

*systemy organizacji wiedzy, Internet, biblioteki cyfrowe, intranet,
listy słownictwa kontrolowanego, klasyfikacje, kategoryzacje, taksonomie,
tezaury, sieci semantyczne, ontologie formalne*

BARBARA SOSIŃSKA-KALATA

Instytut Informacji Naukowej i Studiów Bibliologicznych UW

SYSTEMY ORGANIZACJI WIEDZY W ŚRODOWISKU SIECIOWYM

Termin *system organizacji wiedzy* (SOW) oznacza wszelkiego typu narzędzia organizowania informacji, które wspierają zarządzanie wiedzą w różnych środowiskach i dla różnych celów. Najbardziej reprezentatywnymi obszarami współczesnych poszukiwań i zastosowań SOW w środowisku sieciowym są biblioteki cyfrowe, internetowe i intranetowe systemy i serwisy informacyjne rozmaitych organizacji oraz projekt Semantycznego Weba. SOW stosowane w tych systemach obejmują narzędzia bardzo różnorodne, od stosunkowo prostych list słownictwa kontrolowanego do taksonomii, schematów klasyfikacyjnych i kategoryzacyjnych oraz tezaurów, sieci semantycznych i ontologii formalnych. W artykule wskazano podstawowe cechy wspólne wszystkim SOW oraz wyróżniono i omówiono ich najważniejsze typy.

1. WSTĘP

Termin *organizacja wiedzy* (ang. *knowledge organization*) wszedł do języka bibliotekoznawstwa i informacji naukowej za sprawą monografii z zakresu teorii klasyfikacji piśmiennictwa Henry'ego E. Blissa pt. *Organization of Knowledge in Libraries and the Subject Approach to Books* (Bliss, 1933). W zamierzeniu autora wskazywał najważniejsze zadanie, którego realizację powinno zapewnić rzeczowe uporządkowanie kolekcji bibliotecznej: umożliwienie wglądu w zawartość wiedzy zapisanej w składających się na nią dokumentach. Przez blisko sześćdziesiąt lat termin ten był rzadko stosowany, choć sporadycznie sięgali do niego kolejni autorzy, np. Christopher Needham w tytule książki *Organizing knowledge in libraries* wydanej w 1964 r. (Needham, 1964). Renesans Blissowskiego patrzenia na narzędzia porządkowania zasobów informacji, który nastąpił w końcu lat osiemdziesiątych XX wieku, wiąże się z co najmniej trzema wzajemnie powiązanymi czynnikami:

- dążeniem do uniwersalizacji metod wypracowanych w bibliotekarstwie i bibliografii,
- coraz szerszym zainteresowaniem metodami organizacji dostępu do informacji w środowiskach pozabibliotecznych, w tym przede wszystkim informatyków i specjalistów zarządzania działalnością biznesową,
- coraz silniejszą świadomością interdyscyplinarnej natury metodologii porządkowania informacji i potrzeby wykorzystania w jej rozwoju doświadczeń różnych dziedzin wiedzy.

Od 1989 r. istnieje International Society for Knowledge Organization (ISKO), międzynarodowe towarzystwo naukowe, które integruje środowisko badań nad metodami i narzędziami konceptualnego porządkowania wiedzy, wypracowanymi i stosowanymi w różnych dziedzinach, takich jak np. organizacja zasobów baz danych, bibliotek, słowników i Internetu (ISKO, 2005). Widoczny obecnie rozwój zainteresowania tą problematyką wiąże się jednak przede wszystkim z doskonaleniem informatycznych narzędzi zarządzania treścią zasobów informacyjnych oraz z oczekiwaniem jak najbardziej efektywnego wykorzystania tych narzędzi w działalności bibliotecznej i informacyjnej, która w coraz większym stopniu migruje do środowiska sieci komputerowych.

Od końca lat dziewięćdziesiątych coraz większą popularność zdobywa termin *systemy organizacji wiedzy* (SOW; ang. *knowledge organization systems, KOSs*), który oznacza wszelkiego typu schematy porządkowania informacji ułatwiające zarządzanie wiedzą w różnych środowiskach i dla różnych celów (Hodge, 2000). SOW obejmują zatem niejednorodny zestaw narzędzi o różnym stopniu złożoności, tworzonych i wykorzystywanych w różnych zastosowaniach przez specjalistów wywodzących się z różnych dziedzin. Stąd też narzędzia podobne, a czasami takie same pod względem cech funkcjonalnych i strukturalnych, są różnie nazywane. Bywa również, że SOW istotnie różniące się, w uproszczony sposób są ze sobą identyfikowane. Wytworzony chaos może prowadzić do niepotrzebnych nieporozumień, toteż warto podjąć próbę rozróżnienia i scharakteryzowania podstawowych typów systemów organizacji wiedzy stosowanych do porządkowania informacji w sieciowym środowisku cyfrowym.

2. POCHODZENIE METOD ORGANIZACJI INFORMACJI CYFROWEJ

Najbardziej reprezentatywnymi obszarami współczesnych poszukiwań i zastosowań różnego typu systemów organizacji wiedzy zapisanej w formie cyfrowej są biblioteki cyfrowe, internetowe i intranetowe systemy i serwisy informacyjne rozmaitych organizacji oraz projekt Semantycznego Webu.

Od bibliotekarzy i specjalistów informacji wymaga się dziś nie tylko gromadzenia informacji w formie elektronicznej, ale też jej organizowania w postaci bibliotek cyfrowych. Jednym z najważniejszych aspektów tego organizowania jest zapewnienie zrozumiałego, łatwego i możliwie precyzyjnego dostępu do zgromadzonych zasobów. Niezależnie bowiem od wiążących się z tym wątpliwości, biblioteki cyfrowe są postrzegane przede wszystkim właśnie jako punkty dostępu do informacji, z których użytkownicy korzystają samodzielnie, bez tradycyjnego pośrednictwa bibliotekarzy (Chowdhury, 2002).

Biblioteki cyfrowe przyjmują różną formę. Ich zasoby mogą być tworzone i przechowywane lokalnie, albo też mogą być rozproszone w globalnej sieci i integrowane za pomocą bibliotek wirtualnych. Zarówno wielkość, jak i zakres zasobów bibliotek cyfrowych są w praktyce bardzo zróżnicowane. Zasięg bibliotek cyfrowych może być lokalny, narodowy, a także międzynarodowy. Założoną grupą odbiorców mogą być bardziej lub mniej liczne środowiska o dobrze zdefiniowanych, specyficznych potrzebach informacyjnych, albo też szeroka i niejednorodna publiczność. Bez względu na rodzaj biblioteki cyfrowej, sposób gromadzenia i udostępniania jej zasobów informacyjnych, ich zakres tematyczny i przeznaczenie, w każdym przypadku podstawą jej udanej implementacji jest skuteczne zarządzanie tymi zasobami. Warunkiem *sine qua non* tej skuteczności jest profesjonalna organizacja zasobów biblioteki za pomocą jednego lub wielu systemów organizacji wiedzy.

Częstym zjawiskiem jest przenoszenie do bibliotek cyfrowych metod organizacji tradycyjnych kolekcji bibliotecznych, takich jak powszechnie znane klasyfikacje i systemy haseł przedmiotowych. Środowisko sieciowe otwiera przed nimi nowe możliwości zastosowań, ale też stawia nowe warunki. Tradycyjne SOW były opracowywane jako narzędzia pracy bibliotekarzy i specjalistów informacji, a umiejętność posługiwania się nimi należała do ich podstawowych kompetencji zawodowych. Udostępnienie zasobów biblioteki cyfrowej do samodzielnego przeszukiwania przez użytkowników wymaga, aby zastosowane w niej systemy organizacji wiedzy ułatwiały to przeszukiwanie bez konieczności uprzedniego opanowania złożonych umiejętności operowania samym SOW. Stąd też systemy tradycyjne stosowane jako narzędzia organizacji informacji cyfrowej są modyfikowane lub rozbudowywane, tak aby spełniały ten istotny warunek.

Tradycyjne SOW przenikają do środowiska sieciowego także za sprawą migracji do niego serwisów bibliograficznych. Zjawisko to odnosi się przede wszystkim do takich SOW jak tezaury i specjalistyczne klasyfikacje. Podobnie jak w przypadku systemów wykorzystywanych w bibliotekarstwie, systemy organizacji wiedzy przez lata rozwijane dla potrzeb serwisów bibliograficznych cechuje duży stopień złożoności. W latach

sześćdziesiątych serwisy bibliograficzne zaczęto udostępniać w formie elektronicznej za pośrednictwem sprzedawców usług informacyjnych takich jak DIALOG lub BRS. Później w dystrybucji ich produktów zastosowano również CD-ROM-y, a obecnie coraz więcej producentów informacji bibliograficznej udostępnia ją w Internecie. W wielu przypadkach zapewnione jest również przeglądanie i przeszukiwanie w sieci elektronicznej wersji specjalistycznych systemów organizacji wiedzy wykorzystywanych w tych serwisach. Coraz lepsze wyszukiwarki internetowe, umożliwiające realizację coraz bardziej złożonych pytań oraz rosnąca niezależność użytkowników przeszukujących zasoby serwisów bibliograficznych, prowadzą do modyfikowania stosowanych w nich tradycyjnych SOW. Większość tych narzędzi zachowała jednak swoją wartość, a w sieciowych zastosowaniach wyposażana jest przede wszystkim w dodatkowe możliwości prezentowania zawartego w nich słownictwa wyszukiwawczego i objaśniania zasad jego użycia.

Bibliograficzne serwisy internetowe prowadzą również wydawcy książek i czasopism naukowych, organizując je często za pomocą specjalnie tworzonych systemów tematycznego porządkowania i przeszukiwania. Wiele takich serwisów, poza informacją bibliograficzną i streszczeniem, udostępnia także pełne teksty artykułów naukowych. Dynamiczny wzrost zasobów czasopism online wymusza dziś odchodzenie od – charakterystycznego do niedawna dla tych serwisów – wyłącznego stosowania technik prostego przeglądania spisów treści i stosowanie również zaawansowanych narzędzi wyszukiwania pełnotekstowego oraz systemów tematycznej organizacji wiedzy. Dobrą ilustracją tego procesu jest rozwój serwisu wydawnictwa Elsevier (http://www.elsevier.com/wps/find/homepage.cws_home), zapewniającego dostęp do informacji bibliograficznej i abstraktów kilkudziesięciu tysięcy publikacji naukowych oraz do pełnych tekstów artykułów z kilkuset czasopism.

