub w zakresie pewnego specjalnego obszaru [tematycznego] - przez integrowanie zwłaszcza konceptualnych podejść [właściwych dla] badań klasyfikacyjnych i sztucznej inteligencji". Towarzystwo kładzie nacisk na "filozoficzno-logiczne, psychologiczne i semantyczne podejście w zakresie uporządkowania konceptualnego". Organem służącym do komunikowania informacji

i jej wymiany pomiędzy członkami ISKO jest czasopismo "International Classification", które od 1990 r. ukazuje się jako kwartalnik. ISKO chce współpracować z organizacjami o pokrewnej sferze zainteresowań, np. ze wspomnianym już komitetem FID/CR, Sekcją Klasyfikacji i Indeksowania IFLA oraz INFOTERM-em.

Powyższe informacje zaczerpnięto z notatki poświęconej założeniu ISKO^{1/}. W notatce tej można znaleźć także wyliczenie korzyści, które istnienie towarzystwa przyniesie specjalistom z zakresu organizacji wiedzy. Jedną z tych korzyści jest możliwość częstszego zwoływania międzynarodowych konferencji naukowych w wymienionej dziedzinie. Zaproponowano organizowanie takich konferencji co dwa lata, za każdym razem w innym kraju. Działania w tym kierunku podjęto natychmiast - już we wspomnianej notatce o założeniu nowego towarzystwa zapowiedziano zorganizowanie pierwszej międzynarodowej konferencji pod hasłem "Organizacja wiedzy a interfejs użytkownika". Termin konferencji ustalono na 14-17 sierpnia 1990 r. na krótko przed konferencją IFLA w Sztokholmie.

Konferencja odbyła się w wyznaczonym terminie w Darmstadcie pod nieznacznie zmodyfikowanym tytułem "Narzędzia organizacji wiedzy a interfejs użytkownika". Konferencję zorganizował Sekretariat Generalny ISKO we współpracy z Niemiecką Sekcją ISKO, Komitetem Badań nad Klasyfikacją i Tezaurusami Niemieckiego Towarzystwa Dokumentacji i Komitetem FID/CR. W dniu 14 sierpnia 1990 r. - bezpośrednio przed otwarciem konferencji - odbyło się w Darmstadcie seminarium przedkonferencyjne na temat oprogramowania komputerowego służącego do tworzenia, utrzymania, aktualizacji i integracji tezaurusów.

^{1/}The Founding of the International Society for Knowledge Organization. Frankfurt, 22 July 1989. "International Classification" 1990 vol. 16 nr 2 s. 71-72.

Materiały pierwszej międzynarodowej konferencji ISKO ukazały się jako publikacja dwutomowa, nakładem wydawnictwa INDEKS VERLAG^{2/}.

W każdym z obu tomów umieszczono najpierw referaty wygłoszone na sesji plenarnej, a następnie referaty przedstawione w ramach dwunastu następujących zespołów problemowych (workshops):

- A. Zagadnienia ogólne
- B. Algorytmiczna analiza tekstu
- C. Terminologia
- D. Organizacja wiedzy za pomocą systemów uniwersalnych
- E. Organizacja wiedzy w systemach specjalistycznych [dotyczących określonych dziedzin]
- F. Wyszukiwanie online
- G. Organizacja wiedzy w systemach specjalnych [przeznaczonych dla specyficznych typów placówek lub rodzajów informacji]
- J. Hiperśrodki (ang. Hypermedia)
- K. Wyszukiwanie z zastosowaniem systemów uniwersalnych
- L. Techniki wyszukiwawcze
- M. Indeksowanie

Zespół J: zajmował się głównie zastosowaniami hipertekstu. Przez "systemy uniwersalne" były rozumiane głównie systemy klasyfikacji o zakresie uniwersalnym, a mianowicie radziecka Klasyfikacja Biblioteczno-Bibliograficzna (KBB), Uniwersalna Klasyfikacja Dziesiętna (UKD), Klasyfikacja Biblioteki Kongresu (K BK) i Klasyfikacja Dwukropkowa S.R. Ranganathana.

Tom 1 materiałów konferencyjnych zawiera abstrakty referatów lub ich pełne teksty, tom 2 - wyłącznie pełne teksty. W tomie 1 za-

^{2/} Tools for Knowledge Organization and the Human Interface. Proc. 1st Int. ISKO Conference, Darmstadt, 14-17 Aug. 1990. Ed. R. Fagmann Frankfurt/Main: INDEKS Verlag. Vol. 1:1990. Vol. 2:1991.

sygnalizowano 21 referatów, których pełne teksty pojawiły się dopiero w tomie 2. Ogółem w obu tomach opublikowano 51 pełnych tekstów referatów; w wypadku trzech referatów opublikowano (w tomie 1) wyłącznie abstrakty. Na ogólną liczbę 51 referatów opublikowanych in extenso osiem przedstawiono na sesji plenarnej, a pozostałe - tj. 43 referaty - w ramach zespołów problemowych. W obrębie sesji plenarnej zaprezentowano następujące referaty:

Charles R. HILDRETH: Użytkownicy końcowi a ustrukturalizowane przeszukiwanie katalogów dostępnych online: najnowsze wyniki badań naukowych ;

Masanobu FUJIKAWA: Teoria pojęć i analiza fasetowa jednostek wiedzy - ze szczególnym uwzględnieniem badań w zakresie sztucznej inteligencji ;

Alan GILCHRIST: Organizacja wiedzy a interfejs użytkownika ;

Otto SECHSER: Zagadnienia klasyfikacji w bazach danych utworzonych z czytelnych dla maszyny danych tekstowych ;

Brian C. VICKERY: Zasady klasyfikacyjne w inteligentnych interfejsach ;

Roland HJERPPE: Struktura do charakteryzowania systemów organizacji wiedzy: początkowa (wyjściowa) baza danych do porównań i oceny ;

Karen Markey DRABENSTOTT: Doświadczenia z katalogami online w USA przy użyciu systemu klasyfikacji jako narzędzia wyszukiwania tematycznego ;

Hanne ALBRECHTSEN: Pojęcia związane oprogramowaniem: Organizacja wiedzy a interfejs użytkownika.

