

Lidia Derfert-Wolf

Biblioteka Główna Akademii Techniczno-Rolniczej w Bydgoszczy

Serwisy tematyczne o kontrolowanej jakości w Internecie - *subject gateways*

Początki korzystania z Internetu to etap entuzjazmu nad niewiarygodnymi możliwościami komunikowania, publikowania i wyszukiwania informacji w sieci. Ci, którzy nabyli umiejętności odnajdywania dokumentów, poznali działanie wyszukiwarek i wiedzą ile cennych informacji jest poza ich zasięgiem, przekonali się szybko, że narzędzia typu Google nie wystarczają do przeszukiwania sieci w celu dostępu do informacji związanej z określoną tematyką. Ponadto pojawił się problem porządkowania i jakości informacji, zwłaszcza potrzebnej dla nauki i edukacji. Rozwiązaniem we wczesnej fazie rozwoju Internetu stały się katalogi tematyczne typu *directory* (np. Yahoo!) grupujące witryny wg popularnych dziedzin. Doborem witryn zajmują się ludzie, którzy na podstawie przeglądu strony decydują o jej umieszczeniu w danej kategorii. Narzędzia te rejestrują zasoby całej sieci i nie posługują się określonymi kryteriami doboru źródeł czy ich opisu i klasyfikacji. Rzadko również uwzględniają zasoby tzw. sieci niewidzialnej (*deep web, invisible web*), która pod względem wielkości ok. 500-krotnie przewyższa sieć "widoczną" (*surface web*) [2]. Równoległe do katalogów tematycznych, które obecnie przekształciły się w wielkie portale, powstawały niezależne listy linków uporządkowanych wg tematów lub rodzajów informacji (katalogi biblioteczne, encyklopedie, informatory). Najwięcej z nich spotkać można do tej pory na witrynach bibliotek, również polskich [4, 6, 9]. Ich jakość niekiedy pozostawia wiele do życzenia, a metodologia doboru linków, ich opisu i aktualizacji jest bardzo zróżnicowana i rzadko opisana w serwisie. Działy tematyczne najczęściej odzwierciedlają dziedziny reprezentowane w danej uczelni bądź wszystkie tematy, jeśli wykaz tworzy biblioteka publiczna [6]. Do rzadkości należy możliwość wyszukiwania poprzez zadawanie pytań. Najczęściej mamy do wyboru jedynie wertowanie wg tematów. Ciekawe rozwiązanie zastosowano na witrynie Działu Informacji Bibliografii WiMBP w Łodzi (<http://www.wimbp.lodz.pl/informacja>), gdzie linki uporządkowane są wg UKD, zaopatrzone w krótkie adnotacje i jednocześnie można je przeszukiwać zadając pytania prostej wyszukiwarce [6]. Tworzenie i stałe aktualizowanie tego typu przewodników po zasobach sieci jest z pewnością bardzo pracochłonne i obecny ich kształt na witrynach polskich bibliotek nie spełnia do końca wymagań profesjonalnych narzędzi tego rodzaju, posługujących się np. kontrolowanym słownikiem haseł przedmiotowych. Warto byłoby z pewnością połączyć wysiłki np. w ramach uczelni o podobnym profilu kształcenia i tworzyć jeden serwis krajowy z danej dziedziny bazując oczywiście na najlepszych wzorcach już istniejących wykazach. Podobnie jak w sytuacji zaniechania tworzenia linków do witryn i katalogów bibliotek na rzecz umieszczania adresu do wykazu linków w EBIB.

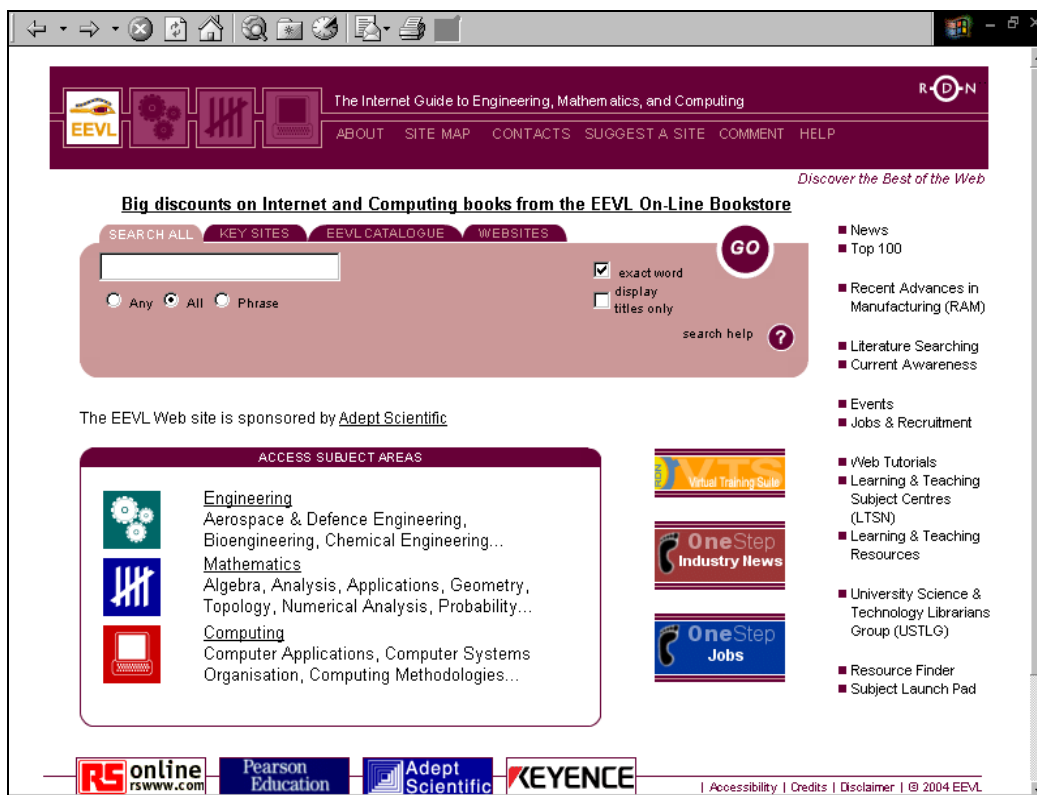
Kolejnym etapem w światowym Internecie było projektowanie i rozwijanie przewodników po sieci, później portali tematycznych, współtworzonych głównie przez bibliotekarzy. W odróżnieniu od katalogów typu Yahoo! te serwisy zawierają w większości krótkie opisy cytowanych źródeł sieciowych oraz często są uporządkowane wg określonej klasyfikacji. Najstarszy tego typu katalog to WWW Virtual Library (<http://www.vlib.org/>) tworzony przez wolontariuszy z całego świata i zawierający 312 odrębnych tematycznych - „bibliotek wirtualnych”. Kolejne indeksy linków do zasobów internetowych - Librarians' Index to the Internet oraz BUBL LINK / 5:15 – zostaną omówione w dalszej części, ponieważ są przykładami serwisów typu *subject gateways*, omawianymi w niniejszym artykule. Tego typu serwisy różnią się od wspomnianych wyżej przede wszystkim poziomem jakości prezentowanych linków oraz strukturą organizacyjną i metodami gromadzenia danych.