Środowisko sieciowe integruje działalność informacyjną prowadzoną w różnych dziedzinach, w których od dawna funkcjonują specyficzne systemy organizacji wiedzy. Należą do nich wykazy znormalizowanego nazewnictwa (np. Polska Scalona Nomenklatura Towarowa Handlu Zagranicznego, Urzędowy Wykaz Leków, wykazy nazw geograficznych polskiej Komisji Standaryzacji Nazw Geograficznych, rozmaite słowniki terminologii specjalistycznej) oraz specjalistyczne schematy klasyfikacyjne, np. standaryzowane w skali krajowej i międzynarodowej klasyfikacje gospodarcze: (International Standard Industrial Classification of Economic Activities (ISICEA), Standard Occupational Classification (SOC), Polska Klasyfikacja Wyrobów i Usług (PKWU), Polska Klasyfikacja Działalności (PKD), Europejska Klasyfikacja Działalności (EKD,) Standard Industrial Classification (SIC), North American Industrial Classification (NAIC) itp.) lub klasyfikacje chorób stosowane w

wymianie informacji medycznej (Międzynarodowa Klasyfikacja Chorób i Problemów Zdrowotnych ICD-10, Klasyfikacja Procedur Medycznych ICD-9). W miarę rozwoju internetowych serwisów informacyjnych różnego typu organizacji do porządkowania sieciowych zasobów informacyjnych wykorzystuje się coraz więcej tego rodzaju SOW, powstaje też wiele systemów nowych.

Środowiskiem, w którym stosowana jest prawdopodobnie największa różnorodność systemów organizacji wiedzy są korporacyjne serwisy intranetowe i ekstranetowe. Projektanci systemów zarządzania informacją i wiedzą w korporacjach oraz systemów komunikacji z ich środowiskiem biznesowym dostosowują do lokalnych potrzeb lub tworzą setki specjalistycznych klasyfikacji, słowników terminologicznych i schematów kategoryzacyjnych. Za specyfikę klasyfikacji i kategoryzacji lokalnie tworzonych dla takich systemów uznać można powiązanie ich struktury ze strukturą organizacyjną firmy, albo z procesami charakterystycznymi dla jej działalności lub też z głównymi kategoriami gromadzonej informacji (*raporty, projekty, propozycje, pracownicy, klienci* itp). Innymi popularnymi metodami organizacji wiedzy w korporacyjnych intranetach są tzw. *think-maps* oraz infomapy, wykorzystujące model relacyjnych baz danych i umożliwiające grupowanie informacji w klasach oraz wiązanie relacjami zarówno tych klas jak i ich elementów (Zaliwski, 2000, s. 57-77). Na potrzeby intranetów i ekstranetów wielkich korporacji implementowane są też najbardziej złożone SOW – rozbudowane taksonomie i ontologie formalne wyrosłe z koncepcji Semantycznego Webu. Przykładem takiej rozbudowanej ontologii jest system Ariadne Genomics, wykorzystywany do formalizacji danych o komórkach proteinowych. Ontologia ta definiuje rodzaje białek, klasyfikuje je oraz kodyfikuje relacje zachodzące między różnymi ich kategoriami, składnikami, własnościami, etc. (Ariadne Genomics, 2005).

Obszarem kształtowania koncepcji nowych, zaawansowanych systemów organizacji wiedzy przeznaczonych do porządkowania, identyfikowania i selekcionowania sieciowych zasobów informacyjnych jest właśnie wspomniany projekt Semantycznego Webu. Bogate archiwum tych działań zawiera witryna prowadzona przez W3C (World Wide Web Consortium; (<http://www.w3.org/2001/sw/>)).

Według pierwotnego pomysłu Tima Bernersa-Lee, system WWW miał być globalną siecią komputerową zawierającą dane w formie, którą mogą automatycznie interpretować komputery, albo też w formie, z której można automatycznie przeprowadzić konwersję na zapis komputerowo interpretowalny. Przez interpretację rozumie się tu automatyczną identyfikację zawartości informacyjnej danych, czyli znaczenia (treści) zapisów

przechowywanych w formie cyfrowej. Zdaniem sceptyków, jeśli w ogóle realizacja takiej wizji globalnej sieci jest możliwa, to należy do odległej przyszłości. Optymiści, obserwując postęp w projektowaniu i implementacji kolejnych narzędzi Semantycznego Webu twierdzą, że powinien być on postrzegany jako logiczna konsekwencja obecnego Webu, która już dziś stopniowo się urzeczywistnia.

W oryginalnej koncepcji T. Bernersa-Lee można wyróżnić dwie części realizacji globalnej sieci informacyjnej. Pierwszą stanowi przekształcenie jej w medium zapewniające skuteczne współdziałanie. Druga część to przekształcenie jego zawartości w dane zrozumiałe dla maszyn, a zatem przetwarzalne komputerowo. Już wstępny pomysł Bernersa-Lee zakładał więc znacznie więcej niż wyszukiwanie stron HTML z serwerów sieci. Na często przywoływanym schemacie pierwotnej koncepcji Webu, opracowanym w 1989 r. dla CERN, widoczna jest specyfikacja związków między rozproszonymi w sieci informacjami. Są tam takie relacje, jak: „zawiera”, „opisuje”, „napisał”, „odnosi się do”, „na przykład”. Kolekcje dokumentów współpracujących ze sobą organizacje porządkują współwykorzystywane systemy hierarchiczne (zob. Berners-Lee, 1989). W rzeczywistości tego typu relacje między różnymi elementami zasobów sieci nie są obecnie zarejestrowane, a wspólne narzędzia są dopiero stopniowo rozwijane. Coraz szerzej stosowaną technologią, która służy identyfikacji i zapisaniu związków między informacjami jest RDF – Resource Description Framework. Warto jednak odnotować, że już oryginalna wersja Webu obejmowała ideę dodatkowych metadanych tworzonych ponad i poza tym, co wykorzystywane jest we współczesnym Webie. To właśnie te dodatkowe metadane są podstawowym narzędziem organizacji wiedzy w Internecie, które warunkuje przetwarzanie zawartej w nim wiedzy przez komputery.

Idea rozproszonego współwykorzystywania i komputerowej interpretowalności danych wymaga, aby identyfikujące je metadane były niezależne od programów, które je wykorzystują. Warunkiem komputerowej interpretacji danych jest tworzenie „danych inteligentnych”, czy też „samosterowalnych” (ang. *smarter data*). Na kolejnych poziomach złożoności służyć mają temu takie narzędzia, jak: XML, RDF, taksonomie i ontologie formalne (Daconta i in., 2003, s. 3). Dwa pierwsze standaryzują zapis charakterystyk zasobów informacyjnych - niezależnie od dziedziny zastosowań. Taksonomie i ontologie formalne są systemami organizacji wiedzy, zapewniającymi kodyfikację pojęć zapisanych w różnorodnych dokumentach oraz zachodzących między nimi relacji, które stanowią podstawę obliczeń logicznych i automatycznego wnioskowania. W projekcie Semantycznego Webu oczywiście i one budowane są w sposób znormalizowany, co warunkuje komunikację między korzystającymi z nich programami.

3. ZAKRES TERMINU *SYSTEM ORGANIZACJI WIEDZY*

Według przytoczonej we wstępie definicji Networked Knowledge Organization Systems Working Group, termin *system organizacji wiedzy* (SOW) oznacza różnego rodzaju schematy porządkowania informacji ułatwiające zarządzanie wiedzą (Witten et al., eds., 1998). Nazwa ta obejmuje więc zestaw narzędzi bardzo różnych, wywodzących się z różnych tradycji działalności praktycznej i przystosowanych do stosowania w różnych środowiskach technicznych.

Systemy organizacji wiedzy wykorzystywane w określonej sytuacji mogą odwoływać się do tradycji bibliotecznej, jak np. znane klasyfikacje, odpowiadające zakresowi tematycznemu kolekcji, własnościom składających się na nią materiałów i potrzebom założonego typu użytkowników. Liczne przykłady pokazują, że rolę SOW organizujących zasoby bibliotek cyfrowych skutecznie pełnią takie systemy klasyfikacyjne, jak: Klasyfikacja Dziesiąta Deweya, Uniwersalna Klasyfikacja Dziesiąta i Klasyfikacja Biblioteki Kongresu, a także: SAB Classification (Klassifikationssystem for Svenska bibliotek), duńska MOS Class, ACM Computing Classification System, Chinese Library Classification, Nederlandse Basisclassificatie, Mathematical Subject Classification, INSPEC Thesaurus, AVEL Thesaurus (por. Sosińska-Kalata, 2002, s. 221-225; Zeng, 2003). Bywa, że projektanci bibliotek cyfrowych sięgają do metod organizacji działalności komercyjnej, takich jak schematy kategoryzacyjne portali Yahoo i Excite lub internetowej księgarni Amazon. Systemy organizacji wiedzy są też tworzone lokalnie, odpowiednio do specyficznych, zwykle wąskich, specjalistycznych potrzeb określonej społeczności użytkowników. Jeśli organizowane zasoby informacyjne mają służyć szerokiej publiczności, najkorzystniejsze jest stosowanie takich SOW, które albo są tej publiczności znane z innych popularnych zastosowań, albo z których korzystanie nie wymaga od użytkownika specjalnej wiedzy i dodatkowego wysiłku. Dostosowane do specyficznych potrzeb lokalne systemy organizacji wiedzy dominują wśród narzędzi stosowanych do organizowania cyfrowych zasobów intranetów wspierających zarządzanie wiedzą w rozmaitych instytucjach i organizacjach.

System organizacji wiedzy może być wykorzystywany do tworzenia rekordów metadanych dla każdego elementu zasobu informacyjnego i wyrażany za pomocą metaznaczników, przechowywanych razem z obiektami przez nie opisanymi. SOW może też być przechowywany niezależnie od cyfrowych zasobów informacyjnych, stanowiąc część mechanizmu dostępu do nich. Bez względu jednak na sposób wykorzystywania, pochodzenie

i rodzaj, każdy system organizacji wiedzy ma zawsze to samo zadanie: organizować treść (zawartość) zasobów informacyjnych i dzięki temu umożliwiać identyfikację i selekcję informacji relewantnych do pytań i potrzeb wyszukujących.

