

2006-05-25

Sabina Cisek, Remigiusz Sapa
Instytut Informatyki i Bibliotekoznawstwa
Uniwersytet Jagielloński

Komunikacja naukowa w Internecie – mity i rzeczywistość

Abstrakt

Artykuł poświęcony jest wybranym nieporozumieniom na temat Internetu i jego funkcji w informacji i komunikacji naukowej. Opisanie w tekście mity dotyczą bezpłatnego dostępu do zasobów naukowych, publikowania w WWW oraz wyszukiwania informacji w Sieci. Wskazano ich źródła oraz skutki. Zwrócono uwagę na rolę informacji naukowej i bibliotekoznawstwa w rozwiązywaniu problemów z nimi związanych.

Abstract

The paper depicts some chosen misconceptions about the Internet and its function in the scholarly information and communication processes. The myths described are connected with the open access to scholarly resources, scientific publishing on the Web and information retrieval and searching. The implications of those myths are indicated. The role of information and library science and profession in coping with the problem is stressed.

Wstęp

Internet jako nowe i dynamiczne środowisko informacyjne człowieka, z jednej strony z pewnością otwiera nowe możliwości przed komunikacją naukową, ale jednocześnie stawia bardzo poważne wyzwania wynikające z konieczności opanowania umiejętności korzystania z jego zasobów i usług, redefiniowania relacji między różnymi uczestnikami procesu komunikacji oraz zbadania nowych zjawisk i określenie ich konsekwencji dla praktyki informacji naukowej i bibliotekoznawstwa. Mimo upływu lat nadal w powszechnej opinii funkcjonuje wiele mitów i nieporozumień na temat wpływu Internetu na komunikację naukową i kierunki jej rozwoju. Częściowo wynikają one z nadmiernego upraszczania samej komunikacji naukowej i postrzegania jej w kategoriach jednego uniwersalnego modelu, utrzymujących się problemów i niejasności terminologicznych czy wreszcie przeceniania wpływu technologii na zachowania naukowców. Co więcej, publikowanie elektroniczne, mimo oczywistych postępów, nadal jest w stanie ciągłych eksperymentów i poszukiwania optymalnych rozwiązań [Cronin 2004, dok. elektr.]. Choć dysponujemy już dzisiaj wynikami wielu badań i doświadczeniami wynikającymi z wdrażania nowych koncepcji w zakresie komunikacji naukowej, nadal nasza wiedza jest niewystarczająca do pewnego prognozowania kierunków jej rozwoju i optymalizacji usług informacyjnych i bibliotecznych oferowanych naukowcom.

Celem autorów nie jest przedstawienie kompletnego katalogu nieporozumień na temat Internetu i jego roli w komunikacji naukowej, ale zwrócenie uwagi na kilka szczególnie popularnych błędnych przekonań na ten temat, które stanowią, oprócz wielu innych, przedmiot badań podstawowych i stosowanych w zakresie informacji naukowej i bibliotekoznawstwa.

Wszystko za darmo dla wszystkich?

