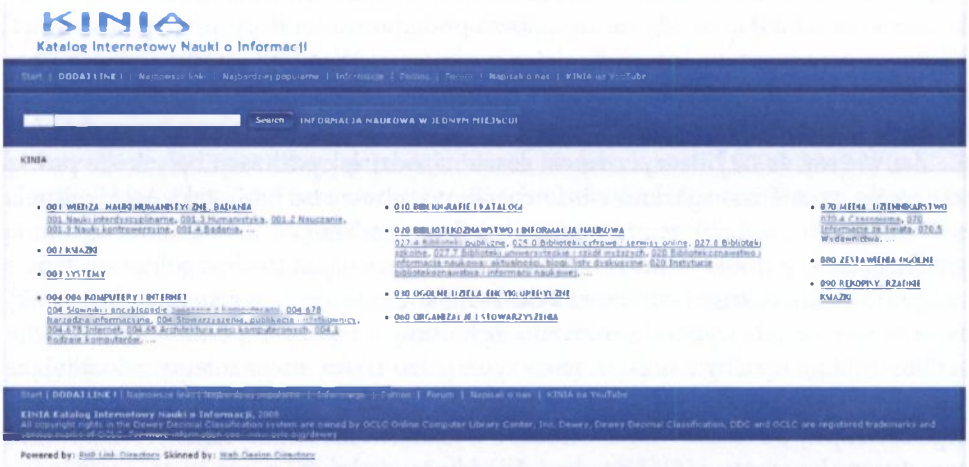
Jarosław Pacek, Sebastian D. Kotuła, Piotr G. Zieliński

PROJEKT, ZAŁOŻENIA I REALIZACJA INTERNETOWEGO KATALOGU NAUKI O INFORMACJI (KINIA) – HTTP://KINIA.CZYTELNIA.NET/



Fot. 1. Strona główna katalogu KINIA

Przyczyna utworzenia katalogu

Coraz większa dostępność do internetowego medium[1] i stosunkowa łatwość w budowaniu zasobów (dziś oprogramowanie do tworzenia stron WWW działa na zasadzie

[1] „Internet stanowi zbiór wszystkich narzędzi technicznych i programowych, które są wykorzystywane do przesyłania informacji między komputerami – krótko mówiąc jest to jedna z form łączności elektronicznej. WWW, Web, Sieć tworzą natomiast połączone ze sobą w skali świata – Internetem – różnej wielkości lokalne sieci (wraz z zasobami informacyjnymi i usługami wchodzących w ich skład komputerów)" [Gogołek, 2006: 18].

„upuść i wklej") powoduje, że po pierwsze, dla wszystkich aktywnych poszukiwaczy informacji, czyli dla tych o sprecyzowanych potrzebach informacyjnych Internet jest źródłem pierwszym i po drugie, zasobów jego stale przybywa[2]. W gąszczu informacji wyrażanych poprzez strony WWW, rozproszonych po licznych serwerach, niekiedy trudno się poruszać. Stąd wiele osób określa sieć jako wielki śmietnik. Dzieje się tak z kilku przyczyn. Do najczęstszych można zaliczyć nieumiejętność budowania prawidłowych kwerend, nieznajomość opcji zaawansowanego wyszukiwania i wreszcie nieznajomość architektury cyberprzestrzeni. Ta ostatnia może wyrażać się poprzez brak umiejętności korzystania z serwisów społecznościowych lub w ogóle ich ignorowania jako źródeł rzetelnej informacji. Pamiętajmy, że dla niektórych badaczy, jak np. Jarona Laniera współtworzenie treści (tzw. mądrość tłumu) zbytnio odznacza się kolektywizmem, znanym z pewnych utopii społecznych, jak np. komunizm [Basara, 2007: 15], co ostatecznie prowadzi niektórych do omijania internetowych skupisk wiedzy. Choć z powodzeniem można przyjąć za Małgorzatą Biedroń [2006: 245], że „mądrość to nie jest zbiór nagromadzonych przez nas informacji, ale zdolność do korzystania z nich do wydawania sądów i podejmowania decyzji w oparciu o nie". W dobie paradygmatu Internetu drugiej generacji Web 2.0[3] istotne jest nie potępienie amatorów za współtworzenie treści, lecz w miarę możliwości jak najszersze wykorzystywanie tych wartościowych, w tym tworzonych także przez amatorów.