„Subject gateways” – pojęcia ogólne, definicje

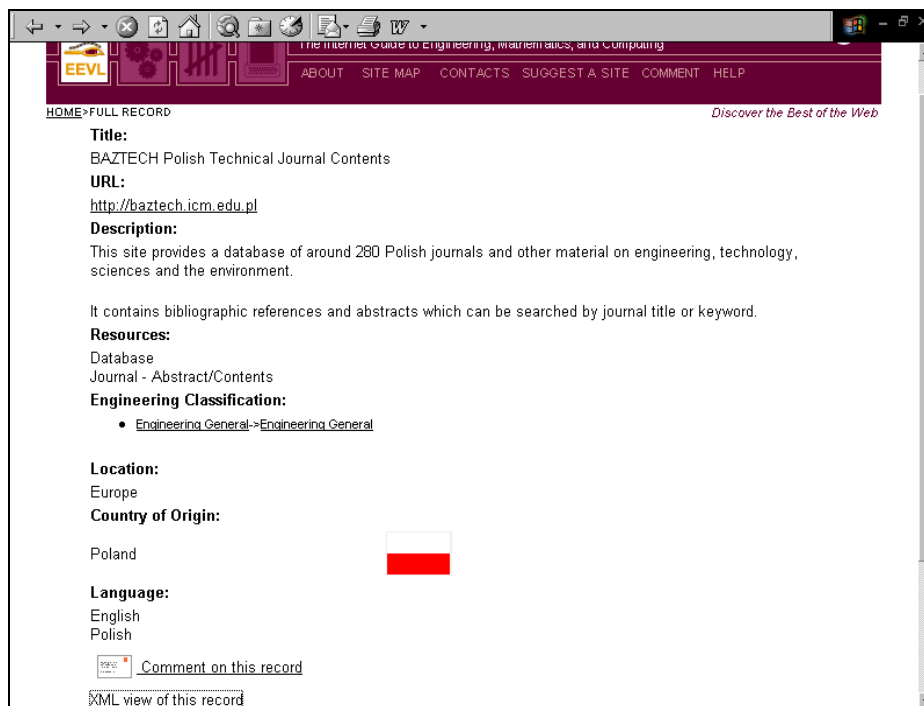
W odpowiedzi na zapotrzebowanie na uporządkowany tematycznie dostęp do wysokiej jakości źródeł, głównie dla środowisk naukowych, powstały nowe typy serwisów, nazywane ogólnie *subject gateways*. Są to dziedzinowe przewodniki (miejsca startowe) po zasobach internetowych (dokumentach, obiektach, witrynach, serwisach), uporządkowane wg dziedzin. Zasoby są selekcyjonowane, oceniane, opisywane i katalogowane przez bibliotekarzy lub ekspertów z danej dziedziny. *Subject gateways* są najczęściej ukierunkowane na dziedziny związane z obszarami zainteresowań akademickich. Linki zgromadzone w tych serwisach dobiera się zgodnie z oficjalnie opublikowaną listą kryteriów oceny jakości, a później kataloguje i opisuje również wg powszechnie stosowanych systemów klasyfikacyjnych. W porównaniu z ogólnymi i niekiedy przypadkowymi

wykazami linków, *subject gateways* są bardziej pracochłonne, ale w rezultacie dają zbiór (bazę danych) o kontrolowanej jakości do przeszukiwania wg słów kluczowych i/lub wertowania wg kategorii tematycznych. Rezultaty zawierają standardowe opisy, które umożliwiają użytkownikom ocenić przydatność informacji. Ogólnie rzecz ujmując, tego typu serwisy pomagają w wyszukiwaniu informacji w sieci na dany temat i dają gwarancję odnajdywania źródeł o wysokiej jakości. Termin *subject gateways* przyjęty dla potrzeb niniejszego artykułu w oryginale z powodu braku dobrego polskiego tłumaczenia, jest pewnym uproszczeniem. Podana wyżej krótka charakterystyka odnosi się do serwisów, których przykłady podano poniżej, zwanych w literaturze również: *subject-based information gateways (SBIGs)*, *quality-controlled subject gateways*, *subject-based gateways*, *subject index gateways*, *virtual libraries*, *clearing houses*, *subject trees*. Najczęściej spotykanym i zdaje się najwłaściwszym jest pojęcie *subject-based information gateways (SBIGs)*, wprowadzone w trakcie realizacji projektu DESIRE (opis poniżej). Natomiast najczęściej cytowana definicja tego typu serwisów jest autorstwa T. Kocha i ujmuje ona *quality-controlled subject gateways* jako serwisy internetowe, wykorzystujące bogaty zestaw wskaźników jakości pomagający systematycznie wyszukiwać źródła. Do zapewnienia doboru źródeł zgodnie z kryteriami jakościowymi oraz ich opisu na podstawie standardowych metadanych wkłada się znaczny trud intelektualny. Natomiast gwarancją regularnej aktualizacji i kontroli linków jest profesjonalne zarządzanie ich kolekcją. Głównym celem tych serwisów jest umożliwienie dziedzinowego dostępu do źródeł o wysokiej jakości, poprzez indeksowanie zasobów Internetu przy pomocy słowników kontrolowanych oraz poprzez organizację linków wg głębokiej struktury tematycznej, umożliwiającej jej przeglądanie (wertowanie) oraz przeszukiwanie [8].

Dla zobrazowania charakterystyki *subject gateways*, na poniższych rysunkach przedstawiono witrynę przykładowego serwisu (rys. 1) oraz rekord bibliograficzny z tego samego serwisu opisujący przykładowe źródło internetowe (rys. 2). Ilustracje te będą z pewnością pomocne przy lekturze dalszej części artykułu.



Rys. 1 Witryna serwisu EEVL - the Internet Guide to Engineering, Mathematics and Computing.



Rys. 2. Rekord bibliograficzny źródła internetowego w serwisie EEVL - the Internet Guide to Engineering, Mathematics and Computing.