W bibliotekach, archiwach i muzeach cyfrowych oraz w korporacyjnych bazach wiedzy systemy organizacji wiedzy używane są do organizowania materiałów w celu ich wyszukiwania i zarządzania ich kolekcją. SOW służy jako pomost między potrzebą informacyjną użytkownika i materiałem w kolekcji. Za jego pomocą użytkownik może zidentyfikować interesujący go obiekt wcześniej nie mając o nim wystarczających danych. Czy to poprzez przeglądanie, czy bezpośrednio przeszukiwanie, czy poprzez tematy na stronach internetowych, czy za pomocą wyszukiwarki, SOW kieruje użytkownikiem w jego procesie odkrywania wiedzy. SOW pozwala też organizować informację stanowiącą wynik poszukiwań, odpowiednio do jej zakresu i przedmiotu.

Decyzja o wyborze systemu organizacji wiedzy ma zasadnicze znaczenie dla realizacji koncepcji i rozwoju każdej biblioteki cyfrowej i każdego systemu zarządzania wiedzą. SOW musi być odpowiedni do zadań, jakie stawia się przed nimi i do charakteru zgromadzonej w nich informacji. Musi być też odpowiedni do założonego sposobu opracowania zasobów za pomocą metod automatycznego indeksowania albo za pomocą indeksowania intelektualnego, czyli prowadzonego przez klasyfikatorów i indeksatorów. Kwestią ważną jest również to, że SOW, którym w poszukiwaniach informacji mają posługiwać się ludzie, musi być dla nich zrozumiały i dostosowany do ich wiedzy i umiejętności wyszukiwania informacji. Ma im pomagać w przeglądaniu jej zasobów i ich przeszukiwaniu, a nie tworzyć bariery utrudniające to zadanie.

Jest oczywiste, że nie istnieje system organizacji wiedzy, który akceptowaliby wszyscy. Choć opracowanie jednego, powszechnie przyjętego systemu, byłoby może ze względów praktycznych korzystne, to jest mało prawdopodobne, aby kiedykolwiek taki system powstał. SOW, jak wszystkie wytwory ludzkiego intelektu, są uwarunkowane kulturą, z której i dla potrzeb której wyrosły. Co jest istotne dla jednej kultury niekoniecznie musi być istotne dla innej. Stąd żyjemy w świecie wielości i różnorodności, i dotyczy to również sposobów organizowania wiedzy. Mimo tej różnorodności dla wszystkich systemów organizacji wiedzy wspólne jest to, że:

- każdy jest emanacją pewnego określonego sposobu patrzenia na rzeczywistość, który implikuje sposób patrzenia na organizowany zasób informacji i jego elementy;
- każdy jest uporządkowanym zbiorem wyrażeń reprezentujących pojęcia, które składają się na tę projekcję rzeczywistości;

– za pomocą różnych SOW ten sam obiekt (tekst, obraz, zapis dźwiękowy lub filmowy) może być w różny sposób scharakteryzowany, a zatem różne SOW mogą się wzajemnie uzupełniać;

– swobodne i poprawne posługiwanie się każdym SOW przez człowieka posiadającego pewną wiedzę w dziedzinie poszukiwań wymaga zachowania wystarczającej jedności między pojęciem wyrażanym za pomocą SOW i obiektem, do którego pojęcie to się odnosi; osoba szukająca określonych materiałów używając SOW musi być w stanie powiązać zawartą w nim projekcję pojęć z własnym wyobrażeniem świata.

Zakres terminu można określić przez wyliczenie podporządkowanych mu terminów węższych. Innymi słowy, rodzaj definiować można przez wyliczenie należących do niego gatunków. Zabieg taki w odniesieniu do terminu *systemy organizacji wiedzy* wymaga jednak dodatkowych wyjaśnień. Na wstępie zasygnalizowany został problem niejednorodnego stosowania nazw różnych typów systemów organizacji wiedzy przez różne środowiska zajmujące się tą problematyką. Jest to problem dotyczący zarówno terminologii polskiej, jak i w innych językach. Nazwy różnych rodzajów SOW mają wprowadzić swoje definicje, niekiedy wprowadzane nawet za pomocą szczegółowych norm, nie są to jednak definicje ogólnie przyjęte, a autorzy wielu z nich zdają się po prostu nie wiedzieć, że wprowadzany przez nich termin ma już znaczenie wcześniej przypisane w innych dyscyplinach. Odrębną kwestią jest to, że w wyniku coraz szerszego i coraz bardziej różnorodnego zastosowania przede wszystkim w środowisku cyfrowym, obiekty desygnowane przez te terminy podlegają dużym zmianom.

Reakcją na narastający chaos terminologiczny w nazewnictwie współczesnych SOW było zorganizowanie w 1999 r. specjalnego seminarium amerykańskiego National Information Standards Organization (NISO), na którym rozważano zakres tego problemu i możliwości jego rozwiązania (NISO, 1999). W lutym 2003 r. NISO przeprowadziła analizę stosowania w praktyce najważniejszej dla konstruktorów SOW normy Z39.19-1993 (R.1998) *Guidelines for the Construction, Format, and Management of Monolingual Thesauri*. Jak wynika z tytułu, odnosi się ona do budowy tezaurusów i to w tradycyjnym znaczeniu, przyjętym w działalności dokumentacyjnej. Nie istnieją jednak krajowe ani międzynarodowe standardy budowy SOW innego typu.

Na podstawie ankiety stwierdzono, że norma Z39.19-1993 w wersji zrewidowanej w 1998 r. w dużym stopniu nie spełnia oczekiwań twórców systemów organizacji wiedzy w serwisach elektronicznych. Respondenci zwracali uwagę na konieczność przystosowania normy do potrzeb tworzenia tezaurusów online i taksonomii, wyznaczających systematykę

terminów gromadzonych w tezaurusach. Podkreślali też potrzebę uwzględnienia w tej normie takich standardów, jak: Dublin Core (Z39.85), normy wielojęzyczne, brytyjska i międzynarodowa norma budowy tezaurusów jedno- i wielojęzycznych (BSI 5723, ISO 2788, ISO 5964), standardy dotyczące XFML i XML, a także ustalenia przyjmowane w projekcie Semantycznego Weba (ANSI/NISO, 2003).

Wiele z postulatów zgłoszonych na seminarium w 1999 r. oraz zgromadzonych w badaniu z lutego 2003 r. wykorzystano przygotowując rewizję normy Z39.19-2003. Nadal jednak odczuwana jest potrzeba pełniejszego uporządkowania terminologii i reguł tworzenia przede wszystkim systemów organizacji wiedzy w środowisku sieciowym. Wychodząc z założenia, że tworzenie i stosowanie narzędzi wyszukiwania i przeglądania informacji nie jest już dziś domeną wyłącznie bibliotekarzy, indeksatorów i innych specjalistów informacji, lecz stało się obszarem działań specjalistów różnych dziedzin i zawodów oraz służy różnym celom szczegółowym, NISO podjęło prace nad tzw. „nową normą tezaurusową”. Kolejna rewizja normy Z39.19-2003 ma dostosować zalecenia budowy SOW do potrzeb twórców internetowych i intranetowych stron WWW, baz danych i serwisów informacyjnych. Nowy standard ma też uregulować zasady tworzenia systemów metadanych ułatwiających wyszukiwanie informacji niespecjalistom (NISO, 2005).

4. TYPY SYSTEMÓW ORGANIZACJI WIEDZY

Wszystkie systemy organizacji wiedzy są uporządkowanymi zbiorami wyrażeń reprezentujących pojęcia, które składają się na pewną semantyczną wiedzę dziedzinową. Zakres tej wiedzy wyznacza zakres tematyczny porządkowanych zasobów informacji. Podstawą rozróżnienia typów SOW, zarówno stosowanych w środowisku tradycyjnych zbiorów dokumentów, jak i w środowisku sieciowych zasobów informacyjnych, może być stopień i sposób uporządkowania tych wyrażeń, czyli przyjęta metoda strukturyzacji wiedzy dziedzinowej. O sposobie uporządkowania wiedzy wyrażonym w strukturze SOW decyduje zakres wykorzystania trzech podstawowych narzędzi: porządku formalnego oraz podziału dyscyplinarnego lub kategoryjnego. O stopniu strukturyzacji wiedzy decyduje natomiast zakres identyfikacji i reprezentacji relacji, zachodzących między pojęciami reprezentowanymi przez wyrażenia zgromadzone w SOW (por. Sosińska-Kalata, 1999, s. 127-173).

W piśmiennictwie dotyczącym porządkowania cyfrowych zasobów informacyjnych spotyka się rozróżnienie dwóch głównych typów systemów organizacji wiedzy: dokładnych schematów organizacyjnych i niejednoznacznych schematów organizacyjnych (Rosenfeld i Morville, 2003, s. 77). Generalnie odpowiada ono znanemu z teorii i metodologii

opracowania rzeczowego podziałowi na systemy stanowiące podstawę alfabetycznych układów klasowych i przedmiotowych oraz logicznych układów działowych i systematycznych.

Schematy dokładne operują formalnymi metodami porządkowania wyrażeń reprezentujących wyodrębnione, zwykle szczegółowe, tematy lub pojęcia. Każde z tych wyrażeń wyznacza klasę z założenia jednoznacznie identyfikującą określone obiekty lub określoną treść, a zakres każdej takiej klasy jest rozłączny wobec zakresu klas pozostałych. Najczęściej podstawą porządkowania jest układ alfabetyczny, rzadziej – chronologiczny lub geograficzny. Tego typu SOW zapewniają szybkie i łatwe odnajdywanie informacji za pomocą znanych użytkownikowi nazw interesujących go obiektów, np.: nazwisk pracowników pewnej organizacji, nazw towarów lub tytułów dokumentów zgromadzonych w bibliotece cyfrowej. Podobnie jak biblioteczne katalogi przedmiotowe, dokładne schematy organizacyjne nie umożliwiają jednak całościowego wglądu w zawartość zasobów informacyjnych zorganizowanych za ich pomocą. Wyszczególniając informację o obiektach jednostkowych nie zapewniają jej generalizacji.