Kiedy prawie 20 lat temu Internet zaczął podbijać świat nauki, odczuwalne były już pierwsze konsekwencje tak zwanego „kryzysu czasopism”. Stały wzrost cen periodyków naukowych oraz kosztów ich gromadzenia, opracowywania, przechowywania i udostępniania w bibliotekach groził poważnym ograniczeniem lub przynajmniej utrudnieniem dostępu do literatury naukowej. W polskich warunkach problem ten był dodatkowo potęgowany przez trudną sytuację ekonomiczną i niski kurs złotówki w stosunku do dolara, a w początkowym okresie także przez bariery polityczne. Rezygnowanie z prenumeraty zbyt drogiego czasopism przez niektórych nabywców spowodowało jeszcze szybszy wzrost ich cen, gdyż wydawcy musieli w jakiś sposób kompensować straty wynikające ze spadającej sprzedaży. Z drugiej strony komercjalizacja nauki, wprowadzanie mechanizmów rynkowych lub quasi rynkowych do zarządzania nauką i związana z nimi presja na efektywność badań i wydajność samych naukowców połączona ze wzmocnieniem oceny ich pracy, oznaczały w konsekwencji wzrost zainteresowania liczbą i jakością publikacji. Nikt przecież nie rozlicza naukowców z liczby przeczytanych artykułów i książek. Ich zatrudnienie, kariera i sława zależą bezpośrednio nie od tego, ile przeczytają, ale ile napiszą. Jak dowodzą niektóre badania, w porównaniu do swoich poprzedników, dzisiejsi naukowcy mniej czytają, a więcej publikują [Björk, Turk 2000, dok. elektr.]. Obie podstawowe role naukowca stały w wyraźnej opozycji. Naukowiec-autor chciał jak najwięcej publikować, a naukowcowi-czytelnikowi zaczynało brakować pieniędzy na zakup tych publikacji. Samonakręcająca się spirala negatywnych zjawisk musiała rodzić frustrację i przecucie radykalnych zmian. Pojawienie się w tej sytuacji Internetu, środowiska na początku tyleż nieznanego, co niezwykle obiecującego dla komunikacji naukowej, zadziało jak katalizator. Wielu zaczęło wieszczyc radykalne zmiany. Tak jak radio miało spowodować całkowity upadek prasy, upowszechnienie się telewizji wyeliminować kino, Internet miał radykalnie zrewolucjonizować komunikację naukową. Dodatkowo, Internet w swoim pionierskim okresie stanowił ziemię niczyją, budowany był oddolnie i spontanicznie. Nikt nie kontrolował komunikacji naukowej w tym środowisku (tak, jak kontrolowana była przez wydawców w rzeczywistości realnej). Nic zatem dziwnego, że w takiej sytuacji wśród wielu różnych koncepcji reformatorskich nabrał mocy jeden z najważniejszych i chyba najczęściej powtarzanych postulatów – żądanie wykorzystania Internetu do zapewnienia darmowego, powszechnego dostępu do wyników badań! Nie trzeba dodawać, że hasło darmowego dostępu do literatury naukowej musiało spotkać się ze szczególnie przychylnym przyjęciem w tych społecznościach, które przede wszystkim „importowały” osiągnięcia naukowe i jednocześnie dysponowały bardzo ograniczonymi środkami na ich zakup. Okres pionierski Internetu był z pewnością okresem entuzjazmu, w którym znalazły ujście nabrzmiałe problemy, oczekiwania i marzenia, jak się jednak miało okazać często nazbyt idealistyczne lub po prostu utopijne.

Pojawiły się czasopisma elektroniczne. Szybko jednak okazały się dwie rzeczy: ktoś płacić musi, a rzeczywistość nie znosi próżni. Mniej więcej w roku 1996 wielcy wydawcy zdecydowali się „wejść” do Internetu [Mogge 1999, dok. elektr.] i przenieść do tego środowiska nieco zmodyfikowany system komunikacji naukowej, w którym grali kluczową rolę i który oznaczał koniec mrzonek o darmowych czasopismach. A to oznaczało pryśnięcie nadmiernie wybujałych marzeń i nieuchronną w takich sytuacjach frustrację. Całkowicie darmowe (dla czytelników) powszechnie dostępne czasopisma elektroniczne nie wyszły poza margines komunikacji naukowej. Zresztą nadal pozostają na marginesie i nie stanowią głównego nurtu komunikacji naukowców [Björk 2005, dok. elektr.] [Harnad i in. 2004, dok. elektr.]. Mimo, iż już wiele lat wcześniej wydawało się, że dalszego wzrostu cen czasopism rynek nie wytrzyma, czasopisma drożały nadal, a środowiska naukowe płaciły coraz więcej.

Co gorsza, dostępność wielu tytułów w postaci elektronicznej i drukowanej jednocześnie musiała przyczynić się do wzrostu kosztów po stronie odbiorcy.