Cechy charakterystyczne i metodologia tworzenia

Szczegółowa charakterystyka *subject gateways* wymagałaby kilku publikacji. Niniejszy artykuł wskazuje jedynie na najważniejsze aspekty ich funkcjonowania, w dużym skrócie. W literaturze krajowej brak dokładnych i analitycznych opracowań na ten temat. Materiał zebrano na podstawie literatury zagranicznej [1, 3, 5, 8] i przeglądu witryn internetowych. Przykładowe serwisy typu *subject gateways* opisano w dalszej części. Pełniejsze zestawienie adresów umieszczono w niniejszym numerze EBIBa w dziale „Warto wiedzieć” pn. „Serwisy tematyczne o kontrolowanej jakości w Internecie - *subject gateways*” [<http://ebib.oss.wroc.pl/2004/57/wykaz.php>], gdzie znaleźć można adresy wszystkich cytowanych *subject gateways* i projektów międzynarodowych. Omówione poniżej cechy stanowiły podstawę wypracowania definicji serwisów *quality-controlled subject gateways* przez T. Kocha oraz stanowią zestaw kryteriów, wg których można stwierdzić czy jakiegokolwiek serwis grupujący linki do ciekawych stron internetowych spełnia wymagania *subject gateways*, a dokładniej *quality-controlled subject gateways* czy *subject-based information gateways (SBIGs)* [8].

1) Zakres, wielkość baz danych, zawartość

Subject gateways są w większości tworzone na terenie jednego kraju/regionu bądź przez pojedyncze instytucje. W przeważającej liczbie gromadzą źródła angielskojęzyczne i/lub źródła w języku danego kraju. Najczęściej reprezentowane typy indeksowanych źródeł to pełne teksty dokumentów, witryny organizacji, bazy danych, spisy linków, serwisy, materiały komercyjne, oprogramowania, książki elektroniczne, dokumenty multimedialne. Dominującym formatem dokumentów jest HTML. Serwisy liczą od ok. 2 tys. do ponad 10 tys. linków. Większość z nich nie ogranicza doboru do miejsca posadowienia dokumentu (terytorium). Tematyka jest zazwyczaj ograniczona do jednej dziedziny wiedzy, np. rolnictwo, medycyna. Rzadziej spotykane są indeksy ogólne.

2) Zarządzanie kolekcją linków

Zarządzanie kolekcją jest w tym przypadku pojęciem bardzo szerokim i obejmuje zarówno gromadzenie źródeł (wyszukiwanie w sieci), opracowanie oraz utrzymanie zbioru opisów bibliograficznych w bazie danych. W tym miejscu warto wspomnieć o tym, kto tworzy *subject gateways*. Pierwsze tego typu serwisy powstawały w wielkich bibliotekach z inicjatywy bibliotekarzy (BUBL, LII). Ze względu na wymaganą wysoką jakość przy wyszukiwaniu i katalogowaniu źródeł do współpracy przystąpili eksperci z odpowiednich dziedzin. Wiele serwisów umożliwia zgłaszanie źródeł przez użytkowników sieci, którzy muszą podać adres URL, krótki opis oraz słowa kluczowe. Jednak zawsze w takim przypadku zgłoszenie jest weryfikowane, opis uzupełniany i dodawane dodatkowe

elementy do rekordu bibliograficznego. Zajmują się tym redaktorzy serwisów. Korzystając z takiej możliwości zgłoszono np. polską bazę danych BAZTECH do kilku prestiżowych *subject gateways* z zakresu nauk technicznych (rys. 2). Odpowiedzialnymi za tworzenie są najczęściej uczelnie, a finansowanie pozostaje w gestii odpowiednich ministerstw, agencji rządowych czy fundacji i dodatkowo sponsorów. Bardzo dużo *subject gateways* powstało lub było wspieranych metodologicznie w ramach projektów finansowanych przez UE. Polityka gromadzenia i opracowania źródeł internetowych oraz utrzymania zbioru opisów opiera się na opublikowanych zasadach (niekiedy dostępne w sieci). Należy pamiętać, że selekcja zasobów nie odbywa się automatycznie, ale manualnie – ludzie gromadzą i oceniają dokumenty. Jeśli chodzi o dobór źródeł zasady dotyczą głównie kompletności i zakresu kolekcji biorąc pod uwagę odbiorców, dla których jest serwis tworzony. Należy sprecyzować zakres tematyczny, akceptowane typy dokumentów, poziom prezentowanej informacji, rodzaj dostępu (koszty, wymagane hasła), stosowane technologie na witrynach, terytorium, język. Kolejne ważne dokumenty i instrukcje koncentrują się wokół jasnych i klarownych kryteriów oceny dokumentów. Najczęściej są to:

- kryteria oceny zawartości (ocena informacji zawartych w źródle) – wiarygodność, autorstwo i reputacja, dokładność, kompletność, unikalność, układ i organizacja treści, aktualność, zgodność z wymaganą tematyką,
- kryteria oceny formy – łatwość nawigacji, pomoce dla użytkowników, korzystanie ze standardowych rozwiązań technicznych, estetyka.

Z kolei polityka utrzymywania kolekcji linków i ich opisów powinna gwarantować aktualność całego serwisu. Dotyczy to systematycznego sprawdzania rekordów pod kątem zmian i kontroli czy opis jest nadal adekwatny do źródła (np. uzupełniania adnotacji o nowości w wyszukanych źródłach), regularnego aktualizowania kolekcji (dodawania nowych linków z określoną częstotliwością), sprawdzania linków już zarejestrowanych w bazie.

3) Oprogramowanie

Wdrożenie i eksploatacja *subject gateways* wymaga korzystania ze specjalistycznych programów do przechowywania danych (niekiedy kopii witryn), indeksowania źródeł oraz wyszukiwania. W zależności od serwisu są to programy darmowe lub komercyjne. Większość europejskich systemów wykorzystuje pakiet opracowany w ramach projektu ROADS, o którym będzie mowa w dalszej części artykułu. Opisy (adnotacje, abstrakty, symbole klasyfikacyjne) przechowywane są w bazie danych. Wyszukiwanie odbywa się zazwyczaj przy pomocy interfejsu opartego na przeglądarkach WWW i daje do dyspozycji użytkownikom kilka poziomów zaawansowania oraz techniki przypominające wyszukiwanie w profesjonalnych bazach danych. Witryny będące rezultatem wyszukiwania są najczęściej w HTML lub Java.

4) Metadane i opis zawartości

Źródła sieciowe w *subject gateways* są katalogowane pod względem formalnym (np. autor/twórca, tytuł, data, adres URL, typ dokumentu, język, data skatalogowania, nazwisko katalogującego) wg standardowych formatów. Zwykle jest to Dublin Core albo RDF [10, 11, 12]. Nie wszystkie serwisy ujawniają to na swoich witrynach. W niektórych można po wyszukaniu opisu źródła wybrać opcję wyświetlania rekordu w formacie np. Dublin Core, podobnie jak w systemach bibliotecznych – MARC. Opis przedmiotowy tworzony jest dla indywidualnego źródła również zgodnie ze standardami, ale o zdecydowanie wyższej jakości niż ma to miejsce w przypadku stron internetowych, gdzie w dokumencie HTML, w polu „keywords” czy „description” wpisywane są dowolne treści, niekiedy nie mające związku z dokumentem. W *subject gateways* obowiązują publikowane zasady opisu dokumentu (źródła sieciowego), który jest przeważnie sformalizowaną i spójną adnotacją, streszczeniem lub recenzją opracowywaną zawsze przez specjalistę z danej dziedziny. To właśnie jedna z cech odróżniających te serwisy od zwykłych wyszukiwarek, gdzie opisem jest najczęściej uzyskiwany automatycznie początek albo fragment tekstu dokumentu. Większość metadanych w *subject gateways* jest wyświetlana jako rezultat wyszukiwania (opis bibliograficzny źródła) i już na tym etapie – zanim połączymy się ze wskazanym źródłem – możemy na podstawie metadanych ocenić wstępnie jego wartość i przydatność do dalszej pracy.