Do niejednoznacznych schematów organizacyjnych zalicza się SOW, które dzielą informacje na grupy tematyczne (klasy, kategorie) uporządkowane na podstawie kryteriów semantycznych (logicznych). Należą do nich różne odmiany struktur hierarchicznych, w których klasy reprezentowane są przez wyrażenia naturalne. Większość tych wyrażeń cechuje niejednoznaczność, toteż sposób hierarchizacji wyodrębnionych klas zwykle odwołuje się do jednej, arbitralnie przyjętej interpretacji reprezentującego ją wyrażenia, albo też – częściej – dąży do objęcia systematyzacją wielu lub wszystkich jego interpretacji. Używając słów jako etykiet kategorii tematycznych ryzykuje się, że różni użytkownicy mogą je różnie rozmieć. Odrębną kwestią jest wybór kryteriów porządkowania, implikujący określoną interpretację porządkowanych wyrażeń, która nie zawsze konweniuje z interpretacją znaną lub preferowaną przez użytkownika. W rezultacie w oczach użytkowników niejednoznaczny może stać się cały układ hierarchiczny. Stąd też wśród tego rodzaju SOW dominuje dążenie do wieloaspektowej organizacji polihierarchicznej. Niejednoznaczne schematy organizacyjne mają zapewnić wgląd w całą zawartość organizowanych zasobów informacyjnych i poprzez jej przeglądanie umożliwić odnalezienie informacji, której użytkownik nie potrafi wcześniej precyzyjnie zdefiniować.

Podział SOW według zakresu odwzorowania relacji zachodzących między kodyfikowanymi w nich pojęciami zdaje się być najbliższy powszechnej intuicji. W wielu bowiem opracowaniach, w których omawia się możliwe do zastosowania metody

organizowania wiedzy w bibliotekach cyfrowych albo korporacyjnych systemach zarządzania wiedzą, najczęściej wskazuje się trzy grupy narzędzi: wykazy słownictwa, struktury hierarchiczne oraz tezaury (por. Hodge, 2000; Rosenfeld i Morville, 2003).

Generalny podział struktur relacyjnych określających modele organizacji wiedzy za podstawę przyjmuje kryterium liczby relacji wykorzystanych w projekcji organizacji wiedzy oraz kryterium sposobu wykorzystania związków hierarchicznych w projekcji organizacji wiedzy. Ze względu na pierwsze kryterium wyodrębnia się modele monorelacyjne i polirelacyjne; ze względu na drugie – modele ahierarchiczne, monohierarchiczne i polihierarchiczne. Dokładniejszą analizę sieciowych SOW umożliwi wprowadzenie dodatkowego kryterium stopnia wykorzystania niehierarchicznych związków semantycznych.

Najprostszy model organizacyjny reprezentują SOW w postaci tzw. list terminów, w których relacyjna struktura pojęć wyrażanych przez zgromadzone słownictwo w ogóle nie jest *explicite* prezentowana lub zredukowana jest do projekcji grup synonimicznych, czyli związku równoważnościowego. W niektórych z nich związki znaczeniowe między poszczególnymi wyrażeniami można ustalić na podstawie zawartości definicji objaśniających, które towarzyszą wyrażeniom, nie umożliwiają one jednak bezpośredniej projekcji systemu pojęciowego dziedziny czy sortowania słownictwa według kryteriów semantycznych. Ponieważ wszystkie wyrażenia zawarte na listach terminów z punktu widzenia projekcji organizacji wiedzy uznać można za współrzędne, listy terminów realizują model monorelacyjny i ahierarchiczny.

Modelem bardziej złożonym są struktury hierarchiczne, w których terminy lub symbole i reprezentowane przez nie pojęcia łączone są w wielopoziomowe gniazda tematyczne. Głównym związkiem pojęciowym, który konstytuuje projekcję organizacji wiedzy jest tu związek hierarchiczny. Jeśli wszystkie pojęcia uporządkowane są w jednym drzewiastym układzie hierarchicznym, projekcja struktury wiedzy przyjmuje postać klasyfikacji monohierarchicznej. Jeśli uwzględnia się więcej niż jeden układ hierarchiczny, to projekcja struktury wiedzy ma formę polihierarchii. Warto zauważyć, że pojęcie polihierarchii nie jest jednolicie interpretowane. W rozumieniu klasycznym, związanym z teorią klasyfikacji piśmiennictwa, polihierarchia oznacza organizację elementów uniwersum w więcej niż jednym drzewie hierarchicznym, co jednak nie musi implikować wielokrotnego przyporządkowywania pojęć szczegółowych różnym pojęciom nadrzędnym, czyli tzw. polipozycyjności. W odniesieniu do sieciowych SOW, zwłaszcza systemów kategoryzacyjnych i taksonomii porządkujących stosunkowo złożone grupy tematyczne, najczęściej polihierarchią nazywa się właśnie to wielokrotne przyporządkowanie

poszczególnych grup węższych zagadnieniom o szerszym zakresie, stanowiące konsekwencję akceptowania „krzyżowych” zależności hierarchicznych.

Najbardziej złożony model realizują takie systemy organizacji wiedzy, w których wyrażenia reprezentujące określone pojęcia powiązane są różnymi typami związków, nie tylko hierarchicznymi, ale także mniej lub bardziej bogatym zestawem innych relacji asocjacyjnych. Do grupy tej należą więc takie zbiory wyrażeń, które umożliwiają wieloaspektową nawigację po systemie terminologicznym poszczególnych dziedzin. Jak w pozostałych przypadkach, klasa tego rodzaju SOW jest wewnątrznie zróżnicowana. W niektórych oznacza się różne rodzaje relacji, ale nie są one dokładnie specyfikowane. Systemy najbardziej złożone dążą do identyfikacji i dokładnej specyfikacji możliwie wszystkich rodzajów pojęciowych powiązań, zapewniając każdemu reprezentowanemu pojęciu bogaty kontekst relacyjny definiujący sposób jego interpretacji.

4.1. LISTY TERMINÓW

Najczęściej stosowanymi odmianami SOW, które określa się listami lub zbiorami słownictwa kontrolowanego, są kartoteki haseł wzorcowych (autorytarnych), tzw. glosariusze, słowniki i indeksy nazw geograficznych. Charakterystyczną cechą tych systemów organizacji wiedzy jest porządek formalny zgromadzonego słownictwa.

Najogólniej mówiąc, k a r t o t e k i h a s e ł w z o r c o w y c h są listami terminów używanymi do kontroli nazw odnoszących się do tych samych desygnatów. Przykładami są nazwy państw, osób czy organizacji. Terminy odrzucone mogą być powiązane z terminami preferowanymi. W tego rodzaju SOW cechuje płytka organizacja i zwykle prosta składnia. Uporządkowane są najczęściej alfabetycznie, rzadziej za pomocą płytkiego schematu klasyfikacyjnego. Niekiedy stosuje się tzw. ograniczoną hierarchizację, która umożliwia prostą nawigację, szczególnie wtedy, gdy kartoteka przeglądana jest manualnie albo jest bardzo duża. Najbardziej znanymi przykładami takich SOW są: Library of Congress Name Authority File, czy też Centralna Kartoteka Haseł Wzorcowych wykorzystywana przez polski katalog centralny NUKAT.

G l o s a r i u s z a m i lub słownikami specjalistycznymi nazywa się listę terminów należących do wąskiej specjalności, zwykle z ich definicjami. Terminy mogą pochodzić z pewnej dziedziny, albo z określonego dzieła czy zbioru dzieł określonego autora. Przykładami glosariuszy pierwszego rodzaju są Glosariusz Internetowego Serwisu Głównego Negocjatora opracowany w ramach programu „Zrozumieć negocjacje” (ISGN, 2003) albo Słownik serwisu

SearchEngines.pl, będący listą terminów dotyczących wyszukiwarek internetowych (SearchEngines, b.d.). Glosariuszem drugiego rodzaju jest np. słownik terminów Arystotelesowskich. Terminy zdefiniowane są tu przez określony kontekst dziedzinowy i w przypadku polisemów na ogół nie zawierają znaczeń spoza tego kontekstu.

Nazwę *s ł o w n i k* odnosi się najczęściej do alfabetycznych wykazów wyrazów i fraz wraz z ich definicjami. W przypadku wyrażen polisemicznych w słownikach podaje się definicje wszystkich ich znaczeń. Zwykle zakres tego typu SOW jest ogólniejszy niż zakres glosariuszy. Często zapewniają one także informacje o etymologii wyrażen oraz ich wariantach ortograficznych i morfologicznych. W słownikach często wskazuje się też synonimy, a poprzez definicję objaśniającą znaczenie, również wyrażenia o znaczeniu pokrewnym. Słowniki nie umożliwiają jednak bezpośredniej prezentacji struktury hierarchicznej systemu leksykalnego ani pojęciowego grupowania wyrażen. Dobrym przykładem takich narzędzi są serwisy Miriam–Webster Online (Miriam-Webster Online, 2005) oraz Słownik języka polskiego PWN (PWN, 2004).

I n d e k s y g e o g r a f i c z n e są wykazami nazw jednostek geograficznych, np. miejscowości, rzek, jezior etc. Tradycyjne indeksy tego typu publikuje się w formie książek lub jako indeksy w atlasach geograficznych. W wykazach takich nazwy uzupełniane bywają informacją o typie jednostki geograficznej, do której należy desygnowany przez nie obiekt, np. rzeka lub miasto. Przykładem tego rodzaju listy jest GNIS US Geographic Names System (USGS, 2004). Znacznie rozszerzoną wersją indeksu geograficznego są wykazy nazw geograficznych polskiej Komisji Standaryzacji Nazw Geograficznych (KSNG), które poza nazwami państw świata i ich stolic obejmują też nazwy ich mieszkańców (KSNG, 2003).

4.2. KLASYFIKACJE, KATEGORYZACJE I TAKSONOMIE

Drugą grupę systemów organizacji charakteryzuje cecha bezpośredniego oznaczania związków hierarchicznych między pojęciami reprezentowanymi przez terminy lub symbole klasyfikacyjne. Związkom tym podporządkowana jest projekcja struktury wiedzy i zapewniają one systematyczne przeglądanie zawartości zasobów informacyjnych. Na oznaczenie tego typu SOW w piśmiennictwie używa się nazw *klasyfikacja*, *kategoryzacja* i *taksonomia*, często operując nimi zamiennie. Choć z teoretycznego punktu widzenia można w sposób jednoznaczny rozróżnić zakres tych trzech nazw (por. Sosińska-Kalata, 2002, s. 21-26), to w praktyce różnice te nie są ostro identyfikowane. Za najważniejszą i konstytutywną własność wszystkich SOW zaliczanych do tej grupy uważa się organizowanie pojęć w

szerszych grupach tematycznych, które poddawane są najczęściej wielostopniowym podziałom, segmentującym zasoby informacyjne na coraz węższe klasy.