Obrady konferencji podsumował Eric de Grolier, który na tle jej wyników sformułował dziesięć "problemów" (cudzysłów E. de Groliera). Warto odnotować przynajmniej pierwsze trzy problemy dostrzeżone przez

wspomnianego autora. Pierwszy dotyczy granic automatyzacji, tego co moglibyśmy nazwać "robotyzacją" systemów wyszukiwania informacji. Wydaje się, że realizacja całkowicie automatycznego, a zarazem w pełni zadowalającego systemu informacyjno-wyszukiawczego jest taką samą iluzją jak urzeczywistnienie w pełni automatycznego przekładu; dlatego należałoby mówić skromniej o wspomaganym komputerowo przekładzie i o wspomaganym komputerowo systemach wyszukiwania informacji. Drugim problemem jest użytek, jaki możemy uczynić z nowych technologii, takich jak sztuczna inteligencja, publicznie dostępne katalogi online, hipertekst, hiperśrodki (hypermedia) itp. W tym zakresie możliwe są dwa rozwiązania: 1) przystosowanie do nowych środków starych, tradycyjnych, konwencjonalnych systemów, odziedziczonych po XIX w. lub nawet po wcześniejszych stuleciach; 2) zbudowanie zupełnie nowych schematów.

Trzecią sprawą poruszoną przez E. de Groliera jest znaczenie informacji faktograficznej.

Tom I omawianych materiałów konferencyjnych oprócz referatów zawiera przedmowę (napisaną wspólnie przez przewodniczącą Rady Wykonawczej ISKO I. Dahlberg i wiceprzewodniczącego Rady Roberta Fugmanna), wprowadzenie (pióra R. Fugmanna) oraz alfabetyczną listę autorów referatów. Tom 2 obok referatów i omówionego wyżej podsumowania zawiera przedmowę autorstwa R. Fugmanna, alfabetyczną listę uczestników Konferencji (153 osoby), indeks nazw (tj. indeks nazw osobowych oraz nazw własnych instytucji, systemów klasyfikacji, projektów badawczych, programów komputerowych, wydarzeń itp.) oraz indeks rzeczowy (Subject Index).

Druga Międzynarodowa Konferencja ISKO odbędzie się w Madrasie (Indie) w dniach 26-29 sierpnia 1992 r., a jej tematem będą "Paradygmaty poznawcze w organizacji wiedzy" (Cognitive Paradigms

in Knowledge Organization). ISKO było też współsponsorem Piątej Międzynarodowej Konferencji Badawczej na temat Badań Klasyfikacyjnych (5th International Study Conference on Classification Research), zorganizowanej głównie przez Komitet FID/CR. Konferencja odbyła się w Toronto (Kanada) w dniach 24-28 czerwca 1991 r., a jej tematem przewodnim były "Badania klasyfikacyjne na potrzeby reprezentacji i organizacji wiedzy" (Classification Research for Knowledge Representation and Organization).

Konferencje naukowe są organizowane lub współorganizowane także przez krajowe (narodowe) sekcje ISKO. Krajowa sekcja ISKO w Czechosłowacji (National Chapter of ISKO in Czechoslovakia) zorganizowała we współpracy z centralą ISKO oraz innymi organizacjami i instytucjami międzynarodową konferencję poświęconą "Organizacji wiedzy, technologii i zarządzaniu dostępem do informacji" (Knowledge Organization, Technology and Information Access Management). Konferencja ta odbyła się 14-15 maja 1991 r. w Bratysławie. Niemiecka sekcja ISKO organizuje swoją drugą sesję w Weilburgu (Hesja) w dniach 16-18 października 1991 r.; tematem sesji są "Poznawcze fundamenty porządkowania i przedstawiania wiedzy" (Kognitive Ausätze zum Ordnen und Darstellen von Wissen).

Eugeniusz Ścibor

18 września 1991 r.

KRONIKA

KOMPUTEROWE BAZY DANYCH W OŚRODKU INFORMACJI NAUKOWEJ PAN

W 1990 r. DIN PAN zainicjował prace zmierzające do utworzenia systemu usług informacyjnych opartego na zagranicznych komputerowych bibliograficznych i faktograficznych bazach danych. Wybór i zakup baz danych poprzedził sondaż przeprowadzony w środowisku naukowym mający na celu wytypowanie baz danych najbardziej przydatnych i najlepiej odpowiadających potrzebom pracowników nauki.

Dla nauk przyrodniczych wybrano następujące bazy bibliograficzne:

Current Contents - Life Sciences

Current Contents - Agriculture, Biology and Environmental Sciences
oraz bazy faktograficzne - CAB-Abstracts, Biological Abstracts.

Dla potrzeb szerokiego grona użytkowników informacji naukowej wybrano także renomowane bazy cytowań naukowych szeroko rozpowszechnione w świecie - Social Science Citation Index oraz Science Citation Index.

Bazy typu Current Contents wydawane przez Instytut Informacji Naukowej (ISI) w Filadelfii są najbardziej popularnym wydawnictwem zawierającym dane bibliograficzne artykułów publikowanych w wiodących czasopiśmie naukowych na całym świecie. Bazy te były wydawane do roku 1989 w formie drukowanych zeszytów, aktualizowanych cotygodniowo. od 1989 r. ISI wzbogacił ofertę wydawniczą o cotygodniowe dyskiety. Current Contents są najczęściej wykorzystywanym źródłem bieżącej informacji bibliograficznej, nierzadko prezentującej tytuły prac naukowych zaledwie w 2 tygodnie po

ich opublikowaniu, co umożliwia najszybszy przepływ informacji naukowej. Current Contents wydawane są w 7 wersjach tematycznych - Nauki Biologiczne (Life Sciences), Medycyna Techniczna (Clinical Medicine), Rolnictwo (Agriculture), Biologia i Ochrona Środowiska (Biology and Environmental Sciences), Nauki Fizyczne, Chemiczne i Nauki o Ziemi (Physical, Chemical and Earth Sciences), Nauki Techniczne i Stosowane (Engineering, Technology and Applied Sciences), Socjologia i Nauki Społeczne (Social and Behavioral Sciences) oraz Sztuka i Nauki Humanistyczne (Art and Humanities).