5) Dostęp dziedzinowy

Jako minimalne wymaganie dla przedmiotowego opisu źródła i dostępu do źródeł w *subject gateways* uważa się wysoki poziom jakościowy kategorii tematycznych porządkujących zasób i dostępnych dla użytkownika do ich przeglądania w celu odnajdywania źródeł. Drugim warunkiem podstawowym jest indeksowanie źródeł wg niekontrolowanych słów kluczowych wspomagających proces wyszukiwania.

Głęboka struktura hierarchiczna oparta na znanych systemach klasyfikacyjnych i indeksowanie na podstawie kontrolowanych słowników (np. tezaury, biblioteczne systemy klasyfikacyjne) wskazuje na wysoki poziom *subject gateways* [8, 16]. Do najczęściej wykorzystywanych klasyfikacji należą:

- klasyfikacje uniwersalne, np. Uniwersalna Klasyfikacja Dziesiętna, Klasyfikacja Dziesiętna Deweya, Klasyfikacja Biblioteki Kongresu,
- ogólne klasyfikacje narodowe, np. holenderska *Nederlandse Basisclassificatie*, szwedzka *Sveriges Allmänna Biblioteksforening*
- klasyfikacje specjalistyczne, np. klasyfikacja NLM (National Library of Medicine), Ei Engineering Information Classification Codes, MSC Mathematics Subject Classification.

Natomiast najpowszechniej stosowanymi tezaurusami są: Ei thesaurus (nauki techniczne), HASSET thesaurus (nauki społeczne), MeSH thesaurus (medycyna), Arts and Architecture Thesaurus (architektura i sztuka), Thesaurus of Geographic Names, hasła przedmiotowe Biblioteki Kongresu LCSH i CAB (rolnictwo).

6) Dodatkowe możliwości

Większość *subject gateways* umożliwia zarówno wyszukiwanie wg terminów jak również przeglądanie bazy danych opisów źródeł. Końcowym rezultatem jest bezpośrednie połączenie (link) do określonej witryny czy serwisu. Podstawowe techniki wyszukiwania są zbliżone do technik powszechnie stosowanych w wyszukiwarkach. Jednakże na poziomie danego *subject gateway* wpływa umożliwienie wyszukiwania przy pomocy złożonych zapytań z użyciem operatorów Boole'a, możliwość „maskowania” końcówek, korzystania z tezaurusów, ograniczania rezultatów do ściśle określonego terytorium, języka czy typu dokumentu. Niektóre pozwalają nawet ograniczyć wybór do poziomu źródła czyli wybrać materiały odpowiednie dla uczniów, studentów czy naukowców. Do innych możliwości i elementów wybranych *subject gateways*, poza interfejsem do wyszukiwania należą:

- lokalne przechowywanie kopii skatalogowanych źródeł internetowych (dotyczy to wyłącznie statycznych witryn),
- podręczniki online na tematy wyszukiwania informacji w sieci,
- listy dyskusyjne dla użytkowników, kalendarze wydarzeń, linki do witryn ekspertów,
- linki do księgarni, serwisów o pracy i wiele innych.

7) Współdziałanie

Kolejnym etapem w rozwoju *subject gateways* jest tworzenie wspólnych „bramek” wyszukiwawczych dla kilku serwisów. Istnieje już kilka projektów (np. Renardus, RDN) jednoczących *subject gateways* i tworzących serwisy typu cross-searching, których celem jest jednoczesne przeszukiwanie kilku *gateway* [7, 13]. Może to się odbywać na poziomie krajowym albo tematycznym. Współpraca polega albo wyłącznie na umożliwianiu wyszukiwania w kilku bazach jednocześnie, albo dodatkowo na współkatalogowaniu zasobów internetowych. W pierwszym przypadku poszczególne serwisy nie zmieniają zasad swojego funkcjonowania i nadal pozostają niezależne. Jednak dla użytkowników w wielu sytuacjach może być cenniejszy dostęp do wielu indeksów z jednego miejsca. Podstawowym i koniecznym warunkiem do współdziałania różnych serwisów jest stosowanie standardowych metod opisów danych (np. Dublin Core czy RDF) oraz standardów umożliwiających jednoczesne wyszukiwanie (np. Z39.50). Wymieniane niżej europejskie projekty (ROADS, DESIRE) dotyczyły wypracowywania metod i narzędzi programowych pomocnych w tworzeniu wspólnych interfejsów dla kilku samodzielnych *subject gateways*.

Projekty wspomagające tworzenie *subject gateways*, współpraca międzynarodowa

Pierwsze typowe serwisy *subject gateways* z określonej dziedziny wiedzy powstawały w Europie w latach 90-tych, głównie w Wielkiej Brytanii w ramach programu eLib (ADAM, EEVL, Sosig) czy w Szwecji (Engineering E-Library, EELS). W celu ujednoczenia prac, wspomagania ich rozwoju oraz opracowywania wskazówek dla twórców nowych serwisów, realizowano kilka ważnych projektów międzynarodowych [15]. Poniżej omówione zostaną cztery z nich: DESIRE, ROADS, IMesh oraz Renardus.

ROADS <http://www.ilrt.bris.ac.uk/roads/>

Program był realizowany w ramach eLib przy współpracy Wydziału Informatyki w Politechnice w Loughborough, Instytutu Kształcenia i Badań w Uniwersytecie w Bristol oraz UKOLN (UK Office for Library and Information Networking) w Uniwersytecie w Bath i miał następujące cele:

- opracowanie pakietu oprogramowania i narzędzi do tworzenia i utrzymania *subject-specific gateways*,
- opracowanie metod dla serwisów typu cross-searching i współdziałania pomiędzy *gateways*,

- udział w rozwijaniu standardów indeksowania, katalogowania i wyszukiwania tematycznego źródeł internetowych.

Oprogramowania ROADS są wykorzystywane przez wiele europejskich subject gateways, min. ADAM, Biz/ed, EELS, Finnish Virtual Library, OMNI, Sosig. Mimo, że projekt został zakończony, narzędzia i metody wypracowane w trakcie jego realizacji nadal służą wszystkim nowym twórcom *subject gateway* i są udostępniane poprzez witrynę internetową.