Kiedy nie zdarzył się cud i minęła związana z tym frustracja, przyszedł czas na pragmatyzm. Jego szczególnym wyrazem jest tak zwana „zielona droga” do otwartego, darmowego dostępu do publikacji naukowych, uznana za jedną z właściwych dróg realizacji idei *Open Access*. Oznacza ona archiwizowanie i darmowe udostępnianie w Internecie artykułów wcześniej opublikowanych w czasopiśmie dystrybuowanych lub udostępnianych w Internecie na zasadach komercyjnych. Istnieją przynajmniej trzy podstawowe kanały takiej formy udostępniania wyników badań: archiwa e-printów (zazwyczaj oddzielne dla różnych dyscyplin), repozytoria instytucji (np. uczelni) oraz indywidualne strony WWW naukowców. Jeśli obecnie jedynie nieliczne spośród około 24 000 recenzowanych czasopism naukowych wydawanych na świecie są dostępne w Internecie całkowicie za darmo dla wszystkich zainteresowanych (około 5%), to zdecydowana większość redakcji (aż 92%) zgadza się na różnych warunkach na darmowe upublicznienie w Internecie artykułów opublikowanych wcześniej na ich łamach [Harnad 2005, dok. elektr.]. Ta droga, choć nie rewolucyjna i nie radykalna, zaczyna już dzisiaj przynosić wymierne efekty. Jeśli interesujący nas artykuł jest poza naszym zasięgiem, bo nasza biblioteka nie prenumeruje danego czasopisma, a dostęp do wersji elektronicznej jest płatny, nie należy od razu się poddawać. Szanse znalezienia pełnego tekstu takiego artykułu dostępnego za darmo w Internecie (lub jego wersji), oczywiście przy wykorzystaniu wyspecjalizowanych narzędzi, są dzisiaj z pewnością znacznie większe niż 5 lub 10 lat temu. Wyraźny postęp w tym zakresie nie oznacza jednak całkowitego zaniechania przez środowisko naukowe poszukiwania nowych modeli komunikacji. Mamy nawet do czynienia z prawdziwym bogactwem i różnorodnością koncepcji, która utrudnia przewidywanie przyszłości w tym zakresie. Nawet jednak najdalej idące propozycje reform cechuje świadomość konieczności racjonalnego rozwiązania sposobu finansowania czasopism naukowych. Jest wielu autorów, którzy są przekonani o rychłym końcu dominacji komercyjnych wydawców, ale swoje przekonania opierają na racjonalnych przesłankach i szukają realnych dróg prowadzących do urzeczywistnienia takiego celu [np. Nentwich 2001]. Przede wszystkim dostrzegają jednak konieczność znalezienia źródeł finansowania lub radykalnych oszczędności i utrzymania odpłatności za dostęp na poziomie możliwym do zaakceptowania. Nie ma tu oczywiście miejsca na przedstawienie wszystkich propozycji w tym względzie, ale warto zauważyć, że większość z nich zmierza do przejęcia funkcji wydawnictw komercyjnych przez innych uczestników komunikacji naukowej: towarzystwa naukowe, biblioteki akademickie, uczelnie itp. W zasadzie w każdym tego typu projekcie poświęca się dużo miejsca właśnie na analizę sposobu finansowania nowych rozwiązań w zakresie komunikacji naukowej. Proponuje się, by informację naukową (wiedzę) traktować nie jak towar, którym się handluje, ale jak prezent, który się darowuje. Nikt już jednak nie wierzy w św. Mikołaja – wszyscy wiemy, że prezenty trzeba kupić. Płaci dalej naukowiec (lub raczej stojące za nim instytucje, sponsorzy, budżet państwa) tyle, że przedtem płacił jako czytelnik, a w proponowanych nowych modelach komunikacji płaci jako autor [Niedźwiedzka 2005, dok. elektr.].

Jakkolwiek potoczą się dalsze zmiany w komunikacji naukowej, jedno jest dzisiaj pewne – prosta, stabilna i w miarę jednolita struktura uległa zaburzeniu. Są czasopisma dystrybuowane za darmo, sprzedawane i takie, w których to autorzy płacą za opublikowanie tekstów. Co więcej, różne wersje tego samego artykułu (a czasem nawet te same) można jednocześnie kupić i znaleźć w Internecie za darmo. Konieczne jest regularne monitorowanie zjawisk zachodzących w tym środowisku, szukanie optymalnych z punktu widzenia finansowego sposobów docierania do publikacji i publikowania oraz badanie potrzeb samych naukowców (jako czytelników i jako autorów). W przeciwnym razie biblioteki akademickie będą coraz bardziej kosztowne i jednocześnie coraz mniej przydatne, a naukowcy

(szczególnie ci z biedniejszych państw) nie będą w stanie efektywnie uczestniczyć w komunikacji naukowej. To wydaje się dzisiaj jednym z ważniejszych zadań, jakie ma do spełnienia informacja naukowa i bibliotekoznawstwo.

Łatwe i szybkie publikowanie?

Spośród niespełnionych nadziei na rewolucyjne zmiany w komunikacji naukowej, które miało przynieść upowszechnienie Internetu, jeszcze przynajmniej jedna zasługuje na kilka słów – nadzieja na łatwość i szybkość publikowania. Wspomniane wcześniej trudności w połączeniu z presją wywieraną na autorów stanowiły podatny grunt dla rozwoju przekonania, że oto dzięki Internetowi (a szczególnie środowisku WWW) wystarczy wcisnąć kilka odpowiednich klawiszy na swojej klawiaturze, by „opublikować” artykuł w Internecie. Powstał mit łatwości publikowania w sieci powtarzany często bezrefleksyjnie i traktowany przez wielu jak pewnik, nie wymagający dowodzenia.