Od 1990 r. OIN PAN udostępnia bazę Current Contents - Life Sciences zawierającą dane bibliograficzne 1200 czasopism rocznie w następujących zbiorach tematycznych: Nauki Zootechniczne i Nauki o Roślinach (Animal and Plant Sciences), Biochemia (Biochemistry), Biofizyka (Biophysics), Chemia (Chemistry), Medycyna Kliniczna (Clinical Medicine), Biologia Eksperymentalna i Medycyna (Experimental Biology and Medicine), Immunologia (Immunology), Mikrobiologia i Biologia Komórki (Microbiology and Cell Biology), Biologia Molekularna i Geny (Molecular Biology and Gene), Wielodyscyplinarny (Multidisciplinary), Neurologia i Zachowanie (Neurosciences and Behaviour), Farmakologia (Pharmacology), Fizjologia (Physiology).

Na szeroko rozpowszechnioną przez OIN PAN w środowisku naukowym ofertę usług informacyjnych odpowiedziało około 70 zakładów naukowych, zgłaszając zamówienia na wyszukiwanie danych wg tzw. profili wyszukiwawczych. Z tego 25 użytkowników zgłosiło chęć stałego korzystania z zaoferowanych usług.

W 1991 r. OIN rozszerzył zakres usług informacyjnych o kolejne bazy z serii Current Contents - Rolnictwo, Biologia i Ochrona Środowiska (Agriculture Biology and Environmental Sciences).

Nauki Techniczne i Stosowane (Engineering Technology and Applied Sciences), Nauki Fizyczne, Chemiczne i Nauki o Ziemi (Physical, Chemical and Earth Sciences), przy czym eksploatowaną już bazę Life Sciences, a także wymienione powyżej, za wyjątkiem Current Contents - Engineering, Technology and Applied Sciences wydawca wzbogacił o abstrakty.

Wszystkie bazy są cotygodniowo aktualizowane co stanowi o ich wysokiej wartości dla użytkowników.

Baza CC-Agriculture Biology and Environmental Sciences zawiera dane bibliograficzne wraz z abstraktami z 900 czasopism w następujących działach tematycznych: Chemia Rolna (Agricultural-Chemistry), Rolnictwo (Agriculture), Agronomia (Agronomy), Nauki Zootechniczne (Animal Sciences), Nauki o Środowisku Wodnym (Aquatic Sciences), Biologia (Biology), Biotechnologia i Nauki Stosowane (Biotechnology and Applied Science), Entomologia (Entomology/Pest Control), Ekologia (Environmental/Ecology), Żywność i Żywienie (Food Nutrition), Wielodyscyplinary (Multidisciplinary), Nauka o Roślinach (Plant Sciences), Medycyna Weterynaryjna (Veterinary Medicine).

Baza CC-Engineering, Technology and Applied Sciences referuje bibliografie 800 czasopism rocznie w działach tematycznych: Przestrzeń Kosmiczna (Aerospace), Chemia (Chemical), Komunikacja/Informacja (Communication, Information), Elektryka i Elektronika (Electric and Electronic), Inżynieria Lądowa i Ochrona Środowiska (Environmental/Civil Engineering), Górnictwo Węglowe i Naftowe (Geo/Petro/Mining), Aparatura Kontrolno-Pomiarowa (Instrumentation/Control Materials), Mechanika (Mechanics), Metalurgia (Metallurgy), Badania Jądrowe (Nuclear), Optyka i Akustyka (Optics and Acoustics).

Baza CC-Physical, Chemical and Earth Sciences zawiera dane bibliograficzne wraz z abstraktami z 800 czasopism z następujących działów tematycznych: Analityczno-Nieorganiczny (Analytical-Inorganic), Fizyka Stosowana (Applied Physics), Chemia (Chemistry), Nauki o Ziemi (Earth Sciences), Matematyka (Mathematics), Wielodyscyplinarny (Multidisciplinary), Chemia Organiczna i Chemia Polimerów (Organic Chemistry/Polimer), Chemia Fizyczna/Fizyka (Physical-Chemistry/Physics), Nauki o Kosmosie (Space Sciences).

Eksploatowane bazy mają swoich stałych użytkowników, szczególnym zainteresowaniem cieszy się baza Life-Sciences wykorzystywana najszerszej przez środowisko naukowe akademii medycznych w całym kraju.

Również w 1991 r. do pakietu usług informacyjnych włączono bazy na dyskach optycznych CAB-Abstracts oraz Science Citation Index i Social Science Citation Index.

CAB-Abstracts wydawane przez Commonwealth Agriculture Bureau jest wiodącą w świecie bazą faktograficzną nauk rolniczych zawierającą abstrakty w języku angielskim z ponad 10 tysięcy czasopism, monografii, materiałów konferencyjnych, raportów technicznych i patentów wydawanych w 100 krajach. Abstraktowane czasopisma dotyczą następujących działów tematycznych: Nauki Zootechniczne (Animal Science and Production), Weterynaria i Choroby Pasożytnicze (Veterinary Science and Parasitic Diseases), Nauki o Roślinach Uprawnych (Crop Science and Production), Leśnictwo (Forestry), Ochrona Roślin (Crop Protection), Rozwój Ekonomiczny i Socjologia (Economics Development and Sociology), Maszyny i Konstrukcje (Machinery and Buildings), Żywnienie Człowieka (Human Nutrition), Biotechnologia i Biodegradacja (Biotechnology/Biodeterioration).

Roczna kumulacja zawiera 130 tys. abstraktów mieszczących się w 53 zbiorach tematycznych z czego 65% pochodzi z czasopism anglojęzycznych a pozostałe 35% z czasopism wydawanych w 75 innych językach.

Baza CAB-Abstracts zawiera około 10 mln abstraktów z lat 1984-1990 i jest uzupełniana na początku każdego roku kalendarzowego danymi za rok ubiegły. Z bazy dotychczas skorzystało 27 użytkowników, głównie z Akademii Rolniczych.

Indeksy cytowań naukowych Social Science Citation Index i Science Citation Index są bazami wydawanymi przez Instytut Informacji Naukowej w Filadelfii. Stanowią one uporządkowany wykaz cytowanych dokumentów, któremu towarzyszy wykaz dokumentów cytujących. Każda roczna kumulacja cytuje od 25-50% spośród 5-10 milionów artykułów i książek z całego świata.

Science Citation Index indeksuje około 3300 najważniejszych naukowych i technicznych czasopism z całego świata w następujących obszarach wiedzy: Nauki Biologiczne (Life Sciences), Nauki Fizyczne (Physical Sciences), Chemia (Chemistry), Nauki o Ziemi (Earth Sciences), Rolnictwo (Agriculture), Ochrona Środowiska (Environmental Sciences), Medycyna Kliniczna (Clinical Medicine), Nauki Techniczne i Stosowane (Engineering Technology and Applied Sciences), Matematyka i Informatyka (Mathematic and Computer Sciences).