DESIRE <http://www.desire.org/>

Projekt DESIRE (Development of a European Service for Information on Research and Education), realizowany był w latach 1998-2000 w ramach EU Telematics Application Programme, we współpracy pomiędzy 10 instytucjami z krajów europejskich (Holandia, Norwegia, Szwecja i Wielka Brytania). Skupiał się na promowaniu, ulepszaniu i łączeniu istniejących sieci informacyjnych dla użytkowników z obszarów nauki w Europie. Badania prowadzono w trzech kierunkach: przechowywanie kopii dokumentów online najczęściej wykorzystywanych przez użytkowników (caching), łączenie manualnego katalogowania zasobów sieciowych z automatycznym indeksowaniem WWW (resource discovery) oraz łączenie niezależnych *subject gateways* w serwisy na wielką skalę dające jeden punkt startowy do wyszukiwania informacji (directory services). Z jednej strony rezultaty DESIRE mają dać naukowcom efektywny dostęp do zasobów sieciowych o wysokiej jakości, a z drugiej strony – „miejsce startowe” dla bibliotekarzy i instytucji zainteresowanych problemami katalogowania, indeksowania zasobów sieciowych oraz w ogóle tworzenia *subject gateways* i ich międzynarodowych sieci [18]. Opracowano bogatą kolekcję narzędzi programowych, raportów, procedur i poradników, które są bezpłatnie dostępne w sieci. Do godnych polecenia i wykorzystania należą min.:

- DESIRE Information Gateways Handbook <http://www.desire.org/handbook/> - podręcznik dla twórców portali sieciowych o kontrolowanej jakości, składający się z trzech części: zagadnienia strategiczne dla zarządzających i sponsorów, zagadnienia związane z gromadzeniem i opracowaniem informacji dla ekspertów dziedzinowych i specjalistów ds. informacji, zagadnienia techniczne. W podręczniku są też wymienione kryteria oceny jakości źródeł (Quality selection: ensuring the quality of your collection <http://www.desire.org/handbook/2-1.html>) oraz zasady zarządzania kolekcją linków („Collection management” <http://www.desire.org/handbook/2-6.html>)
- Podręcznik online „Internet Detective” o tym jak szukać i jak oceniać informacje w sieci <http://www.sosig.ac.uk/desire/internet-detective.html>.

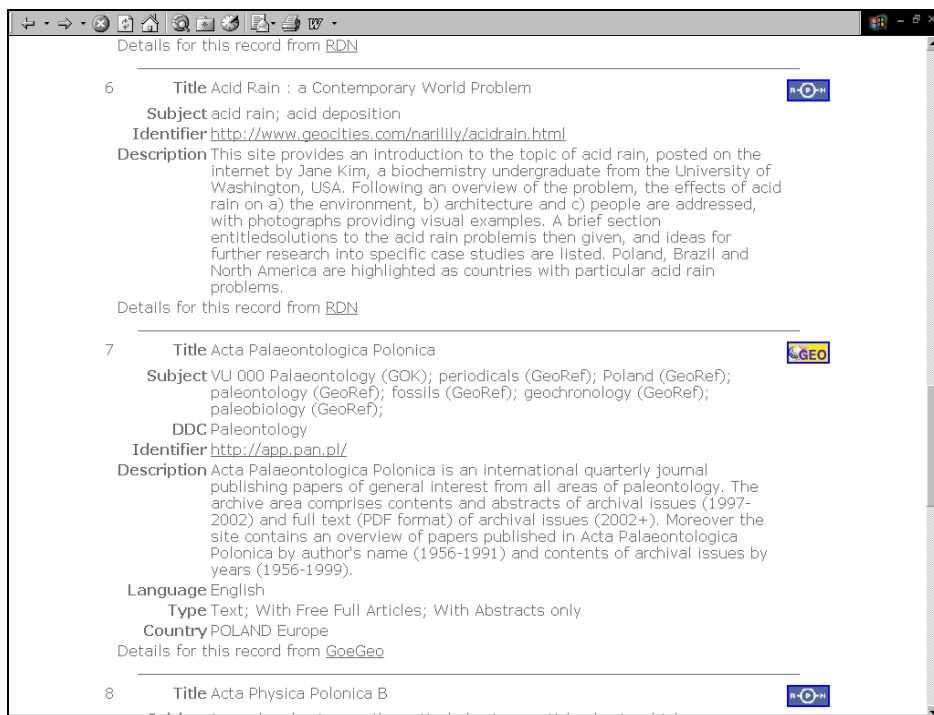
Imesh: International Collaboration on Internet Subject Gateways <http://www.imesh.org/>

Międzynarodowa inicjatywa powstała w celu wspierania komunikacji i współpracy pomiędzy twórcami *subject gateway* i innymi grupami instytucji. Skupia się na następujących działaniach i inicjatywach::

- IMesh Toolkit <http://www.imesh.org/toolkit/> - poradnik dotyczący architektury i oprogramowania dla *subject gateways*,
- IMesh Community <http://www.desire.org/html/subjectgateways/community/imesh/> - m.in. lista dyskusyjna oraz zestawienie publikacji i ciekawych linków,
- międzynarodowe warsztaty IMesh Workshops.

Renardus <http://www.renardus.org/>

Renardus umożliwia wyszukiwanie wśród źródeł internetowych zebranych i dokładnie opisanych przez *subject gateways* z różnych krajów europejskich. Szukać można poprzez zadawanie pytań albo przeglądanie wg kategorii tematycznych uporządkowanych na podstawie UKD. Pytania kierowane są do bazy opisów źródeł (metadanych), a nie ich pełnych tekstów. W rezultacie wyszukiwania otrzymujemy tytuł źródła, hasła przedmiotowe, niekiedy hasło UKD, język, kraj, typ dokumentu, adres URL, opis, nazwa serwisu subject gateway, który źródło zindeksował (rys. 3). Specjalną funkcją serwisu Renardus jest opcja przeglądania (Browse by Subject) wg kategorii i podkategorii tematycznych w wersji tekstowej i graficznej. Serwisy uczestniczące w projekcie Renardus (min. DutchESS, Danmarks Elektroniske Forskningsbibliotek, Finnish Virtual Library, Geo-Guide, History Guide, MathGuide, NOVAGate, Online theses – Germany, RDN Resource Discovery Network) dostarczają łącznie ok. 64.000 źródeł sieciowych, w przeważającej części zdigitalizowanych, z prawie wszystkich dziedzin nauki, głównie w języku angielskim. Serwis powstał w ramach projektu finansowanego przez Komisję Europejską i jest utrzymywany przez Niedersächsische Staats- und Universitätsbibliothek, Göttingen, Germany. Biblioteki uczestniczące tworzą obecnie konsorcjum, którego grupę zarządzającą stanowią: Niedersächsische Staats- und Universitätsbibliothek w Getyndze, NetLab przy szwedzkiej uniwersytecie w Lund, biblioteka uniwersytecka w Jyväskylä w Finlandii oraz Institute for Learning and Research Technology w uniwersytecie w Bristolu.