Niektóre spośród tych systemów mają postać hierarchicznego uporządkowania pojęć lub grup tematycznych reprezentowanych przez symbole numeryczne, alfabetyczne lub alfanumeryczne. Struktura tych symboli wskazuje zależności hierarchiczne między reprezentowanymi przez nie klasami. W takich przypadkach najczęściej mówi się o *k l a s y f i k a c j a c h*, jakkolwiek wiele tego typu SOW nie spełnia i dla sprawnego funkcjonowania nie musi spełniać wymagań klasyfikacji w sensie logicznym. Schematy klasyfikacyjne mają w praktyce bardzo zróżnicowaną strukturę, od prostych układów monohierarchicznych do złożonych i wieloaspektowych polihierarchii, wykorzystujących równolegle podział pojęć (tematów) na kategorie i klasy dyscyplinarne. Charakterystyczną cechą tego typu SOW jest tzw. podział aspektowy – pojęcia szczegółowe podporządkowane są klasom nadrzędnym wyodrębnianym na podstawie pewnego aspektu generalizacji, a zatem klasy najogólniejsze identyfikują sposób ujęcia zagadnień szczegółowych. Każdy taki sposób ujęcia może prowadzić do utworzenia odrębnej hierarchii. Tradycyjnie najczęściej stosowanym aspektem porządkowania jest kwalifikacja dziedzinowa lub zagadnieniowa. Aspektem organizacji pojęć szczegółowych może być też ich przynależność kategorialna. Najbardziej reprezentatywnymi przykładami SOW tego typu są dobrze znane w bibliotekarstwie, ale też szeroko stosowane w środowisku sieciowym, tradycyjne klasyfikacje biblioteczno-bibliograficzne i dokumentacyjne, takie jak Klasyfikacja Biblioteki Kongresu, Klasyfikacja Dziesiąta Deweya i Uniwersalna Klasyfikacja Dziesiąta, tematologiczna Broad System of Ordering (BSO), a także specjalistyczne klasyfikacje jak Mathematical Subject Classification lub ICONCLASS. We wszystkich SOW tego rodzaju relacje hierarchiczne tworzą szkielet strukturyzacji wiedzy, nie jest w nich jednak identyfikowany typ hierarchicznych powiązań, nie specyfikuje się też innych związków semantycznych między pojęciami reprezentowanymi przez symbole. Coraz częściej możliwość nawigacji po takich systemach zapewniana jest za pomocą wyrażen naturalnych, z pominięciem sztucznych symboli, operowanie którymi użytkownikom nieprzygotowanym zwykle sprawia trudność. Doskonałym przykładem takiego właśnie wykorzystania tradycyjnej Klasyfikacji Dziesiątej Deweya jest brytyjski serwis BUBL LINK (<http://bubl.ac.uk/link/index.html>).

Do systemów hierarchicznych zaliczymy też *p o d z i a ł z a p o m o c ą h a s e ł p r z e d m i o t o w y c h*, które pozwalają wyodrębnić zwykle dość szczegółowe kategorie tematyczne, szeregując je następnie na podstawie porządku alfabetycznego. Zaznaczyć trzeba, że struktura takiego podziału jest na ogół znacznie płytsza niż struktura klasyfikacji

aspektowej. Wykazy haseł przedmiotowych nie zapewniają projekcji hierarchicznej organizacji pojęć reprezentowanych przez hasła (tematy i określniki), jakkolwiek informacje o związkach hierarchicznych zamieszczane są w nich w artykułach przedmiotowych. Same takie wykazy należałoby zatem zaliczyć do SOW pierwszego typu, czyli list terminów kontrolowanych. Jednak podział zasobów informacyjnych (dokumentów) za pomocą haseł przedmiotowych prowadzi do utworzenia struktury hierarchicznej, w której klasy najogólniejsze reprezentowane są przez hasła najprostsze, identyfikujące przedmiot informacji. Podporządkowane im klasy szczegółowe wyodrębniane są za pomocą określników wskazujących specyficzne aspekty rozważania danego przedmiotu. Określniki dołączane do haseł generują więc hierarchiczny układ zagadnień szczegółowych związanych z przedmiotem rozważań reprezentowanym przez temat. Warto przypomnieć, że w tradycyjnej terminologii bibliotekarskiej tego typu systemy organizacji wiedzy w odróżnieniu od klasyfikacji aspektowych nazywane były klasyfikacjami przedmiotowymi, klasy najogólniejsze wyznaczone są tu bowiem przez identyfikację przedmiotu porządkowanej informacji, klasy szczegółowe natomiast na podstawie aspektów jego rozważania. Dobrze znanymi przykładami SOW tego rodzaju są Library of Congress Subject Headings (LCSH) czy kartoteka wzorcowych haseł przedmiotowych Biblioteki Narodowej (SJHP BN).

Od pewnego czasu dużą popularnością cieszy się określenie *k a t e g o r y z a c j a*, które zwykle stosuje się w odniesieniu do hierarchicznych układów nazw grup tematycznych. Pod tym względem kategoryzacja podobna jest więc do podziału za pomocą haseł przedmiotowych, różni je jednak to, że kategorie podrzędne z reguły nie są w niej wyodrębniane za pomocą dodatkowych atrybutów lub aspektów, których nazwy dołączane byłyby do nazw kategorii nadrzędnych. Wyznaczają je najczęściej pojęcia bardziej szczegółowe, pozostające wobec kategorii nadrzędnej w pewnym związku hierarchicznym, np. *Aktualności* > *Media* > *Prasa* > *Prasa tematyczna* > *Prasa kobieca*. Kategorie nadrzędne mogą mieć charakter zarówno przedmiotowy jak i aspektowy (ujęciowy). De facto więc kategoryzacje stanowią pewnego typu klasyfikacje, zwykle tworzone indukcyjnie, w których klasy reprezentowane są przez nazwy naturalne.

Systemy kategoryzacyjne wykorzystywane są do zestawiania terminów w klasy o pewnej wspólnej cesze ogólnej. Sam termin *kategoria*, w sensie wywodzącym się od Arystotelesa, oznacza „orzecznik i orzeczenie” oraz „sposób orzekania” o dowolnym przedmiocie [por. Arystoteles, 2003, s. 34; Narecki, 2003, s.63]. Orzekanie to może przyjmować różne formy, bo o każdej rzeczy można powiedzieć, że jest czymś i że jest jakaś. Wyróżniony przez Stagirytę zestaw dziesięciu kategorii ogólnych, naczelnej kategorii

substancja oraz dziewięciu kategorii tzw. ogólnych przypadłości (przysługujące substancji: *jakość, ilość, relacja, miejsce, czas, położenie, stan, czynność, doznawanie*), stanowi najbardziej znany wzorzec podziału kategorialnego tego, co o bycie można orzec, innymi słowy – wszelkich pojęć, którymi człowiek posługuje się konceptualizując świat.

Kategoryzację uważa się za fundamentalny składnik działalności językowej. Pojęcie *kategoria językowa* jest rozumiane bardzo szeroko i w różnych znaczeniach. Najogólniej oznacza zbiór elementów pełniących tę samą funkcję w systemie językowym, a więc znaczy tyle, co *klasa*. W językoznawstwie kategoryzacja jest wykorzystywana na dwóch poziomach. Na poziomie pierwszym jest dla językoznawcy, jak i dla badacza każdej innej dyscypliny, instrumentem porządkującego opisu przedmiotu jego badań. Kategoryzowaniu podlegają więc język i jego elementy. Poziom drugi to analiza tego, jak za pomocą wyrażeń językowych kategoryzowane są elementy świata w nim opisywanego. Wyrazy, morfemy, struktury składniowe nie tylko same są kategoriami, ale także reprezentują kategorie fizycznie i zmysłowo odrębnych właściwości świata rzeczywistego (Taylor, 2001, s.19-20). Na tym drugim poziomie kategoryzacja językowa jest projekcją pojęciowej organizacji wiedzy o świecie. W pewnym sensie tak również można patrzeć na nazywane kategoryzacjami systematyzujące układy hierarchiczne nazw grup tematycznych, stosowane w internetowych serwisach informacyjnych.

Bardziej lub mniej rozbudowane układy kategorii tematycznych należą dziś do najpopularniejszych metod organizacji różnego typu serwisów internetowych. Stosowane są zarówno w wielkich portalach o zakresie nieograniczonym, jak i w specjalistycznych wortalach lub korporacyjnych witrynach. Dobrze znany przykład to wzorcowy system kategorii portalu Yahoo oraz jego swoiste modyfikacje, takie jak kategoryzacja wykorzystywana w portalu Onet i jego katalogu WWW. Inne przykłady to schematy kategoryzacyjne internetowych księgarni, jak Amazon, Merlin lub ekonomiczna.pl, serwisów tematycznych, jak brytyjski ADAM zapewniający uporządkowany dostęp do sieciowych zasobów z zakresu sztuk pięknych i architektury (<http://adam.ac.uk/>). Stopień rozbudowy tych schematów waha się od kilku do kilkunastu poziomów hierarchii, w zależności od wielkości serwisu i liczebności porządkowanych zasobów oraz od kilkuset do kilkudziesięciu tysięcy kategorii tematycznych. Normą jest alfabetyczne porządkowanie grup tematycznych na każdym szczeblu podziału, na ogół też zestaw kategorii najogólniejszych nie przekracza liczby kilkunastu. Charakterystyczna dla nich jest też polihierarchia rozumiana jako polipozycyjność. Te same grupy tematyczne, często wraz z całą dalszą rozbudową, powtarzane są w podziale różnych kategorii nadrzędnych, np. w katalogu WWW portalu Onet

kategoria *prasa medyczna* podporządkowana jest zarówno kategorii *Prasa tematyczna* (a poprzez nią kategoriom *Aktualności* > *Media* > *Prasa* > *Prasa tematyczna*), jak i kategorii *Prasa naukowa* (a poprzez nią kategorii *Nauka*). Podobnie jak w klasyfikacjach, również w strukturze systemów kategoryzacyjnych nie identyfikuje się rodzaju powiązań hierarchicznych, ani innego typu asocjacji.