Social Science Citation Index zawiera 1400 indeksowanych czasopism odnoszących się do następujących dziedzin wiedzy: Antropologia i Archeologia (Anthropology and Archeology), Etnografia (Area and Ethnic Group Studies), Gospodarka i Finanse (Business and Finance), Komunikacja (Communications), Zastosowanie Komputerów i Cybernetyka (Computer Applications and Cybernetics), Kryminologia i Penologia (Criminology and Penology), Demografia (Demo-

graphy), Badania nad Szkolnictwem (Educational Research), Ekonomika (Economics), Geografia (Geography), Higiena i Zdrowie Społeczne (Hygiene and Public Health), Informacja i Bibliotekoznawstwo (Information and Library Sciences), Stosunki Międzynarodowe (International Relations), Prawo (Law), Zarządzanie i Marketing (Management and Marketing), Polityka (Policy Studies), Nauki Polityczne (Political Science), Psychologia i Psychiatria (Psychology and Psychiatry), Socjologia (Sociology), Planowanie i Zagospodarowanie Przestrzeni (Urban Planning and Development).

Do obydwu baz dołączony jest Journal Citation Reports, który zawiera roczną statystykę cytowań czasopism pozwalającą na ocenę ich wartości.

Przydatności indeksów cytowań dla potrzeb nauki nie sposób przecenić, stanowią one bowiem źródło wszechstronnych informacji o badaniach naukowych, ośrodkach specjalizujących się w poszczególnych problemach badawczych, czy też o pracach inspirowanych rozwój danyh kierunków nauki.

Analiza publikacji znajdujących największą oddziały w cytowaniach może być przydatna do rozpoznawania kierunków rozwojowych nauki i wytyczenia własnej tematyki badawczej, a także kształtowania polityki naukowej.

Informacje o liczbie cytowań poszczególnych autorów mogą stanowić elementy oceny wartości osiągnięć naukowych, a tym samym kadry naukowej. Ośrodek dysponuje indeksami cytowań za lata 1980-1991, każdy rok jest zarejestrowany na oddzielnym dysku, baza jest uzupełniana kwartalnie. Użytkownikami bazy jest szerokie grono pracowników naukowych z całego kraju.

Szybka, właściwa i dokładna informacja naukowa ma duży wpływ na realizację prac badawczych w badaniach podstawowych i stosowa-

nych a także na decyzje odnośnie podejmowania badań decydujących o postępie nauki.

Mając na uwadze powyższe cele Ośrodek prowadzi stałe prace zmierzające do poszerzenia i udoskonalenia działalności informacyjnej zgodnie z aktualnymi i przyszłymi potrzebami środowiska naukowego.

H. Majewska

12.11.1991 r.

POSIEDZENIE KOLEGIUM DYREKTORÓW I KIEROWNIKÓW PLACÓWEK ODDZIAŁU PAN W KRAKOWIE

Dnia 4 kwietnia 1991 r. odbyło się w lokalu krakowskiej Pracowni OIN PAN (ul. Św. Tomasza 30/6) posiedzenie Kolegium Dyrektorów i Kierowników Placówek PAN w Krakowie. Zebraniu przewodniczył prof.dr hab. J. Haber, dyrektor Instytutu Katalizy PAN w Krakowie. W posiedzeniu wziął udział również Dyrektor OIN PAN w Warszawie - doc.dr hab. A. Gromek oraz inni zaproszeni goście (łącznie około 40 osób).

Zebrań było poświęcone omówieniu zagadnień dotyczących automatyzacji procesów informacyjno-bibliotecznych w bibliotekach instytutowych i zakładowych Oddziału PAN w Krakowie. Obradom towarzyszyła prezentacja działalności OIN PAN w tej dziedzinie (Dyr. A. Gromek) oraz pokaz działania baz danych na dyskach CD-ROM (Current Contents, Science Citation Index, Social Science Citation

Index, System Spirs) przeprowadzony przez mgr I. Boguską oraz mgr D. Wywiórka z OIN PAN w Warszawie.

W trakcie spotkania padło wiele interesujących propozycji. Informacje na temat sieci EARN oraz możliwości włączenia się do niej przedstawił koordynator tej sieci w ramach krakowskiego Oddziału PAN - dr M. Kruszyński (Instytut Mechaniki Górotworu PAN). Z propozycjami włączenia się do tej sieci wystąpili m.in. dr A. Fitowa (OIN PAN) i mgr B. Morawska-Nowak (Instytut Farmakologii PAN). Dr M.T. Czerwenka (Instytut Katalizy i Fizykochemii Powierzchni PAN) wystąpił z propozycją koordynowania przez PAN wszelkich zakupów sprzętu i programów komputerowych oraz licencji na niektóre programy komputerowe dla wszystkich placówek PAN. Dr A. Fitowa zaproponowała powołanie przy Kolegium Dyrektorów Placówek PAN specjalnej Komisji d/s wspomaganie automatyzacji bibliotek. Komisję powołano w składzie: dr M.T. Czerwenka, dr M. Kruszyński, dr W. Kraus, dr A. Fitowa, mgr inż. J. Paśnik. Ustalono, że siedziba nowo powołanej komisji znajdować się będzie w krakowskiej Pracowni OIN PAN.

W. Babik

6.06.1992 r.

DOSTĘP KRAKOWSKIEJ PRACOWNI INFORMACJI NAUKOWEJ OIN PAN DO SIECI EARN

W czerwcu br. krakowska Pracownia Informacji Naukowej OIN PAN uzyskała możliwość wykorzystywania krakowskiego węzła EUROPEAN RESEARCH NETWORK (EARN). Węzeł ten znajduje się w Regionalnym

Centrum Obliczeniowym CYFRONET w Krakowie, ul. Reymonta 4a. Na razie Pracownia uzyskała tylko możliwość korzystania z węzła na miejscu. Ponieważ Pracownia dysponuje odpowiednim sprzętem komputerowym, zamierza się w najbliższym czasie zainstalować tu odpowiedni modem, który umożliwi bezpośrednie kontaktowanie się z siecią za pomocą sieci telefonicznej.

W ramach dostępu do EARN można obecnie korzystać wyłącznie z poczty elektronicznej umożliwiającej wysyłanie i odbiór komunikatów i listów pomiędzy użytkownikami. Użytkownicy sieci mogą także dokonywać transferu zbiorów różnych typów (danych, programów, dokumentów).