Rys. 3 Rekordy bibliograficzne źródeł internetowych w serwisie typu cross-searching Renardus.

Przykłady *subject gateways*

Dla zobrazowania materiału zawartego w niniejszym artykule warto przedstawić bliżej kilka serwisów typu *subject gateways*. Nie ma oczywiście żadnych wykazów w sieci rejestrujących wszystkie systemy. Zainteresowanych najważniejszymi odsyłam do załącznika „Serwisy tematyczne o kontrolowanej jakości w Internecie - subject gateways” [<http://ebib.oss.wroc.pl/2004/57/wykaz.php>]. Nie ma również standardowej typologii tych serwisów, zwłaszcza że wiele z nich można zaliczyć do kilku grup. Poza tym niektóre działają niezależnie i jednocześnie są częściami systemów typu *cross-searching*, np. RDN. *Subject gateways* można dzielić wg tematyki, języka, terytorium, współpracy i wielu innych. W zakresie tematyki będą to serwisy rejestrujące źródła ogólnodostępne (np. BUBL, ILL, ScoutReport, DutchESS) lub dotyczące jednej dziedziny (np. SOSIG, AVEL, Agrigate, Biz/ed). W zakresie terytorium można je dzielić z kolei na takie, które są tworzone na terenie jednego kraju we współpracy kilku instytucji i rejestrują zasoby światowe (np. DutchESS) lub wyłącznie z danego kraju (np. AVEL, Agrigate). Stosując podział wg rodzaju współpracy wyróżnić można serwisy, w których dokonano podziału pracy (dobór, katalogowanie) w ramach jednego *subject gateway* (np. NOVAGate, DutchESS), albo współpraca dotyczy jedynie wspólnego interfejsu do wyszukiwania w kilku serwisach (np. RDN, Renardus). Poniższe przykłady są to najciekawszymi zdaniem autorki serwisami, zarówno pod względem tworzenia, jak i współpracy międzynarodowej.

BUBL Information Service <http://bubl.ac.uk/>

Serwis tworzony przez bibliotekarzy (Centre for Digital Library Research, Strathclyde University, UK) składający się z dwóch modułów:

- **BUBL LINK** – linki do wybranych źródeł internetowych ze wszystkich dziedzin wiedzy akademickiej skatalogowanych wg Klasyfikacji Dziesiętnej Deweya i sprawdzanych co miesiąc.
- **BUBL 5:15** - alternatywny interfejs bazujący na hasłach przedmiotowych opartych na hasłach Biblioteki Kongresu. Nazwa 5:15 pochodzi od założenia, żeby do każdej dziedziny było przypisanych przynajmniej 5 źródeł, ale nie więcej niż 15, mimo że ten ostatni warunek nie zawsze udaje się spełnić.

Oczywiście jest również dostępna opcja wyszukiwania wg słów i wyrażeń na trzech poziomach zaawansowania oraz opcja przeglądania wg typów dokumentów oraz krajów. W katalogu BUBL zarejestrowanych jest obecnie ponad 11.000 źródeł.

LIJ Librarian's Index to the Internet <http://lii.org/>

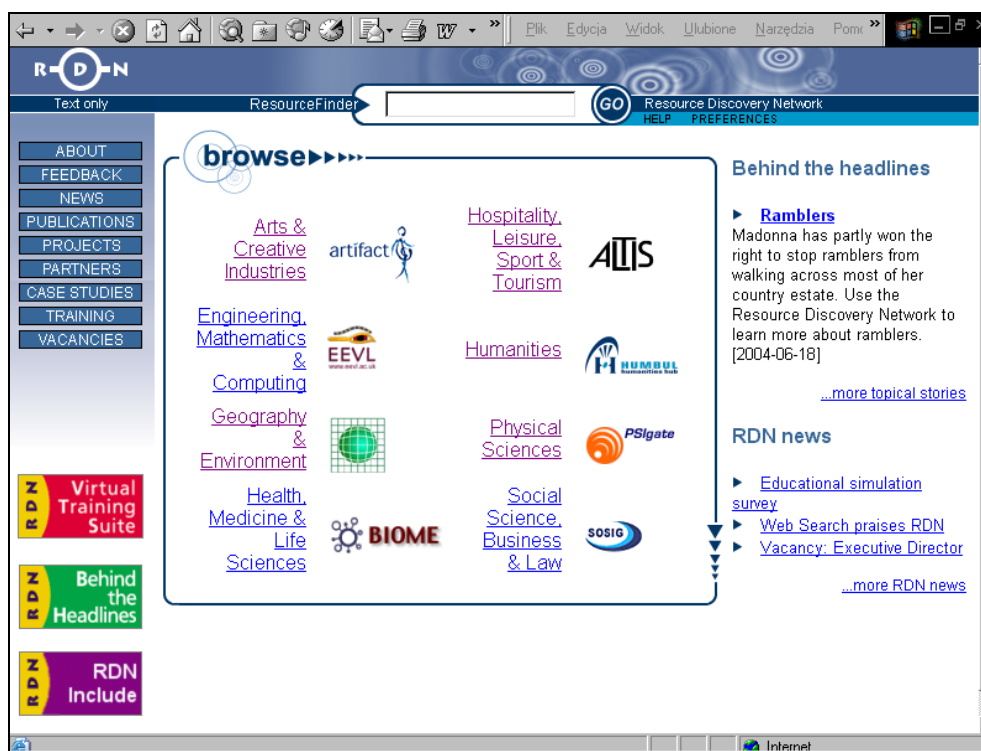
Serwis tworzony przez bibliotekarzy amerykańskich (przede wszystkim ze stanu Kalifornia) i finansowany głównie przez US Institute of Museum and Library Services. Rejestruje ponad 14.000

źródeł internetowych dla potrzeb użytkowników bibliotek publicznych, w układzie przedmiotowym. Wyszukiwanie ułatwia interfejs do zadawania pytań wg 9 różnych kryteriów formalnych i rzeczowych oraz przeglądania zasobów.