Do hierarchicznych SOW należą również t a k s o n o m i e, w coraz szerszym zakresie wykorzystywane w projektowaniu obiektowym i systemach zarządzania wiedzą. Podobnie jak w przypadku klasyfikacji i kategoryzacji, ich zadaniem jest grupowanie dowolnego typu obiektów na podstawie określonych atrybutów. W organizacji zasobów serwisów informacyjnych WWW, taksonomia (niekiedy zwana też hierarchią) stanowi podział zawartości serwisu na kategorie ogólne, z których wyodrębnia się następnie kategorie bardziej szczegółowe. W zaleceniach dla projektantów taksonomii zwraca się uwagę na to, że taksonomia powinna być najważniejszym instrumentem generalnej organizacji zasobów serwisu, który powinien zapewnić łatwość zorientowania się w jego zawartości i przeglądania jej. W zależności od stopnia rozbudowania serwisu i rodzaju jego użytkowników, powinna być ona jednak uzupełniona innymi schematami organizacyjnymi, takimi jak schemat powiązań hipertekstowych, zapewniający większą swobodę przeglądania, lub schemat bazodanowy, umożliwiający dokładne przeszukiwanie serwisu według określonych, reprezentowanych za pomocą metadanych, szczegółowych własności (Rosenfeld i Morville, 2003, s.88-91).

Podobnie jak schematy kategoryzacyjne, taksonomie organizujące serwisy WWW mają zwykle strukturę polihierarchiczną. Niejednoznaczność kategorii tematycznych reprezentowanych nazwami naturalnymi skłania do zapewnienia użytkownikom różnych dróg dotarcia do poszukiwanej informacji. Ze względu na to, że taksonomia ma zapewnić łatwy wgląd w całą zawartość serwisu, kwestią istotną jest wyważenie proporcji między wielością powiązań polihierarchicznych i przejrzystością podziału. Nadmiar zależności krzyżowych prowadzi do rozmycia granic między wyodrębnionymi kategoriami tematycznymi, rygorystyczne ich ograniczenie może utrudnić dotarcie do kategorii tematycznej interpretowanej przez użytkownika inaczej niż przez projektanta taksonomii.

Ważnym aspektem projektowania taksonomii jest też wyważenie proporcji między jej szerokością i głębokością. Przez szerokość taksonomii rozumie się liczbę kategorii wyodrębnionych na jednym poziomie hierarchii, czyli to co w teorii klasyfikacji określa się jako długość szeregu klasyfikacyjnego. Głębokość taksonomii to liczba szczebli podziału kategorii, a więc to co w teorii klasyfikacji nazywa się długością łańcucha klasyfikacyjnego.

Taksonomie nadmiernie głębokie wymagają od użytkownika żmudnego przechodzenie przez szereg kolejnych poziomów podziału zanim ostatecznie odnajdzie poszukiwaną kategorię szczegółową. Z kolei taksonomie nadmiernie szerokie prowadzą do wyodrębnianie na jednym poziomie kategorii zbyt szczegółowych, rozpraszając informację i wymagając od użytkownika uważnej analizy i niekiedy przeglądania dużej liczby oferowanych opcji. Taksonomie głębokie są najczęściej strukturami wysmukłymi, czyli są równocześnie wąskie. Taksonomie szerokie na ogół są równocześnie hierarchiami płaskimi. Projektantom tego typu SOW zaleca się, aby szukając optymalnej szerokości i głębokości schematu kierowali się ludzkimi możliwościami poznawczymi i zdolnościami postrzegania. Generalnie przeważa dziś pogląd, że ograniczenia percepcji zawartości stron WWW wynikają z ograniczeń ludzkich możliwości wizualnego skanowania strony, a nie pojemności pamięci krótkotrwałej. Nie uważa się więc za konieczne przestrzeganie zasady „siedem-plus-minus-dwa”, określającej pojemność pamięci krótkiej, akcentując jedynie potrzebę zwracania uwagi na niebezpieczeństwo dezorientowania lub rozpraszania uwagi użytkowników wprowadzaniem zbyt wielu możliwości wyboru. Mimo to na większości wysoko ocenianych stron WWW, w tym uważanej za wzorcową witryny Microsoftu (www.microsoft.com) liczba kategorii (klas tematycznych) wyróżnionych na stronie głównej nie przekracza 9. Zaleca się też grupowanie i strukturyzację informacji na poziomie strony oraz poddawanie przygotowywanego projektu taksonomii rygorystycznym testom użytkowników. Z kolei optymalną głębokość taksonomii ocenia się na podstawie badania zachowań wyszukiwawczych: wykazują one, że w większości przypadków użytkownicy zmuszeni do przechodzenia przez więcej niż 3-4 poziomy podziału, porzucają dany serwis lub czują się co najmniej sfrustrowani.

4.3. LISTY RELACYJNE

Najbardziej złożone, ale też niejednorodne modele organizacyjne, zawiera klasa SOW nazwana ogólnie listami relacyjnymi. Należą do niej przede wszystkim tezaury, sieci semantyczne, tzw. mapy myśli i ontologie informatyczne.

T e z a u r u s y w sensie najogólniejszym są słownikami pojęciowymi, w których znaczenie terminów opisywane jest za pomocą sieci ich semantycznych powiązań z innymi terminami. Klasycznym wzorem tego rodzaju słowników jest Tezaurus Marka Rogeta, do jego koncepcji nawiązują reguły budowy tezaurusów od ponad czterdziestu lat wykorzystywanych w działalności dokumentacyjnej.

Zestaw relacji powszechnie wyrażanych w tezaursach obejmuje związki hierarchiczne (hiperonimię i hiponimię), równoważnościowe (synonimię) oraz kojarzeniowe (asocjacji, pokrewieństwa semantycznego). Zgodnie z zaleceniami norm tezaursowych, w strukturze artykułów słownikowych relacje te oznaczane są wskaźnikami BT (broader term), NT (narrower term), U/UF (use / used for) i RT (related term). Według normy polskiej odpowiednikami tych oznaczeń są: SD (szerszy deskryptor), WD (węższy deskryptor), U/NU (używaj / nie używaj) i KD (kojarzeniowy deskryptor). W coraz częściej spotykanych tezaursach, w których nie wyklucza się możliwości operowania synonimami w wyszukiwaniu informacji, zamiast wskaźników U/UF (U/NU) stosowany jest wskaźnik SY (synonym / synonim). Zarówno relacje hierarchiczne jak i asocjacyjne mogą być specyfikowane bardziej szczegółowo. Np. standard Z39.19 proponuje rozróżnienie związków BTG (broader term generic), BTI (broader term instance), BTP (broader term partitive), NTG (narrower term generic), NTI (narrower term instance), (NTP = narrower term partitive); analogiczne rozróżnienia przewiduje też polska norma PN-92/N-09018. W wielu nowoczesnych tezaursach wprowadza się również szczegółową specyfikację powiązań kojarzeniowych. Dobrym przykładem jest tezaurus Unified Medical Language System (UMLS) opracowany w amerykańskiej National Library of Medicine. Zdefiniowano w nim ponad 40 relacji, z których większość identyfikuje skojarzenia zachodzące między pojęciami współwystępującymi w różnego typu kontekstach.

Wiele z tezaursów wykorzystywanych w środowisku sieciowym to bardzo duże wykazy słownictwa, niektóre zawierają więcej niż 50 tysięcy terminów. Większość opracowano dla potrzeb szczegółowych dyscyplin i obszarów wiedzy. Przykłady to *Aquatic Sciences and Fisheries* opracowany przez Food and Agricultural Organization lub *National Aeronautic and Space Administration (NASA) Thesaurus*.

W obowiązujących obecnie normach tezaursowych pojęcie tezaursa definiowane jest wąsko, zgodnie z tradycyjnym ujęciem dokumentacyjnym. Jak wcześniej zaznaczono, zakres wykorzystywania tezaursów w sieciowym środowisku informacyjnym ulega wyraźnemu poszerzeniu. Wiele tworzonych obecnie tezaursów nie spełnia wymagań norm tezaursowych, wiele budowanych jest wprost na wzór Tezausa Rogeta. Mimo często daleko idących różnic i naruszania zasad standaryzacyjnych, wszystkie te narzędzia są jednak traktowane jako tezaursy, a to implikuje konieczność weryfikacji tradycyjnych definicji. Weryfikacja ta jest dziś w toku.

Rozwój s i e c i s e m a n t y c z n y c h wiąże się z badaniami nad komputerowym przetwarzaniem języka naturalnego. Tego typu systemy organizacji wiedzy służą do

porządkowania pojęć i terminów za pomocą struktury sieci powiązań, w których wykorzystuje się relacje znaczeniowe różnego typu. Pojęcia reprezentowane są przez węzły tych sieci, a relacje określają rodzaj połączenia między dwoma węzłami. Specyfikacja tych relacji jest znacznie dokładniejsza aniżeli ta, którą zapewniają struktury relacyjne prezentowane w tezaurusach. Obok takich związków jak rodzaj-gatunek i całość-część najczęściej oznaczane są w nich relacje przyczyna – skutek, rodzic – dziecko, obiekt – własność, etc.

Najczęściej wskazywanym SOW w formie sieci semantycznej jest WordNet opracowany w Princeton University. Jak twierdzą autorzy, konstrukcja tej sieci semantycznej inspirowana była najnowszymi teoriami psycholingwistycznymi, dotyczącymi metod porządkowania wiedzy o języku w pamięci człowieka. Sieć organizuje angielskie rzeczowniki, czasowniki, przymiotniki i przysłowki w zbiory synonimów, a każdy taki zbiór reprezentuje węzeł pojęciowy. Powiązania relacyjne łączą grupy synonimów w rozbudowaną strukturę sieciową (WordNet, b.d.). Obecnie WordNet wykorzystuje wiele wyszukiwarek internetowych. Innym, ciekawym przykładem tego rodzaju SOW jest Mimida Multilingual Semantic Network, opracowana przez Maurice Gittens na podstawie słowników 28 języków etnicznych (Gittens, b.d.).