Z ramienia Pracowni osobą odpowiedzialną za korzystanie z wymienionej sieci jest mgr Wiesław Babik. Jako użytkownik sieci Pracownia otrzymała odpowiedni podręcznik oraz będzie regularnie otrzymywać "Biuletyn EARN".

W. Babik

6.01.1992 r.

**SEMINARIUM POLSKO-ROSYJSKIE NT. PROBLEMÓW METODYKI
OPRACOWANIA PRZEDMIOTOWEGO W SYSTEMACH TRADYCYJNYCH
I ZAUTOMATYZOWANYCH**

Warszawa, 6-8 listopada 1991 r.

W listopadzie 1991 r. odbyło się w Bibliotece Narodowej (BN) drugie polsko-rosyjskie seminarium nt. problemów metodyki opracowania przedmiotowego. Pierwsze, zorganizowane przez Bibliotekę im. M.E. Sałtykowa-Szczedrına w Leningradzie, odbyło się w czerwcu 1989 r.^{1/} Zarówno w pierwszym, jak i drugim spotkaniu udział wzięli przedstawiciele BN, Biblioteki im. Sałtykowa-Szczedrına, Wszechzwiązkowej Izby Książki (WIK) w Moskwie oraz Biblioteki INION Akademii Nauk ZSRR.

W tegorocznym seminarium wzięło udział ponad 30 osób, głównie z bibliotek uniwersalnych naukowych, zainteresowanych opracowaniem przedmiotowym.

Głównym zagadnieniem omawianym na spotkaniu było centralne opracowanie przedmiotowe oraz wykorzystanie gotowych opisów przez biblioteki. W Polsce są to m.in. opisy przedmiotowe przygotowane na potrzeby bibliografii narodowej ("Przewodnik Bibliograficzny"); korzysta z nich wiele bibliotek, tworząc na ich podstawie katalogi przedmiotowe. W ZSRR rolę centralnego ośrodka przygotowującego opisy na potrzeby bibliografii pełni Wszechzwiązkowa Izba Książki.

^{1/}Omówienie tego seminarium - "Zagadnienia Informacji Naukowej" 1989 nr 2(55) s. 201-205.

Również jej opisy wykorzystywane są przez biblioteki tworzące katalogi przedmiotowe. Zagadnieniu temu były poświęcone trzy referaty: Eduarda L. Prizmenta (WIK) "Centralne opracowanie przedmiotowe. Spojrzenie od wewnątrz" oraz Walentyny P. Suworowej (Biblioteka Sałtykowa-Szczedrina) "Wykorzystanie haseł przedmiotowych centralnie opracowanych w bibliotekach ZSRR" i Ewy Stępniakowej (BN) "Centralne opracowanie przedmiotowe. Punkt widzenia użytkownika". Wnioski ogólne wynikające z referatów i dyskusji były zbieżne. Mianowicie opisy przedmiotowe przygotowane centralnie mogą być podstawą tworzenia katalogów przedmiotowych, ale w każdej bibliotece wymagają one indywidualnych korekt i rozstrzygnięć. Nie mogą być przejmowane w sposób mechaniczny. Należy je traktować jako adnotacje sformułowane w języku haseł przedmiotowych, które można w pełni lub z modyfikacjami wykorzystać w poszczególnych katalogach.

Ważnym zagadnieniem dyskutowanym na seminarium była sprawa obiektywizmu w wyrażaniu treści dokumentów. Jest to szczególnie istotne w praktyce rosyjskiej, gdzie zakłada się uwzględnianie aspektu polityczno-partyjnego. Zgodzono się, że hasła przedmiotowe powinny być obiektywne w stosunku do treści dokumentów, ale jednocześnie zwracano uwagę na niemożliwość osiągnięcia obiektywizmu, ponieważ te same przedmioty i zjawiska są postrzegane różnie przez różne osoby.

Omawiano też sprawy tematowania literatury pięknej - robi to Wszéchniązdkowa Izba Książki, w Polsce jest to rzadkością i dotyczy tylko gatunków z pogranicza literatury pięknej i antologii o określonej tematyce.

Nieco odbiegającym od przedstawionej wyżej problematyki był referat Janiny Trzecińskiej (BN) "Normalizacja nazw geograficznych w hasłach opisu bibliograficznego".

W drugim dniu obrady koncentrowały się wokół zagadnień automatyzacji katalogów. Referaty nie miały tak jednolitego charakteru, jak w przypadku referatów o katalogach tradycyjnych. Dwa z nich dotyczyły prac prowadzonych we własnych bibliotekach. Był to referat Leonida W. Sacharnego i Andreja C. Masewicza "Lokalne zautomatyzowane systemy informacyjno-wyszukiwawcze wykorzystujące hasła przedmiotowe w Bibliotece im. M.E. Sałtykowa-Szczedrina w Leningradzie" oraz referat Sofii K. Wileńskiej "Niektóre problemy automatyzacji katalogu przedmiotowego (z doświadczeń INION AN ZSRR).

W Bibliotece w Leningradzie tworzy się małe autonomiczne bazy danych. Jest ich kilka: "Architektura Sankt-Petersburga - Leningradu", "Toponimika Sankt-Petersburga - Leningradu", "Inkunabuły GPB", "Nowe piśmiennictwo o Sankt-Petersburgu - Leningradzie", "Memoriał". Są to bazy bibliograficzne i bibliograficzno-faktograficzne działające w oparciu o pakiet programowy micro-ISIS, uzupełniony programami własnymi. W wyszukiwaniu wykorzystuje się hasła przedmiotowe traktowane jak słowa kluczowe łączone za pomocą operatorów algebry Boole'a.

Referat S. Wileńskiej przedstawiał ogólne zagadnienia związane z automatyzacją katalogu przedmiotowego w INION. Autorka skupiła się na słowniku haseł przedmiotowych prowadzonym jako baza danych, w której to bazie oprócz tekstu artykułu przedmiotowego zapisuje się również symbole Rubrykatora GASNTI i umowne oznaczenia tzw. kompleksów tematycznych, do których należą hasła przedmiotowe. Głównym problemem INION jest połączenie katalogu przedmiotowego z

działającym od kilku lat ogólnym zautomatyzowanym systemem INION.