Zauważmy, że na historycznej osi czasu narodziny cywilizacji były ściśle powiązane z tworzeniem magazynów informacji, począwszy od biblioteki Asurbanipala w Niniwie. Idea indeksowania, opisywania i gromadzenia informacji ma ogromną tradycję. Dziś, w dobie społeczeństwa informacyjnego, jest to szczególnie widoczne, czego dowodzą licznie tworzone i udoskonalane internetowe szperacze i wyszukiwarki. Obecnie, gdy wyszukiwanie stało się kulturą, o czym przypomina John Battelle [2006: 9–11], i np. aby z czasem stworzyć bardzo trafne mechanizmy odnajdujące pożądane treści, trzeba zaczynać od podstaw. „Podróż do wyszukiwania doskonałego zacznijmy od raczkowania. Chodzi o wyszukiwanie tematyczne. Wyszukiwarki tematyczne koncentrują się na jednej dziedzinie wiedzy, a z powodu ograniczonego korpusu informacji i jasnych związków między pojęciami zapewniają poszukiwaczom bardzo trafne wyniki" [Battelle, 2006: 202].

Jednym ze sposobów organizowania wiedzy w środowisku cyfrowym są między innymi katalogi typu *subject gateway*, które Adrienne Franco [2003: 228–229] określa także angielskimi terminami: *megasites, metasites, portals, gateways to the internet, web bibliographies, webliographies*. Jerry V. Caswell [2004: 73] gateway'e charakteryzuje jako: zbiór powiązanych narzędzi, które pozwalają użytkownikom identyfikować i lokalizować materiały relewantne do ich potrzeb. Przy czym ważne

[2] Włodzimierz Gogołek w swojej pracy [2006: 20] graficznie przedstawił stały przyrost cyfrowych zasobów informacyjnych.

[3] Więcej na ten temat zob.: S. Kotuła, 2007.

jest, że gateway'e zrywają z utrwalaną przez biblioteczne katalogi linearnością na rzecz nielinearnej architektury Internetu.

Prezentowanym katalogiem próbuje się zlikwidować dwie swoiste bariery informacyjne[4]. Pierwszą jest nadmierne rozproszenie informacji, uniemożliwiające szybkie dotarcie do tej relewantnej, czy ściślej, pertynentnej. Katalog poświęcony jest szeroko rozumianej nauce o informacji (informacji naukowej). Zatem znajdują się w nim opisy dotyczące właśnie tej dziedziny, a umieszczone w polskojęzycznej przestrzeni Internetu. Drugą barierą informacyjną jest próba zdjęcia z sieci etykiety „nieprawdziwe" – wielu użytkowników w ogóle nie uznaje informacji ze świata WWW za prawdziwe ze względu na to, że każdy do sieci może wrzucić cokolwiek[5].

KINIA sprzyja badaniom jakościowym danych stron. Strona umieszczana w katalogu uzyskuje automatycznie wyższy status (jest ważna, ponieważ znalazła się w katalogu stron o kontrolowanej jakości). Przecież jednym ze sposobów sprawdzania wiarygodności danej strony internetowej jest sprawdzenie, jakie serwisy daną stronę linkują[6]. Autorami nadzorującymi katalog są pracownicy naukowi Uniwersytetu, a dane na ich temat są łatwo dostępne z poziomu strony głównej. Jest jasno określona odpowiedzialność za zaakceptowany do zasobów link.

Postęp i rozwój dookolnej rzeczywistości, szybkie zmiany, widoczne zwłaszcza w każdorocznym przyroście informacji wyrażanej w bitach, dziś dzieją się zwłaszcza w Internecie. W 2004 roku, jak podaje Marek Hendrykowski [2005: 38–39] istniało 5 miliardów stron internetowych, a treść ich wszystkich liczona jest na około 100 terabajtów (1 TB = 1024 gigabajtów). Wydawałoby się, że świat wirtualny umknął zupełnie wielu dziedzinom. Jak się jednak okazuje, wypracowanie wspólnej platformy jest zupełnie możliwe. Omawianym projektem próbuje się udowodnić, że można połączyć pragmatyzm Internetu (zawrotna szybkość zmian) z naukową akrybią (cały katalog opiera się na naukowo opracowanej Klasyfikacji Dziesiętnej Deweya).