Do najnowszych rodzajów SOW należą tzw. *m a p y m y ś l i* (ang. *think-maps*). Narzędzia te projektowane są w celu ułatwienia wieloaspektowej organizacji i przeszukiwania zasobów informacji przechowywanych w intranetach lub dostępnych w Internecie. Technologia *think-maps* wykorzystuje ideę uczenia się poprzez tworzenie tzw. map umysłu (ang. *mind-maps*), czyli wielowymiarowych układów pojęć, które mają odwzorowywać naturalne procesy myślowe przebiegające w umyśle człowieka. Ze zbioru różnorodnych informacji tworzy się najczęściej drzewiasty układ tematów szczegółowych, składających się na strukturę bardziej złożonego zagadnienia, np. rozwiązywanego problemu. *Think-maps* są więc narzędziami przekształcającymi statyczny zbiór różnych informacji w dynamiczną strukturę prezentującą powiązania między poszczególnymi jego elementami. W systemach zarządzania wiedzą umożliwiają wyświetlanie na ekranie wybranych odpowiednio do potrzeb fragmentów zbioru informacyjnego w kontekście informacji z nimi powiązanych, co ma ułatwiać to wizualizację rozwiązywanych problemów. Przykładami oprogramowania służącego do tworzenia tego rodzaju narzędzi organizacji wiedzy korporacyjnej przechowywanej w intranetach jest MindManager X5 Pro (Mindjet, 2005), ConceptDraw MindMap (ConceptDraw, 2005) czy ThinkMap firmy ThinkMap, Inc. (ThinkMap, 2004b). Wykorzystując technologię *think map* ostatnia z tych firm przygotowała także produkt o

nazwie Visual Thesaurus, umożliwiającą dynamiczną wizualizację struktur relacyjnych zapisanych w tezaurusach (ThinkMap, 2004a).

W ostatnich latach intensywnie rozwijają się badania nad SOW nazywanymi ontologiami lub *o n t o l o g i a m i f o r m a l n y m i*, przy czym znaczenie tej nazwy ciągle ewoluuje, co powoduje, że przez różnych autorów nie jest ona używana jednolicie. Jedną z najczęściej cytowanych definicji ontologii zaproponował Tom Gruber, według którego ontologia jest specyfikacją pewnej konceptualizacji. A zatem, ontologia jest opisem (podobnie jak formalna specyfikacja pewnego programu komputerowego) pojęć i relacji, które mają być rozpoznawalne dla określonego inteligentnego programu komputerowego – agenta lub pewnego zespołu takich agentów (Gruber, 1993). W procesach wymiany informacji, współkorzystania i wielokrotnego użycia wiedzy zawartej w zasobach sieciowych, ontologie takie mają zapewnić spełnienie Quinowskiej zasady ontologicznego zobowiązania.¹ Oznacza to przyjęcie pewnej umowy, dotyczącej użycia określonego słownictwa w sposób spójny (ale niekoniecznie kompletny) wobec teorii formalnie przedstawionej w określonej ontologii. Ontologie wykorzystywane są przez inteligentne programy (programowych agentów) do interpretacji wiedzy zawartej w przechowywanych zasobach informacji cyfrowej.

Według Johna F. Sowy, przedmiotem ontologii jest badanie kategorii rzeczy istniejących lub mogących istnieć w pewnej domenie. Rezultatem takich badań jest pewien katalog typów rzeczy, co do których przyjęto, że istnieją w określonej domenie D, rozpatrywanej z perspektywy osoby, która domenę tę opisuje w określonym języku L. Katalog taki nazywa się ontologią danej domeny. Wyodrębnione w ontologii typy rzeczy reprezentowane są przez predykaty, sensy słów lub typy pojęć i relacji wyrażanych w L, gdy L jest używane do omawiania tematów należących do domeny D (Sowa, 2000). W takim zatem sensie ontologia jest specyfikacją *explicite* pewnej konceptualizacji. Nawiązując do pierwotnego rozumienia ontologii w filozofii jako systematycznego objaśnienia tego co istnieje, w systemach sztucznej inteligencji przyjmuje się, że „istnieje” oznacza „może być reprezentowane”. Gdy wiedza pewnej domeny jest reprezentowana za pomocą deklaratywnego formalizmu, zbiór obiektów podlegających reprezentacji nazywany jest uniwersum dyskursu. Ten zbiór oraz opisywalne relacje zachodzące między jego elementami odwzorowane są w słownictwie, za pomocą którego program przetwarzania wiedzy

¹ Zasada zobowiązania ontologicznego Willarda Quine’a głosi, że być to być wartością zmiennej związanej, a więc nie mówi, jakie rzeczy istnieją, ale jak ustalić, istnienia jakich rzeczy wymaga dana teoria. Są to rzeczy należące do zakresów zmiennych, występujących w należyście uporządkowanym przedstawieniu formalnym teorii (Blackburn, 1994, s.269).

reprezentuje tę wiedzę. Stąd w odniesieniu do systemów sztucznej inteligencji ontologia opisywana jest przez definiowanie relacji między elementami tego słownictwa. Definicje te łączą nazwy jednostek należących do uniwersum dyskursu (np. klas, relacji, funkcji i innych obiektów) z czytelnymi dla człowieka tekstowymi opisami znaczenia ich nazw oraz formalnymi aksjomatami warunkującymi interpretację prawidłowego używania tych terminów. Z formalnego punktu widzenia, ontologia jest zatem prezentacją pewnej teorii logicznej.

J.F. Sowa rozróżnia dwa rodzaje ontologii: terminologiczne i formalne (Sowa, 2000). Ontologią terminologiczną nazywa zbiór słownictwa, którego znaczenie identyfikowane jest na podstawie sieci relacji semantycznych, a więc *de facto* terminologiczną sieć semantyczną taką jak wcześniej przytaczane sieci semantyczne WordNet i Mimida Multilingual Semantic Network lub japońska sieć EDR opracowana w Japan Electronic Dictionary Research Institute (<http://www2.nict.go.jp/kk/e416/EDR/index.html>). W ontologiach terminologicznych kategorie pojęciowe reprezentowane przez terminy nie muszą być w pełni wyspecyfikowane za pomocą aksjomatów i definicji. Ich specyfikacja oparta jest na identyfikacji relacji takich jak gatunek-rodzaj czy część-całość, które determinują względne miejsce kategorii w odniesieniu do innych pojęć. Dla większości dziedzin nauki, techniki, biznesu i prawa wypracowane zostały systemy terminologiczne czy nomenklatura, klasyfikujące i standaryzujące wyodrębnione w nich pojęcia. Aksjomatyzacja wszystkich tych pojęć w każdej dziedzinie jest zadaniem gigantycznym, ale podzbiory tej terminologii mogą być wykorzystane do zainicjowania ich formalizacji.

Ontologią formalną J.F. Sowa nazywa ontologię terminologiczną, której kategorie są rozróżnione za pomocą aksjomatów i definicji wyrażonych w pewnym języku logiki lub języku programowania, który może być przełożony na formuły logiczne. W ontologiach tego rodzaju specyfikacja obiektów uniwersum dyskursu jest z reguły znacznie bardziej precyzyjna i złożona, aniżeli odwzorowanie związków pojęciowych w sieciach semantycznych. Nie ma ograniczeń co do stopnia złożoności stosowanego języka logiki, w którym formułowane są aksjomaty i definicje. Ontologie formalne zwykle są mniejsze niż terminologiczne, ale ich aksjomaty i definicje zapewniają przeprowadzanie na podstawie systemu pojęciowego odwzorowanego w zgromadzonym i wyspecyfikowanym słownictwie bardziej złożonych procedur obliczeniowych i wnioskowania. Aby zapewnić otwartość systemu ontologie reprezentacji wiedzy nie opierają się na wyznaczonej hierarchii kategorii lecz na pewnym schemacie ich cech dystynktywnych, na podstawie którego hierarchie generowane są automatycznie. Dla każdego szczegółowego zastosowania kategorie są więc definiowane

przez selekcję odpowiedniego podzbioru rozróżnień. Dołączenie do zbioru rozróżnień podstawowych zbioru cech dystynktywnych wyodrębnionych z punktu widzenia określonego zastosowania umożliwia automatyczne wygenerowanie nowej siatki pojęciowej. Przykładami ontologii formalnych są zespoły reguł i ram opracowane dla systemów ekspertowych oraz specyfikacje schematów baz danych wyrażone w SQL-u. Ontologie opisujące wiedzę w pewnej szczegółowej dziedzinie często wykorzystywane są w systemach odkrywania wiedzy (ang. *data mining systems*) oraz w systemach zarządzania wiedzą.

5. ZAKOŃCZENIE

Różne pochodzenie, dziedziny i cele stosowania systemów organizacji wiedzy porządkujących i ułatwiających dostęp do informacji cyfrowej w środowisku sieciowym powoduje dużą proliferację oraz strukturalne i funkcjonalne zróżnicowanie tych narzędzi. Nie ulega jednak wątpliwości, że coraz powszechniejsza staje się świadomość potrzeby ich konstruowania i implementowania w sposób przemyślany i dobrze dostosowany do wymagań i potrzeb użytkowników systemów i serwisów informacyjnych.

W rozwoju SOW stosowanych w bibliotekach cyfrowych i systemach zarządzania wiedzą dostrzec można dwie tendencje. Pierwszą jest poszukiwanie narzędzi przyjaznych, łatwych w stosowaniu przez nieprzygotowanego użytkownika. Najczęściej prowadzi ono albo do upraszczania tradycyjnych, bibliotecznych i dokumentacyjnych systemów klasyfikacyjnych lub opracowania przedmiotowego, albo do tworzenia nowych, prostych kategoryzacji tematycznych. Złożoność współczesnych systemów rzeczowego opracowania zbiorów bibliotecznych jest wynikiem ich stopniowego dostosowywania do potrzeb takiego zorganizowania rozrastających się kolekcji i pogłębiającej się specjalizacji wiedzy, jakie gwarantuje możliwie precyzyjną selekcję dokumentów poszukiwanych przez użytkowników. Wiele bibliotek cyfrowych zawiera już obecnie bogate kolekcje, ale w większości przypadków są one nieporównywalnie mniejsze i mniej zróżnicowane niż zbiory tradycyjnych bibliotek średniej wielkości. Można przypuszczać, że w miarę powiększania się tych zasobów biblioteki cyfrowe będą jednak potrzebowały systemów organizacji wiedzy zapewniających większą selektywność niż oferowana przez proste podziały dziedzinowe lub kategoryzacje. Potrzeba SOW o dużej sile selekcyjnej widoczna jest też w systemach zarządzania wiedzą w organizacjach operujących wielkimi kolekcjami danych.