Referat Jadwigi Sadowskiej (BN) "Problemy wyszukiwania rzeczowego w katalogach online" był opracowany na podstawie piśmiennictwa zagranicznego i wykraczał poza zagadnienia opracowania przedmiotowego. Przedstawiono w nim różnorodność technik wyszukiwawczych, ich efektywność i akceptację przez użytkowników. Z przytoczonych badań amerykańskich wynika, że główny problem to odmienność słownictwa, którym posługują się bibliotekarze formułujący charakterystyki wyszukiwawcze i użytkownicy formułujący pytania informacyjne. W badaniach zwracano też uwagę na przydatność normalizowanych słowników (okazuje się, że użytkownicy w ogóle nie zauważają ich istnienia), rolę języka naturalnego, wykorzystanie symboli klasyfikacyjnych, rozbudowywanie charakterystyk wyszukiwawczych.

W związku z zagadnieniem zgodności języka informacyjno-wyszukiwawczego z oczekiwaniami użytkowników przedstawiono pokrótce wyniki takich badań (J. Sadowska) przeprowadzonych wśród studentów Instytutu Bibliotekoznawstwa i Informacji Naukowej UW oraz użytkowników katalogów Biblioteki Uniwersytetu Łódzkiego i Książnicy Miejskiej w Toruniu.^{2/} Badania dotyczyły akceptacji wybranych reguł języka haseł przedmiotowych.

W dyskusji zwracano uwagę na konieczność tworzenia kartotek autorytatywnych nazw osobowych, geograficznych, ciał zbiorowych, haseł przedmiotowych. Za szczególnie ważne uznano podjęcie prac normalizacyjno-metodycznych związanych z systemami zautomatyzowanymi.

^{2/} Szerzej o wynikach tych badań - bieżący nr ZIN s. 125-134.

Trzeci dzień seminarium przeznaczony był na zwiedzanie Biblioteki Narodowej oraz zapoznanie się z katalogiem przedmiotowym i mikrokomputerowym systemem wyszukiwawczym (MAK), w którym wykorzystuje się hasła przedmiotowe i w którym utrzymywany jest słownik haseł przedmiotowych BN.

Za istotny temat dalszej współpracy uznano sprawy związane z przygotowaniem metodyki opracowania przedmiotowego dla zautomatyzowanych katalogów bibliotecznych.

Jadwiga Sadowska

25.11.1991

HUMANISTA PRZY KOMPUTERZE

Szkoła jesienna w Karpaczu

W dniach 12-15 listopada 1991 r. spotkały się w Karpaczu osoby zainteresowane sprawami automatyzacji bibliotek i kształcenia studentów bibliotekoznawstwa w zakresie podstaw informatyki. Organizatorami imprezy byli: Katedra Bibliotekoznawstwa i Informatyki Naukowej Uniwersytetu Łódzkiego, Instytut Bibliotekoznawstwa Uniwersytetu Wrocławskiego oraz Biblioteka Narodowa. Formuła spotkania odbiegała od tradycyjnego schematu organizacji konferencji - zwracano uwagę na zachowanie interakcyjnego trybu pracy, wzajemne uczenie się, doskonalenie, dzielenie doświadczeniami i refleksjami (stąd nazwa szkoła a nie konferencja). Większości

wystąpień towarzyszyły pokazy komputerowe. Przez cały czas trwania Szkoły uczestnicy mogli też zapoznać się i testować oprogramowania prezentowane w poszczególnych wystąpieniach.

Praca Szkoły rozpoczęła się we Wrocławiu, gdzie miał miejsce między innymi przygotowany przez dr. inż. Kazimierza Chorosia, mgr. inż. Grzegorza Karpowicza i mgr. inż. Dariusza Króla pokaz mikrokomputerowych systemów biblioteczo-informacyjnych opracowanych przez studentów Zakładu Systemów Informacyjnych Politechniki Wrocławskiej. Drugi dzień pracy Szkoły, któremu przewodniczyła prof. dr hab. Hanna Tadeusiewicz z Uniwersytetu Łódzkiego, otworzył mgr Jan Wierzbowski (Biblioteka Narodowa) prezentacją oprogramowania MAK. Dalszy ciąg prac koncentrował się wokół tematu przewodniego "Komputer w służbie historii książki". Z referatami i pokazami wystąpili: mgr Władysław Szczęch (Zakład Bibliotekoznawstwa i Informacji Naukowej UJ) "Badania nad reklamą książki w Polsce w XVIII w.: wspomaganie komputerowe", dr Henryk Hollender (Instytut Bibliotekoznawstwa i Informacji Naukowej UW) "Polska w angielskiej kulturze umysłowej XVIII w.: możliwości poszukiwań w bazie danych ESTC", mgr Katarzyna Ślaska (BN) "Wykorzystanie pakietu MAK w Zakładzie Starych Druków Biblioteki Narodowej", mgr Jolanta Byczkowska-Sztaba (Zakład Zbiorów Muzycznych BN) "Zastosowanie MAKa dla centralnych katalogów rękopisów i starych druków muzycznych BN", mgr Piotr Lewkowicz (Katedra Bibliotekoznawstwa i Informacji Naukowej) "Potrzeby zastosowań komputera w badaniach oraz w dydaktyce historii książki i bibliotek", dr Małgorzata Leyko (Instytut Kulturoznawstwa UŁ) "O programie TEATROLOGIA-ENTER". Gościem Szkoły był dr Michael J. Crump z British Library, który wystąpił z następującymi referatami: "ESTC (Eighteenth Century Short Title Catalogue) - system informacji

o starych drukach angielskich: stan i rozwój oraz strategia wyszukiwania w bazie danych ESTC na CD-ROM", "Poszukiwania w ESTC a badania angielskiej książki XVIII-wiecznej: przykłady, organizacja współpracy, seminaria", "Zarządzanie i problemy kooperacji w systemie ESTC: doświadczenia i propozycje dla strony polskiej".