Wybór metody (dlaczego KDD)

Cytowany wcześniej John Battelle [2006: 30] w kwestii KDD podał, że dziesiętny system Melville'a Deweya „nie sprostał wymaganiom gigantycznej sieci WWW", co oznacza de facto niemożność na jej podstawie zindeksowania zasobów cyfrowego uniwersum. Jednakże, jak pokazuje praktyka, katalogi typu *subject gateway* w zupełności spisują się, gdy z całej przestrzeni świata WWW wybierze się tylko te strony, które treścią odpowiadają danej dziedzinie. Wybór KDD podyktowany był właśnie tym faktem. Wzorem katalogów anglojęzycznych, np. BUBL (http://bubl.

[4] Pojęcie *bariera terminologiczna* stosujemy za: Marzeną Świgoń [2006: 32].

[5] Więcej krytyki na temat tego zjawiska czytelnik odnajdzie w pracy Andrew Keena [2007: 61–65].

[6] Na podstawie: B. Bednarek-Michalska, *Ocena jakości bibliotekarskich serwisów informacyjnych udostępnianych w Internecie*, EBIB, nr 31, 2002, dostęp: http://ebib.oss.wroc.pl/2002/31/michalska.php.

ac.uk/), zdecydowano, że będzie to KDD. Dla potrzeb nauki o informacji wystarczy pierwsza kategoria główna głęboka na trzy podklasy, toteż i zapotrzebowanie na ten fragment KDD zostało zgłoszone OCLC.

Online Computer Library Center

Sam wybór metody stanowił jedno zadanie, odrębnym problemem stała się faktyczna możliwość wykorzystania wybranej klasyfikacji. Jak wiadomo, KDD jest w posiadaniu OCLC. Nauczeni doświadczeniem innych serwisów, wykorzystujących KDD jako podstawę podziału wiedzy cyberświata i faktem odnotowywania przez serwisy na stronie stosownej notatki o treści: „All copyright rights in the Dewey Decimal Classification system are owned by OCLC Online Computer Library Center, Inc. Dewey, Dewey Decimal Classification, DDC and OCLC are registered trademarks and service marks of OCLC. For more information see: www.oclc.org/dewey", co jasno sugeruje, że serwisy te musiały zgodę od OCLC uzyskać. Wystąpiliśmy z oficjalnym pismem do OCLC z zapytaniem o możliwość wykorzystywania owej klasyfikacji. Warto odnotować, że OCLC zareagowała na ten fakt z wielkim entuzjazmem. W korespondencji, jaka wywiązała się między twórcami katalogu a OCLC precyzowaliśmy, o jaki typ działalności chodzi (*non profit*), o klasyfikowanie czego (polskojęzyczna przestrzeń Internetu), jaka idea przyświeca naszej aktywności (udostępnić wszystkim zainteresowanym swobodny dostęp do wartościowych miejsc w sieci). Koniec końców zgodnie z oczekiwaniami i ostateczną aprobującą opinią prawników OCLC 6 grudnia 2006 roku uzyskaliśmy imienną zgodę dla osób: Grzegorz Gmiterek, Sebastian Kotuła, Jarosław Pacek, Piotr Zieliński na wykorzystanie KDD w takim stopniu: pierwszej kategorii głównej OOO wraz z trzema jej podklasami.

W jaki sposób katalog jest wykonany (technologia, domena itp.)

Katalog został uruchomiony przy wykorzystaniu specjalnego skryptu PhpLinkDirectory. Umożliwia on przechowywanie nowo wpisanych rekordów w buforze, w którym oczekują one na ewentualną korektę i zatwierdzenie. Katalog umożliwia również kontrolę aktualności podanych adresów, rejestrowanie użytkowników, automatyczną korespondencję. Przy każdym umieszczonym w katalogu rekordzie, po jego lewej stronie, widoczny jest tzw. wskaźnik PageRank, symbolizowany literami PR wraz z liczbową wartością i symbolicznym zielonym paskiem poniżej. Wskaźnik ten ustalany jest przez algorytmy Google i odnosi się do jakości i popularności docelowej witryny. Serwis wykorzystuje płatne usługi hostingowe, niestety, nie posiada jeszcze żadnego źródła finansowania poza środkami własnymi twórców. Jego adres URL jest darmową domeną, przyznaną przez usługodawcę, w pakiecie wraz z miejscem serwerowym. Jedną z najważniejszych i zakładanych od samego początku cech katalogu była

jego otwartość. W zamierzeniu ma być on miejscem przyjaznym, dla wszystkich użytkowników, zarówno chcących zamieścić tu adres (a przez to w pewien sposób zareklamować swój serwis, swoją instytucję, przedstawić ją zainteresowanej tematyką społeczności), jak i dla użytkowników korzystających z zasobów. Każdy może dodawać linki do katalogu, potrzebne opisy i dodatkowe informacje. Również dalsze wykorzystanie, polegające np. na kopiowaniu, przetwarzaniu zawartych w katalogu rekordów, jest całkowicie dozwolone. Twórcom katalogu zależy, aby takie poczucie przyjazności i łatwości wykorzystania towarzyszyło odwiedzającym serwis.