Co faktycznie Internet zmienił w procesie publikowania? Na ile go ułatwił i przyspieszył z punktu widzenia naukowca-autora? Żeby odpowiedzieć na te pytania, warto prześledzić typowy proces publikowania artykułu w czasopiśmie naukowym. Z perspektywy autora tekstu, w pewnym uproszczeniu wygląda on następująco:

- przesłanie rękopisu artykułu do redakcji (z reguły poprzedzone korespondencją)
- oczekiwanie – proces recenzowania
- wprowadzenie ewentualnych modyfikacji i poprawek i związana z tym korespondencja
- oczekiwanie (lub uczestniczenie) – obróbka redakcyjna, korekta, formatowanie i skład tekstu
- oczekiwanie na publikację w określonym numerze czasopisma
- oczekiwanie – dystrybucja czasopisma lub upublicznienie artykułu w Internecie.

Czy któryś z tych etapów został wyeliminowany przez Internet? Oczywiście, ktoś mógłby powiedzieć, że prawie wszystkie – naukowiec może bowiem zamieścić swój tekst w postaci pierwotnej, nieopracowanej pod względem redakcyjnym i formalnym, na własnej stronie WWW natychmiast po jego napisaniu i uznać go za opublikowany. Być może taki właśnie sposób myślenia jest pożywką dla rozwoju mitu o szybkości i łatwości publikowania w sieci. Ale czy w ogóle można w takim przypadku używać terminu „opublikowany” w kontekście komunikacji naukowej? Czy nierecenzowany tekst, za którym nie stoi żadna szanowana redakcja lub nawet instytucja i co gorsza, o którego istnieniu nikt lub prawie nikt nie wie (przede wszystkim ze względu na nieobecność w powszechnie znanych kanałach dystrybucji), ma szansę zaistnieć w społeczności naukowców? W istniejącym dzisiaj systemie komunikacji naukowej raczej nie.

Internet z pewnością może przyspieszać proces korespondencji i obieg dokumentów między autorem a redakcją. Jednak już przesłanie tekstu w postaci cyfrowej wprawdzie zwalnia redakcję z konieczności jego przepisywania, ale z drugiej strony *de facto* wydłuża i utrudnia pracę autora. Dzisiaj to przecież on jest często redaktorem i korektorem swojego tekstu, a często musi także wykonać znaczną część prac związanych z jego formatowaniem i składem. Choć wydaje się, że od razu i bez większych perturbacji Internet mógłby wyeliminować etap dystrybucji, tak dzieje się tylko w przypadku czasopism wyłącznie elektronicznych. W przypadku tych wydawanych w obu postaciach, wersje drukowane nadal podlegają tradycyjnym formom dystrybucji, a może się zdarzyć, że wersje elektroniczne ukazują się ze względów marketingowych z pewnym wymuszonym opóźnieniem. Jak dowodzą niektóre badania [Kling, Swygart-Hobaugh 2002, dok. elektr.], obserwowane

zmiany nie przyspieszają procesu publikacji, gdyż nie dotyczą najbardziej czasochłonnych elementów.

Dzisiaj, kiedy opadły już silne emocje związane z wkroczeniem Internetu do świata nauki i poszukuje się racjonalnych rozwiązań umożliwiających faktyczne wykorzystanie potencjału tego środowiska dla usprawnienia komunikacji naukowej, zwraca się przede wszystkim uwagę na opóźnienia w istniejącym dotychczas systemie wynikające z okresowego charakteru czasopism oraz czasochłonności procesu recenzowania. W odniesieniu do pierwszego elementu proponuje się najczęściej odejście od periodyczności i publikowanie artykułów od razu po ich ostatecznym opracowaniu (koncepcja znana pod angielskim skrótem ASAP – *as soon as publishable*). Już dzisiaj istnieją tego typu czasopisma, ale czy w odniesieniu do nich nadal można używać terminu „czasopismo”? Dużą uwagę twórców nowych koncepcji przykuwa jednak reorganizacja procesu recenzowania. Ideą, która zrodziła się jeszcze w czasach przedsiwciowych, a nabrała rozmachu dzięki możliwościom Internetu, jest upublicznianie pierwotnych wersji artykułów jeszcze przed ich zrecenzowaniem na przykład na tak zwanych serwerach preprintów. W takim przypadku tradycyjną, anonimową recenzję można zastąpić mniej lub bardziej publiczną dyskusją [Kling, Callahan 2004, dok. elektr.]. Dzięki takim rozwiązaniom środowisko naukowe zyskuje możliwość szybszego dostępu do wyników badań, ale traci jednoznaczność identyfikowania dokumentów uczestniczących w komunikacji naukowej. W miejsce jednego artykułu publikowanego w konkretnym numerze czasopisma, pojawia się wiele wersji tego samego tekstu (i co gorsza może okazać się, że żaden z nich nie będzie wersją ostateczną) ukazujących się w trudno przewidywalnym czasie. Taka zmiana skutkuje nie tylko nowymi problemami związanymi z interpretacją treści publikacji, ale także zupełnie nowymi wyzwaniami dotyczącymi sposobu indeksowania czy formułowania opisów bibliograficznych (np. w przypisach lub bibliografiach) tego typu dokumentów. Czy mają odsyłać do wybranej wersji tekstu przeczytanej przez autora? W takim razie co robić w sytuacji, gdy w danym zasobie wcześniejsze wersje są zastępowane nowymi, a stare albo całkowicie usuwane, albo archiwizowane pod innym adresem? Czy może należałoby zacząć traktować publikację naukową jako pewne kontinuum zmienne w czasie, ze wszystkimi tego konsekwencjami?