RDN Resource Discovery Network <http://www.rdn.ac.uk/>

RDN jest projektem wspólnym ponad 70 brytyjskich instytucji naukowych i edukacyjnych oraz Natural History Museum i British Library, tworzonym na podstawie *subject gateways* finansowanych przez JISC's (Joint Information Systems Committee) w ramach programu eLib. RDN jest siecią składającą się z centralnej instytucji – Resource Discovery Network Centre (RDNC) oraz kilku niezależnych serwisów zwanych hubami. Należy wspomnieć, że huby są tworzone również we współpracy wielu uniwersytetów. System RDN jest dostępny bezpłatnie dla wszystkich użytkowników sieci. Poprzez dostęp do 8 hubów umożliwia wyszukiwanie spośród ponad 30.000 dokumentów. Gromadzi on źródła starannie selekcjonowane, indeksowane i opisywane przez specjalistów z danej dziedziny. RDN ponadto wspiera współpracujące *subject gateways* w ich rozwoju, wypracowuje wspólne dokumenty i standardy, dba o promocję oraz szkolenia. Na witrynie RDN znaleźć można m.in., bardzo interesujący zestaw podręczników online na temat wyszukiwania informacji w Internecie, wg kilkudziesięciu dziedzin, dla studentów i pracowników szkół wyższych, zaopatrzone w quizy i testy [17]. Najważniejszą funkcją dostępną na stronie RDN jest wyszukiwanie jednocześnie we wszystkich współpracujących hubach (*subject gateways*) poprzez zadawanie pytań. W rezultacie otrzymujemy rekordy z opisem źródła i wskazaniem jego pochodzenia z danego serwisu. Nie ma w RDN możliwości przeglądania zasobów wg dziedzin. Wykonanie tej operacji jest możliwe wyłącznie poprzez wybór odpowiedniego *subject gateways* (rys. 4):

- artifact - Arts and Creative Industries <http://www.artifact.ac.uk/>
- EEVL - the Internet Guide to Engineering, Mathematics and Computing <http://www.eevl.ac.uk/>
- GEsorce - Geography and the Environment <http://www.gesource.ac.uk/home.html>
- BIOME - Health and Life Sciences <http://biome.ac.uk/> - składa się z 6 subject gateways: OMNI (Medical and Health Sciences), NMAP (Nursing, Midwifery and Allied Health Professions), VetGate (Animal Health and Veterinary Science), BioResearch (Biological and Biomedical Research), Natural Selection (The Natural World) oraz AgriFor (Agriculture, Food and Forestry)
- ALTIS - Hospitality, Leisure, Sport and Tourism <http://www.altis.ac.uk/>
- Humbul – Humanities Hub <http://www.humbul.ac.uk/>
- PSigate physical sciences (astronomy, chemistry, earth sciences, materials sciences, physics, and science history and policy) <http://www.psigate.ac.uk/newsite/>
- SOSIG – Social Science <http://www.sosig.ac.uk/>



Rys. 4. Witryna serwisu RDN Resource Discovery Network

DutchESS, Dutch Electronic Subject Service <http://www.kb.nl/dutchess/index.html>

Holenderski narodowy subject gateway ukierunkowany na środowisko akademickie, utworzony przez Bibliotekę Narodową (Koninklijke Bibliotheek) w 1998 r. W projekcie uczestniczy 7 bibliotek akademickich na podstawie porozumienia, które mogą podpisywać kolejne biblioteki. DutchESS uczestniczy i korzysta z rezultatów projektów: DESIRE, ROADS i Renardus. Źródła internetowe są dobierane i klasyfikowane wg Nederlandse Basisclassificatie przez specjalistów dziedzinowych z bibliotek współpracujących. W zasadzie nie ma żadnych ograniczeń co do języka czy kraju pochodzenia źródeł, ale szczególną dbałość przywiązuje się do kompletności w zakresie źródeł naukowych z Holandii. Wyszukiwanie w serwisie odbywa się poprzez zadawanie pytań (słowa z dokumentu, opisu, autor, URL, symbol klasyfikacji i jego słowny odpowiednik) albo przeglądanie struktury hierarchicznej wg w/w klasyfikacji. Ciekawą opcją jest anonowanie w odrębnym dziale nowych źródeł dodanych do serwisu w ciągu ostatniego tygodnia.

VASCODA (Niemcy) <http://www.vascoda.de/>

Portal zaprojektowany w celu integracji niemieckich bibliotek wirtualnych, *subject guides* również tworzonych przez biblioteki oraz czasopism elektronicznych (Elektronische Zeitschriftenbibliothek EZB) w ogólny i łatwy w użytkowaniu interdyscyplinarny system [14]. VASCODA różni się od innych serwisów tym, że w przyszłości zamierza stać się ogólnonarodową biblioteką cyfrową (Digitale Bibliothek Deutschland) udostępniającą informacje nie tylko z zasobów *subject gateways*, ale również dającą dostęp do licencjonowanych serwisów czasopism elektronicznych i baz danych dla uprawnionych uczestników projektu. Łącznie w projekcie finansowanym przez Deutsche Forschungsgemeinschafts uczestniczy ponad 30 instytucji naukowych, głównie bibliotek akademickich. Obecnie można wyszukiwać jednocześnie w 36 *subject guides*, spośród których wszystkie nadal pozostają niezależnymi serwisami i tworzone są indywidualnie przez poszczególne biblioteki albo konsorcja.

Australian Subject Gateways Forum (ASGF)

<http://www.nla.gov.au/initiatives/sg/gateways.html>

Utworzona w 2000 r. grupa zrzeszające instytucje tworzące *subject gateways* dla środowiska szkół wyższych w Australii. Celem ASGF jest ułatwienie wymiany informacji pomiędzy istniejącymi i powstającymi serwisami oraz rozwiązywanie wspólnych problemów. Działalność grupy jest koordynowana przez Bibliotekę Narodową Australii. Nie istnieje na razie wspólny interfejs wyszukiwawczy, ale pod w/w adresem wymienione są wszystkie uczestniczące serwisy, do których można uzyskać bezpośrednie połączenie.

Zakończenie

Podsumowując warto przypomnieć główne cechy *subject gateways*:

- zazwyczaj ograniczenie do jednej dziedziny wiedzy,
- rekordy bibliograficzne źródeł internetowych tworzone są manualnie i zawierają m.in. bogaty opis źródła z adekwatnymi informacjami,
- współkatalogowanie – rozproszona grupa specjalistów dziedzinowych zasila bazę danych,
- kryteria jakości – powinien istnieć oficjalny dokument wyjaśniający wg jakich kryteriów oceniane są źródła (najlepiej dostępny dla użytkowników)
- system klasyfikacji – wykorzystywany jako podstawa do przeglądania zasobów,
- stosowanie ogólnodostępnych standardów w celu umożliwienia współpracy z innymi serwisami, np. cross searching.