Motywy przewodnim drugiego dnia Szkoły było hasło "Między praktyką i badaniami". Pracom przewodniczył doc. dr Czesław Daniłowicz z Politechniki Wrocławskiej. Głos zabrali: mgr Zdzisław Dobrowolski (IBIN UW) "Poetyka tzw. helpów i instrukcji programów komputerowych: dlaczego uczenie się nowych programów jest tak irytujące?!", mgr Mariusz Robowski (KBiIN Uł) "Przegląd oprogramowania użytkowego do zastosowania w bibliotekach", doc. dr Czesław Daniłowicz "Użytkownicy - pośrednicy - komputerowe systemy wyszukiwania informacji", ks. Krzysztof Gonet (Seminarium Duchowne w Warszawie) "Zastosowanie pakietu MAK w lokalnej sieci Biblioteki Seminarium Duchownego w Warszawie", mgr Maria Styblińska (Zakład Organizacji Informacji Instytutu INT Uniwersytetu Śląskiego) "Sieci komputerowe ze szczególnym uwzględnieniem EARN", mgr Mariola Nałęcz (Zakład Zbiorów Muzycznych BN) "Zastosowanie oprogramowania MAK w katalogowaniu nut drukowanych i rękopiśmiennych oraz utworów muzycznych publikowanych w czasopiśmie", mgr Aleksander Radwański (Instytut Bibliotekoznawstwa Uniwersytetu Wrocławskiego) "Badanie struktury danych w zbiorach "Przewodnika Bibliograficznego", mgr Maria Śliwińska (Zakład Bibliotekoznawstwa i Informacji Naukowej UMK) "Konwersja danych "Przewodnika Bibliograficznego" w formacie SABINA na format FOKA w systemie ISIS" oraz "Baza danych bibliografii dorobku pracowników Uniwersytetu Mikołaja Kopernika w Toruniu - wynik pracy studenckiej w Zakładzie Bibliotekoznawstwa i Informacji Naukowej UMK", mgr Jolanta Stępniaak (Biblioteka

Naukowa Zamku Królewskiego) "Porównanie pakietów ISIS i MAK", dr Wandla Pindłowa (ZBiIN UJ) "Bazy danych jako temat prac magisterskich kierunku bibliotekoznawstwo i informacja naukowa". Głos zabrała także mgr Małgorzata Kamińska z Wojewódzkiej Biblioteki Publicznej w Białymstoku, która przedstawiła funkcjonujące w Bibliotece dwa zautomatyzowane podsystemy - rejestracji wypożyczeń i katalogowy. Podzieliła się także informacjami dotyczącymi ocen czytelników.

Trzeci dzień Szkoły był poświęcony dydaktycznym problemom informatyki na studiach bibliotekoznawczych. Pracom przewodniczyła prof. dr hab. Anna Sitarska. Rozpoczęto od dyskusji panelowej, w której wzięli udział pracownicy dydaktyczni (mgr M. Śliwińska, mgr Z. Dobrowolski, mgr A. Radwański, mgr M. Robowski, mgr W. Szczęch), prowadzący zajęcia z podstaw informatyki w szkołach wyższych kształcących bibliotekarzy. Następnie wygłoszono cztery komunikaty: dr Jadwiga Woźniak (IBiN UW) "Wykorzystanie bibliograficznej bazy danych w nauczaniu opracowania rzeczowego", dr Magda Kwiatkowska, mgr Małgorzata Mozer, mgr Katarzyna Indrzejczak (KBiIN UŁ) "Jak wprowadzić komputer do zajęć z bibliografii i bibliotekarstwa?", prof. dr hab. Anna Sitarska, dr Stanisława Kurek-Kokocińska (KBiIN UŁ) "Wykorzystanie podstaw informatyki w kursowych zajęciach informacji i w specjalizacji informacyjnej", dr Kazimierz Choroś (Biblioteka Główna Politechniki Wrocławskiej) "Szkolenie w zakresie mikrokomputerowych systemów biblioteczno-informacyjnych na Politechnice Wrocławskiej).

Teksty wszystkich referatów zostaną opublikowane w najbliższym numerze "Acta Universitatis Łodziensis".

Jadwiga Woźniak

3.12.1991 r.

SPIS TREŚCI

1. W. BABIK: Materiały kartograficzne jako źródło informacji.. 3
2. H. POPOWSKA, J. SOLAK: Uczenie się języka naturalnego przez system komputerowy (Na przykładzie systemu DABINAL)..... 37

Materiały i przyczynki

1. M. JANUSZEWSKA: Systemy hipertekstowe - rvs historyczny.... 85
2. E. ARTOWICZ: System informacji hungarologicznej węgierskiej biblioteki narodowej 101
3. J. SADOWSKA: Wybrane reguły metodyki opracowania przedmiotowego w ocenie użytkowników katalogów 125
4. J. ŚLADEWSKI: Problemy metodyki tematowania w bibliotekach uniwersalnych 135
5. J. WOŹNIAK: Kartoteka haseł wzorcowych dla publicznie dostępnego katalogu online sieci bibliotek Uniwersytetu Warszawskiego. Komunikat o stanie prac 149
6. H. SZARSKI: Nowa metoda dostępu do literatury naukowej - ADONIS 155
7. W. BABIK: European Academic Research Network (EARN) 163

Recenzje i omówienia

1. Propozycja nowej klasyfikacji dla materiałów kartograficznych - W. Babik 169
2. Uwagi do "Instrukcji tematowania i katalogu przedmiotowego - J. Woźniak 179
3. Narzędzia organizacji wiedzy a interfejs użytkownika - materiały pierwszej międzynarodowej konferencji Międzynarodowego Towarzystwa Organizacji Wiedzy - E. Ścibor 189

Krońnika

1. Komputerowe bazy danych w Ośrodku Informacji Naukowej PAN - H. Majewska 195
2. Posiedzenie Kolegium Dyrektorów i Kierowników Placówek Oddziału PAN w Krakowie - W. Babik 201

3. Dostęp krakowskiej Pracowni Informacji Naukowej OIN PAN do sieci EARN - W. Babik	202
4. Seminarium polsko-rosyjskie nt. problemów metodyki opracowania przedmiotowego w systemach tradycyjnych i zautomatyzowanych (Warszawa, 6-8 XI 1991 r.) - J. Sadowska	204
5. Humanista przy komputerze. Szkoła jesienna w Karpaczu - J. Woźniak	208

CONTENTS

1. W. BABIK: Cartographic materials as the source of information	3
2. H. POPOWSKA, J. SOLAK: Learning natural language by the computer system (The case of "DABINAL" system).....	37