Budowa katalogu (elementy składowe)

Każdy rekord katalogu składa się z pięciu, elementów: nazwa, czyli Tytuł opisywanego zasobu, który jest jednocześnie hiperłączem uruchamiającym w nowym oknie docelową witrynę. Opis, jak wspomniano wcześniej, zawiera krótkie, podstawowe informacje charakteryzujące dany zasób, w postaci kilku zdań. Opis będący składnikiem wyświetlanego rekordu zawiera dodatkowo jeszcze dwie części. Informację o Odpowiedzialności (element ten dla podniesienia czytelności wyróżniany jest czcionką wytłuszczoną i pochyłą), czyli dane takie jak nazwa osoby czy instytucji odpowiedzialnej za powstanie, aktualizację i treści zawarte w danym serwisie. Następny element to Słowa kluczowe charakteryzujące treść (wyróżnione w podobny sposób jak Opis). Dodanie tych dwóch elementów użytkownik może pozostawić redaktorom serwisu. Jednakże możliwość ich konstrukcji przez samych użytkowników może dostarczać materiału szczególnie interesującego dla obserwacji naukowych. Wpisuje się to zjawisko w obszar znaczeniowy pojęcia „folksonomii". Zwłaszcza słowa kluczowe stanowią interesujący badawczo element. W dalszej perspektywie, w miarę napływu materiału do katalogu, mogą posłużyć jako podstawa do budowy specjalnego źródłosłowu z zakresu nauki o informacji, odzwierciedlającego specyfikę wniesioną do tej dziedziny poprzez rozwój nowoczesnych technologii informacyjnych. Istnieje taka potrzeba, oficjalne zasoby słownictwa z zakresu wymienionych dyscyplin pochodzą sprzed kilku lat. Pojedynczy rekord katalogu kończy (wyróżniony zielonym kolorem) identyfikator – adres sieciowy opisywanego obiektu. Zastosowany do obsługi skrypt PHP pozwala na dołączenie do rekordu wskaźnika tzw. PageRank. W wykorzystaniu katalogu jako źródła informacji może pomóc również możliwość subskrypcji wybranych kategorii za pomocą kanałów RSS (Really Simple Syndication).

Jak dodawać opis

Dodawanie zasobów do katalogu odbywa się poprzez wypełnienie prostego formularza złożonego z sześciu pól: Tytuł (np. Biblioteka Główna UMCS w Lublinie), URL (adres internetowy, np. http://www.bg.umcs.lublin.pl), Opis (krótki, kilkuzdaniowy

opis zawartości strony, można podać informacje o twórcach i sformułować słowa kluczowe charakteryzujące treść strony), Twoje imię (imię, nazwisko, pseudonim osoby dodającej link), Twój e-mail (adres e-mail osoby dodającej link, którego podanie usprawnia redaktorom katalogu ewentualną wymianę informacji dotyczących dodawanego serwisu), Kategoria (należy wybrać najbardziej odpowiednią spośród dostępnych kategorii).

Idea Web 2.0, czyli katalog współtworzą wszyscy

Problem dwuzerowości wkroczył również do świata bibliotek. Library 2.0, fora internetowe i blogi oplecione dookoła problematyki bibliotek i informacji naukowej spowodowały duże zainteresowanie w środowisku. Mimo aktywności bibliotekarzy w serwisach Web 2.0 (de.licio.us, digg.com) i wpisywaniu się ich tym samym w ideę otwartości, potrzebny stał się również specjalistyczny katalog, który by zebrał i uporządkował polskie sieciowe zasoby z dziedzin powiązanych z informacją naukową. KINIA jest próbą połączenia tego, do czego wszyscy bibliotekarze są przyzwyczajeni, czyli porządku z pewnego rodzaju bałaganem, cechującym przecież świat serwisów dwuzerowych. Łatwość dodawania odnośników hipertekstowych i możliwość promocji własnego miejsca pracy winna zdopingować studentów, bibliotekarzy, pracowników instytutów bibliotekoznawstwa, a także tych zatrudnionych w ośrodkach informacji naukowej do współtworzenia katalogu. Niestety, liczba linków dodanych przez wyżej wymienione grupy zainteresowane rozwojem katalogu jest wciąż mniejsza niż liczba dodanych do niego przez jego twórców. Poza tym część z dodanych odnośników musi być przeredagowana z racji częstego braku dwóch elementów opisu: *description* i *keyword*, a także z powodu bezpośredniego kopiowania zawartości stron do opisu.