Jednym z wniosków nasuwających się na podstawie tej charakterystyki – niezbyt budującym – jest stwierdzenie, że nie mamy żadnego krajowego systemu tego typu, ani ogólnodziedzinowego, ani dotyczącego wybranej tematyki. Pierwsze *subject gateways* powstawały ze zwykłych list typu bookmark tworzonych przez bibliotekarzy i

przekształcanych potem w „polecane linki”. Tych ostatnich mamy na witrynach polskich bibliotek bardzo dużo – niektóre na wysokim poziomie kompletności. Warto byłoby na wzór przedstawionych tu przykładów połączyć wysiłki i stworzyć prawdziwe *subject gateways*. Zwłaszcza że kraje bardziej zaawansowane w tej dziedzinie wypracowały metody i środki techniczne, które są w większości bezpłatnie dostępne w sieci. Poza tym można dołączyć do wielu projektów międzynarodowych. Jak zwykle dyskusyjna pozostaje kwestia finansowania. W wielu krajach czynią to odpowiednie ministerstwa lub agencje rządowe. Jednak najważniejsza jest inicjatywa, która wzorem dobrych przykładów (NUKAT, Sympo, BazTech) może pochodzić z jakiegokolwiek biblioteki lub grupy bibliotek. Mamy dobre doświadczenia we współpracy i należy je wykorzystać w kolejnym wyzwaniu dla bibliotekarzy. W tym miejscu należy przyznać rację M. Nahotko, który stwierdził, że „[...] lawinowo rosnąca ilość publikacji internetowych i wszelkich tzw. Informacji „wirtualnych” otwiera nowe i ciekawe możliwości pracy dla bibliotekarzy. Ktoś bowiem musi zapanować nad chaosem, który obecnie cechuje Internet i twierdzą, że będą to bibliotekarze, którzy dysponują prawie gotowymi narzędziami [...]. Zmieni się też pewnie kąt widzenia bibliotekarzy – z dbałości o własne, lokalne zasoby na pracę w ogólnościowej, wirtualnej „bibliotece” zasobów sieciowych” [12].

Bibliografia

1. Bawden D., Robinson L. Internet subject gateways revised. *International Journal of Information Management*, 2002 Nr. 22, s. 157-162.
2. Bergman M. The Deep web: surfacing hidden value. *The Journal of Electronic Publishing* [on-line]. August, 2001 Volume 7, Issue 1 [dostęp 21 czerwca 2004]. Dostępny w World Wide Web: <http://www.press.umich.edu/jep/07-01/bergman.html>.
3. Brümmer Anna, Subject based information gateways [on-line]. December 1997, updated February 2000 [dostęp 21 czerwca 2004]. Dostępny w World Wide Web: <http://www.lub.lu.se/desire/sbig.html>.
4. Czapnik G.: Bibliotekarskie katalogi tematyczne. *Zagadnienia Informacji Naukowej* nr 1 (81), 2003, s. 64-77.
5. Franco A. Gateways to the Internet: Finding quality information on the Internet. *Library Trends*, 2003 Vol. 52, nr 2, s. 228-246.
6. Gawroński, A., Dynkowski, M.: Biblioteczne portale albo wirtualne warsztaty informacyjne. *EBIB* [on-line] 7/2000 (15) [dostęp 21 czerwca 2004]. Dostępny w World Wide Web: <http://ebib.oss.wroc.pl/arc/e015-02.html>.
7. Kirriemuir J., Welsh S., Knight J. Cross-Searching Subject Gateways. The Query Routing and Forward Knowledge Approach. *D-Lib Magazine* [on-line], January 1998. [dostęp 21 czerwca 2004]. Dostępny w World Wide Web: <http://www.dlib.org/dlib/january98/01kirriemuir.html>.
8. Koch T. Quality-controlled subject gateways: definitions, typologies, empirical overview. Manuscript of the article published in „*Online Information Review*”, Vol. 24/1, 2000. [dostęp 21 czerwca 2004]. Dostępny w World Wide Web:..
9. Malesza E. Ogólna charakterystyka stron WWW polskich bibliotek. *EBIB* [on-line] 1999:07 [listopad]. [dostęp 21 czerwca 2004]. Dostępny w World Wide Web: <http://ebib.oss.wroc.pl/arc/e007-02.html>.
10. Nahotko M.: Metadane. *EBIB* [on-line] 6/2000 (14) [dostęp 21 czerwca 2004]. Dostępny w World Wide Web: <http://ebib.oss.wroc.pl/arc/e014-02.html>.
11. Nahotko M.: Metadane jako narzędzie opracowania elektronicznych zasobów sieciowych. *Zagadnienia Informacji Naukowej* nr 2 (82), 2003, s. 101-117.
12. Nahotko M. RDF jako narzędzie uporządkowania zasobów internetowych. *PTIN* 2002 nr 2, s. 3-6.
13. Neuroth, H., Koch T. Cross-browsing and cross-searching in a distributed network of subject gateways: Architecture, data model, and classification. Renardus - The Academic Subject Gateway Service in Europe. [on-line] 19 May 2001 [dostęp 21 czerwca 2004]. Dostępny w World Wide Web: http://renardus.sub.uni-goettingen.de/conferences/elag2001/neuroth_koch_elag2001.html.
14. Pianos T. Vascoda – a portal for scientific resource collection created by German libraries and information centres. W: 65th IFLA Council and General Conference, Bangkok, Thailand [on-line] August 20 - August 28, 1999 [dostęp 21 czerwca 2004]. Dostępny w World Wide Web: <http://www.ifla.org/IV/ifla69/papers/055e-Pianos.pdf>.
15. Place E. International Collaboration on Internet Subject Gateways. W: 65th IFLA Council and General Conference, Bangkok, Thailand [on-line] August 20 - August 28, 1999 [dostęp 21 czerwca 2004]. Dostępny w World Wide Web: <http://www.ifla.org/IV/ifla65/papers/009-143e.htm>.

16. Sosińska-Kalata B. Struktury klasyfikacyjne w organizacji zasobów informacyjnych Internetu. W: III Krajowa Konferencja pt. Multimedialne i Sieciowe Systemy Informacyjne [on-line] 19-20 września 2002 r. [dostęp 21 czerwca 2004]. Dostępny w World Wide Web: <http://www.zsi.pwr.wroc.pl/zsi/missi2002/pdf/s403.pdf>.
17. Virtual Training Suite by RDN. [on-line] Page last updated: Monday March 08, 2004 [dostęp 21 czerwca 2004]. Dostępny w World Wide Web: <http://www.vts.rdn.ac.uk/>.
18. Worldsford E. Subject gateways: fulfilling the DESIRE for knowledge. *Computer Networks and ISDN Systems*, 1998, nr 30, s. 1479-1489.

Serwisy tematyczne o kontrolowanej jakości w Internecie - subject gateways / Lidia Derfert-Wolf// W: Biuletyn EBIB [Dokument elektroniczny] / red. naczelny Bożena Bednarek-Michalska. - Nr 6/2004 (57) czerwiec/lipiec. - Czasopismo elektroniczne. - [Warszawa] : Stowarzyszenie Bibliotekarzy Polskich KWE, 2004. - Tryb dostępu: <http://ebib.oss.wroc.pl/2004/57/derfert.php>. - Tyt. z pierwszego ekranu. - ISSN 1507-7187