Drugi kierunek rozwoju SOW wskazują badania nad Semantycznym Webem, w którym organizację zawartości zasobów cyfrowych zapewnić mają rozbudowane ontologie formalne, wspierające procesy automatycznego wnioskowania oraz rozmaite procedury semantycznego przetwarzania zapisów informacji (ekstrahowania, streszczania, klasyfikowania, syntetyzowania, etc.). Na realizację tego ambitnego zamierzenia trzeba będzie zapewne jeszcze poczekać, już dziś jednak opracowane w jego ramach narzędzia znajdują zastosowanie w środowisku intranetów i w zaawansowanych bibliotekach cyfrowych. Od niedawna obserwować można też konwergencję badań nad zaawansowanymi bibliotecznymi SOW i narzędziami Semantycznego Webu. Przykładem takich działań jest analiza możliwości dostosowania Klasyfikacji Dziesiątej Deweya do wymagań projektu Semantycznego Webu (Miller, 2003).

LITERATURA

- Ariadne Genomics. (2005). *Ontology for Cell Regulatory Networks*. [online] Ariadne Genomics, Inc. [dostęp: 10.02.2005]. Dostępny w WWW: <http://www.ariadnegenomics.com/technology/ontology.html>
- ANSI/NISO. (2003). *ANSI/NISO Z39.19 Summary of Survey Results*. [online] National Information Standards Organization [dostęp: 10.02.2005]. Dostępny w WWW: <http://www.niso.org/committees/TRAG/Z39-19-Survey.pdf>
- Arystoteles. (2003). Kategorie. W: *Dzieła wszystkie*. T. 1. Warszawa: Wydawnictwo Naukowe PWN, s.32-63
- Berners-Lee, T. (1989). *Information Management: A Proposal*. [online] W3C Archive 1990. [dostęp: 10.02.2005]. Dostępny w WWW: <http://www.w3.org/History/1989/proposal.html>
- Bliss, H.E. (1933). *The organization of knowledge in libraries and the subject-approach to books*. New York, Wilson, XVI, 335 p.
- Chowdhury, G. G. (2002). Digital libraries and reference services: present and future. *Journal of Documentation* vol. 58, no 3, pp. 258-283
- ConceptDraw. (2005). *ConceptDraw MINDMAP. Brainstorming, mind mapping and visual thinking*. [online] ConceptDraw. [dostęp: 10.02.2005]. Dostępny w WWW: <http://www.conceptdraw.com/en/products/mindmap/main.php>
- Gittens, M. (b.d.). *The Mimida Project Home Page. A mechanically generated Multilingual Semantic Network*. [online] Gittens Information Systems BV. [dostęp: 10.02.2005]. Dostępny w WWW: <http://www.gittens.nl/SemanticNetworks.html>
- Gruber, T.R. (1993). A translation approach to portable ontologies. *Knowledge Acquisition* vol. 5 no 2, pp.199-220
- Hodge, G. (2000). Executive Summary. In: *Systems of Knowledge Organization for Digital Libraries: Beyond Traditional Authority Files*. [online] Washington, DC: Council of Library and Information Resources [dostęp: 5.03.2005]. Dostępny w World Wide Web: <<http://www.clir.org/pubs/reports/pub91/contents.html#foreword>>

- ISKO. (2005). *International Society for Knowledge Organization*. [online]. CEIT - Centre of Eco-Information & Terminology. [dostęp: 10.02.2005]. Dostępny w World Wide Web: <http://www.ceit.sk/wwwisis/iskoabo.htm>
- ISGN. (2003). *Glosariusz* [online]. Internetowy Serwis Głównego Negocjatora. [dostęp: 10.02.2005]. Dostępny w WWW: <http://www.negocjacje.gov.pl/neg.nsf/xml/glos1>
- KSNG. (2003). *Nazwy państw świata, ich stolic i mieszkańców*. [online] Komisja Standaryzacji Nazw Geograficznych. [dostęp: 10.02.2005]. Dostępny w WWW: http://www.gugik.gov.pl/komisja/pliki/ac_mz_wstep_publ.pdf
- Miller, E. (2003). *Weaving Meaning : An Overview of The Semantic Web*. [online] The Future of the Dewey Decimal Classification. EPC Planning Retreat. OCLC Online Computer Library Center, Inc. [dostęp: 10.02.2005]. Dostępny w WWW: <http://www.w3.org/2004/Talks/0316-semweb-ddc/slide1-0.html>
- Mindjet. (2005). *Software for Visualizing and Managing Information*. [online] Mindjet Visual Thinking. [dostęp: 10.02.2005]. Dostępny w WWW: <http://www.mindjet.com/us/>
- Miriam-Webster. (2005). *Miriam-Webster Online Dictionary*. [online] Miriam-Webster, Incorporated. [dostęp: 10.02.2005]. Dostępny w WWW: <http://www.m-w.com/cgi-bin/dictionary?book=Dictionary&va=home&x=17&y=16>
- Narecki, K. (2003). *Słownik terminów Arystotelesowskich*. W: Arystoteles. *Dzieła wszystkie*. T.7. Warszawa: Wydawnictwo Naukowe PWN, s.11-127
- Needham, Ch.D. (1965). *Organizing knowledge in libraries: an introduction to classification and cataloguing*. [London]: A. Deutsch, 259 p.
- NISO. (1999). *Report on the Workshop on Electronic Thesauri, November 4-5, 1999*. [online] National Information Standards Organization [dostęp: 10.02.2005]. Dostępny w WWW: http://www.niso.org/news/events_workshops/thes99rprt.html
- NISO. (2005). *Developing the Next Generation of Standards for Controlled Vocabularies and Thesauri*. [online]. National Information Standards Organization [dostęp: 10.02.2005]. Dostępny w WWW: <http://www.niso.org/committees/MT-info.html>
- SearchEngine. (b.d.). *Słownik*. [online] SearchEngine.pl. [dostęp: 10.03.2005]. Dostępny w WWW: http://www.searchengines.pl/podstawy_sloownik.php3
- PWN. (2004). *Słownik języka polskiego*. [online] Wydawnictwo Naukowe PWN S.A. [dostęp: 10.02.2005]. Dostępny w WWW: <http://sjp.pwn.pl/>
- Rosenfeld, L.; Morville, P. (2003). *Architektura informacji w serwisach internetowych*. Gliwice: Wydawnictwo Helion, 500 s.
- Sosińska-Kalata, B. (2002). *Klasyfikacja: struktury organizacji wiedzy , piśmiennictwa i zasobów informacyjnych*. Warszawa: Wydaw. SBP, 262 s.
- Sosińska-Kalata, B. (1999). *Modele organizacji wiedzy w systemach wyszukiwania informacji o dokumentach*. Warszawa: Wydawnictwo SBP, 388 s.
- Sowa, J.F. (2000). *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Pacific Grove, CA: Brooks Cole Publishing; także online: <<http://www.jfsowa.com/krbook/index.htm>>
- Sowa, J.F. (2001). *Glossary*. [online] KR Ontology. [dostęp: 5.02.2005]. Dostępny w WWW: <http://www.jfsowa.com/ontology/gloss.htm#Axiomatized>
- Taylor, J.R. (2001). *Kategoryzacja w języku. Prototypy w teorii językoznawczej*. Kraków: Universitas

- ThinkMap. (2004a). *Introducing the Visual Thesaurus Version 3*. [online] ThinkMap Visual Thesaurus. [dostęp: 10.02.2005]. Dostępny w WWW: <http://www.visualthesaurus.com/>
- ThinkMap. (2004b). *Thinkmap Visualization Software Facilitates Communication, Learning and Discovery*. [online] Thinkmap, Inc. [dostęp: 10.02.2005]. Dostępny w WWW: <http://www.thinkmap.com/>
- Tudhope, D.; Koch, T. (2003). New Applications of Knowledge Organization Systems: introduction to a special issue. [online] *Journal of Digital Information* vol. 4 issue 4 article no. 286, [dostęp: 10.02.2004]. Dostępny w World Wide Web: <<http://jodi.ecs.soton.ac.uk/Articles/v04/i04/editorial>>
- USGS. (2004). *Geographic Names Information System (GNIS)*. [online] US Geological Survey. [dostęp: 10.02.2005]. Dostępny w WWW: <http://geonames.usgs.gov/>
- Witten, Ian; Akscyn, Rob; Shipman, Frank M. III, eds. (1998). *International Conference on Digital Libraries. Proceedings of the third ACM conference on Digital libraries*. [online] ACM Portal. Joint Conference on Digital Libraries. [dostęp: 10.02.2005]. Dostępny w World Wide Web: <<http://portal.acm.org/toc.cfm?id=276675&idx=SERIES492&type=proceeding&coll=portal&dl=ACM&part=series&WantType=Proceedings&title=International%20Conference%20on%20Digital%20Libraries&CFID=40169612&CFTOKEN=57901697>>
- WordNet. (b.d.). *WordNet : a lexical database for the English language*. [online] Princeton University. Cognitive Science Laboratory. [dostęp: 10.02.2005]. Dostępny w WWW: <http://wordnet.princeton.edu/>
- Zeng, M.L. (2003). *Subject Access Approaches used by Digital Collections and Information Directories. Report on current progress*. [online] IFLANET. Classification and Indexing Section. Working Group on Subject Access to Web resources. [dostęp: 10.02.2005] Dostępny w WWW: <http://www.ifla.org/VII/s29/wgsaw.htm>

ABSTRACT

The term *knowledge organization systems* (KOSs) means schemas of various kinds, that organize information and support knowledge management in various environments and for various purposes. The most representative areas of contemporary research and implementation of KOSs in networked environment are digital libraries, Internet and intranet systems and services of various organizations and the Semantic Web project. The KOSs applied in those systems include tools that range from relatively simple controlled vocabulary lists to taxonomies, classification and categorization schemes, thesauri, semantic networks and ontologies. In the article the basic features common to all kind of KOSs are indicated, and the most important types of the systems are discussed.