Można wnioskować, że:

- KINIA jako młody katalog jest jeszcze stosunkowo nieznany w środowisku zawodowo związanym z bibliotekarstwem;
- środowisko zawodowo związane z bibliotekarstwem żyje ciągle w świecie odciętym od sieci albo;
- funkcjonuje jeszcze na zasadach wprowadzonych w Internecie przez tzw. Web 1.0.

Dwie ostatnie przyczyny pozostają poza naszym zasięgiem, natomiast pierwszą pokonamy – mamy taką nadzieję – wraz z upływem czasu i współpracując z zainteresowanymi współtworzeniem KINIA.

Perspektywy

Możliwości rozwoju serwisu KINIA mogą wiązać się z rozbudową projektu o nowe elementy i utworzeniem docelowo dużej platformy informacyjnej z zakresu nauki o informacji. Przewidywane jest np. utworzenie słownika lub nawet większej encyklopedii, o charakterze otwartym, działającej w systemie Wiki. Tego rodzaju podręcznych źródeł *online*, praktycznie w polskim Internecie nie ma, a doświadczenie i obserwacja podpowiada, że jest na nie spore zapotrzebowanie, choćby ze strony studentów kierunku informacja naukowa. Istniejący w ramach encyklopedii Wikipedia mniejszy portal specjalistyczny Bibliologia, nie spełnia oczekiwań, jest zbyt ubogi w informację, co wynika z nikłej świadomości jego istnienia w środowisku potencjalnie zainteresowanych. Można liczyć na fakt, że poprzez powiązanie podobnego projektu z coraz lepiej znanym KINIA użytkownicy będą częściej docierać do takiej nowej encyklopedii, wykorzystywać ją i, miejmy nadzieję, współtworzyć. Będzie to choćby częściowa próba uzupełnienia pewnej luki w opracowaniach terminologicznych z zakresu nauki o informacji, zwłaszcza aktualnych, dotyczących sfery technologii informacyjnych, oryginalna poprzez jej otwartość na współtwórców.

Bibliografia

Basara Z., *Czytelnicy do klawiatur*, „Gazeta Wyborcza", 2 marca 2007, s. 15.

Battelle J., *Szukaj. Jak Google i konkurencja wywołali biznesową i kulturową rewolucję*, PWN, Warszawa 2006.

Bednarek-Michalska B., *Ocena jakości bibliotekarskich serwisów informacyjnych udostępnianych w Internecie*, EBIB 2002, nr 31, dostęp: http://ebib.oss.wroc.pl/2002/31/michalska.php.

Biedroń, M., *Homo eligens w globalnej wiosce*, [w:] *Definiowanie McLuhana. Media a perspektywy rozwoju rzeczywistości wirtualnej*, pod red. M. Sokołowskiego, Wyd. ALGRAF, Olsztyn 2006, s. 243–248.

Caswell J. V., *A conceptual frameworks for gateways*, „Information technology and libraries", June 2004, s. 73–81.

Franco A., *Gateways to the Internet: Finding Quality Information on the Internet*, „Library Trends", vol. 52, 2003, nr 2, Fall, s. 228–246.

Gogołek W., *Technologie informacyjne mediów*, ASPRA, Warszawa 2006.

Hendrykowski M., *Metafory Internetu*, Wydawnictwo Naukowe UAM, Poznań 2005.

Keen A., *Kult amatora. Jak internet niszczy kulturę*, WaiP, Warszawa 2007.

Kotuła S., *WEB 2.0 – współczesny paradygmat Internetu*, [w:] *Oblicza Internetu. Architektura komunikacyjna Sieci*, pod red. Marka Sokołowskiego, Wyd. PWSZ, Elbląg 2007, s. 181–188.

Świgoń M., *Bariery informacyjne. Podstawy teoretyczne i próba badań w środowisku naukowym*, SBP, Warszawa 2006.