Materials and Contributions

1. M. JANUSZEWSKA: Hypertext systems - the historical outline.	85
2. E. ARTOWICZ : System of hungarological information of the hungarian national library	101
3. J. SADOWSKA: Some rules of methodology of the subject cataloguing in users' estimation	125
4. J. ŚLADEWSKI: Methodological problems of appointment of subject headings in universal libraries	135
5. J. WOŹNIAK: The index of the pattern headings for the Warsaw University libraries' publicly available online catalogue. The state of the art.....	149
6. H. SZARSKI: ADONIS-a new method of making scientific literature accessible	155
7. W. BABIK: European Academic Research Network (EARN)	163

Reviews and Surveys

1. The proposal of the new classification for cartographic materials - W. Babik	169
2. Comments on the "Directions for appointment of Subject headings and for cataloguing" - J. Woźniak	179
3. Tools for organization of knowledge and user's interface - proceedings of the first international conference of the International Society for Knowledge Organization - E. Ści-bor	189
Chronicles	195

СОДЕРЖАНИЕ

1. В. БАБИК Картографические материалы как источник информации	3
2. X. ПОПОВСКА, Е. СОЛЯК Обучение эвм естественному языку (На примере системы DABINAL)	37

Материалы и примечания

1. М. ЯНУШЕВСКА Гипертекстовые системы - исторический очерк	85
2. Э. АРТОВИЧ Система венгроведческой информации Венгерской национальной библиотеки (Hungarica).....	101
3. Я. САДОВСКА Некоторые правила методики предметизации в оценке потребителей.....	125
4. Я. СЫЛЯДЕВСКИ Проблемы методики предметизации в универсальных библиотеках.....	135
5. Я. ВОЗЬНЯК Карточка образцовых заглавок для общедоступного диаганогов каталога сети библиотек Варшавского университета. Сообщение о состоянии работ.....	149

6. Х. ЖАРСКИ ADONIS - новый метод доступа к научной литературе.....	155
7. В. БАБИК Европейская академическая исследовательская сеть (EARA).....	163

Р е ц е н з и и и о б з о р ы

1. Предложение новой классификации для картографических материалов - В.Бабик.....	169
2. Примечания к "Instrukcji tematowania i katalogu przedmiotowego" - Я. Возняк.....	179
3. Инструменты организации знаний в интерфейсе потребителя - материалы первой международной конференции Международного общества организации знаний - Е.Свибор.....	189
Х р о н и к а	195

Szanowni Czytelnicy

Instytut Bibliotekoznawstwa i Informacji Naukowej Uniwersytetu Warszawskiego rozpoczął badania dotyczące **biblioteki w komunikacji naukowej**. W dniach 4-5 czerwca 1991 r. odbyło się w Warszawie zebranie konsultacyjno-dyskusyjne, w którym wzięło udział około 40 osób z bibliotek oraz ośrodków kształcenia bibliotekarzy na poziomie wyższym. Potwierdzono na nim celowość podjęcia wyżej wymienionych prac badawczych oraz konieczność stworzenia ośrodka dokumentacyjnego bibliotek naukowych. W związku z tym nasz instytut zamierza opracować centralną ogólnopolską kartotekę pracowników naukowych oraz specjalistów-praktyków związanych z bibliotekami i placówkami informacji naukowej. W zamierzeniu kartoteka ta byłaby „bankiem informacji” ułatwiającym kontakty w wypadku organizowania konferencji, zamawiania opracowań i recenzji, pomogłaby w powoływaniu grup ekspertów rozwiązujących konkretne problemy lub wpływających na kierunki polityki czy decyzje dotyczące bibliotekoznawstwa i bibliotek.

Pragniemy podkreślić, że kartotekę będziemy tworzyć na podstawie **dobrowolnych zgłoszeń**. Osoby zainteresowane wyżej przedstawionymi formami współpracy prosimy o przesłanie pod nasz adres (Instytut Bibliotekoznawstwa i Informacji Naukowej UW, Pracownia Bibliotek Naukowych, ul. Nowy Świat 69, 00-046 Warszawa) ankiet, w których będą uwzględnione następujące dane:

1. imię i nazwisko,
2. stopień naukowy,
3. miejsce pracy — nazwa, adres, telefon,
4. zajmowane stanowisko,
5. adres domowy i telefon,
6. kierunek i rok ukończenia studiów wyższych — podyplomowych lub innych,
7. zainteresowania naukowo-badawcze,
8. zainteresowania organizacyjno-zawodowe,
9. aktualnie opracowywane problemy (planowane),
10. ważniejsze publikacje (z ostatnich lat),
11. udział w ważniejszych pracach organizacyjno-metodycznych.

Liczymy na zrozumienie potrzeby istnienia centralnej informacji o osobach i opracowywanych przez nie zagadnieniach z zakresu bibliotekoznawstwa i informacji naukowej.

Wskazówki dla Autorów

Redakcja „Zagadnień Informatyki Naukowej” uprzejmie prosi Autorów o przestrzeganie następujących zasad przy nadsyłaniu materiałów.

Artykuł nie powinien przekraczać 30 stron maszynopisu formatu A4 wraz z przypisami i ewentualnymi tablicami (podwójny odstęp między wierszami, ok. 30 wierszy na stronie, margines 3,5 cm lewy, 1 cm prawy).

Maszynopis artykułu należy dostarczyć w 2 egzemplarzach. Ilustracje (tablice, wykresy) powinny być umieszczone na osobnych ponumerowanych stronach z zaznaczeniem ich miejsca w tekście.

Przypisy i bibliografię należy umieszczać na końcu tekstu i na osobnych stronach. Opis bibliograficzny powinien składać się z: nazwiska i imienia autora, tytułu publikacji, miejsca wydania, nazwy wydawnictwa i roku wydania oraz stron. W wypadku powoływania się na artykuł opis powinien zawierać: nazwisko i imię autora, tytuł artykułu, nazwę czasopisma, rok, tom (numer), strony lub stronę zawierającą cytowany fragment.

Każdy artykuł powinien być zaopatrzony w streszczenie autor-skie ok. 1/2 strony maszynopisu.

Autorzy proszeni są o podanie do wiadomości redakcji (oprócz imienia i nazwiska) również tytułu naukowego, nazwy i adresu instytucji, w której pracują, prywatnego adresu zamieszkania, numeru telefonu i ewentualnie, jeśli honorarium autorskie ma być przekazane na konto, numer konta PKO.

O przyjęciu do druku decyduje Rada Redakcyjna.

Maszynopisów artykułów nie zamówionych nie przyjętych do druku redakcja nie zwraca.

Uwaga! Redakcja przyjmuje materiały również na dyskietkach, najchętniej w kodzie ASCII.