Może okazać się, że w niedalekiej przyszłości przynajmniej w niektórych dyscyplinach naukowych miejsce tradycyjnego artykułu zastąpi zupełnie nowa forma obiektu informacyjnego, a miejsce tradycyjnego czasopisma jakaś odmiana biblioteki cyfrowej [Owen 2002]. Konsekwencje takich zmian dla komunikacji naukowej i rozwoju różnych dyscyplin są dzisiaj trudne do przewidzenia i z pewnością wymagają intensywnych badań zarówno od strony naukoznawczej, jak i z punktu widzenia informacji naukowej i bibliotekoznawstwa. Bez aktualizacji ustaleń teoretycznych i odpowiednich rozwiązań praktycznych w zakresie obsługi informacyjnej naukowców może okazać się, że Internet zamiast radykalnie usprawnić komunikację naukową, w niektórych obszarach spowoduje jej regres.

Wyszukiwanie informacji naukowej w Internecie – „mit Google”

Problematyka wyszukiwania informacji utrwalonej w dokumentach, w tym – zachowań użytkowników, jest jednym z najważniejszych i najbardziej wieloaspektowych zagadnień badanych przez informację naukową i bibliotekoznawstwo. Internet, a przede wszystkim World Wide Web, niewątpliwie przyczynił się do zmiany strategii i techniki wyszukiwania, także – na potrzeby badań naukowych i dydaktyki. Z jednej strony potencjalnie oferuje błyskawiczny dostęp do największych zasobów wiedzy w dziejach ludzkości. Z drugiej – wraz z jego rozwojem problemy przeciążenia informacyjnego (*information overload*), nadmiaru informacji, jej jakości, oceny, selekcji etc. stały się bardziej palące niż kiedykolwiek wcześniej. W rezultacie zaspokojenie potrzeb informacyjnych,

szczególnie tych bardziej kompleksowych, niestety nie stało się dzięki Internetowi, jak można by oczekiwać, zawsze szybsze i łatwiejsze. Również i w tej dziedzinie mamy do czynienia z pewnymi mitami, z których dwa – powiązane ze sobą – są przedmiotem analizy w niniejszym tekście.

Po pierwsze, spora grupa użytkowników sądzi, iż uniwersalne (globalne, standardowe) wyszukiwarki (*general purpose search engines*) – a zwłaszcza Google – umożliwiają dostęp do każdej i wszystkich informacji, funkcjonujących w Sieci. Po drugie, zauważalne jest przekonanie, iż wyszukiwanie informacji w Internecie jest czynnością prostą i niewymagającą żadnych właściwie profesjonalnych umiejętności. W obydwu przypadkach nie tyle chodzi o poglądy głoszone *explicite*, ile – o faktyczne zachowania informacyjne [Brophy, Bawden 2005] [Griffiths, Brophy 2005]. Obydwa też zawierają pewne racjonalne jądro.

Globalne wyszukiwarki na czele z Google ustawicznie doskonalą – we własnym dobrze pojętym interesie – swoją ofertę, poprzez zwiększanie zakresu i zasięgu zbioru indeksowanych dokumentów, usprawnianie mechanizmów rankingu i wyświetlania wyników wyszukiwania oraz coraz lepsze, tj. łatwiejsze w obsłudze i dające bardziej precyzyjne wyniki techniki wyszukiwawcze na poziomie *Basic* i *Advanced Search*. Rzeczywiście, niektóre elektroniczne materiały, jak dokumenty w formacie PDF, multimedialne, wiadomości bieżące etc. – początkowo pomijane przez standardowe wyszukiwarki – dzisiaj są przez większość z nich odnajdywane w odpowiedzi na kwerendy [Notess 1999-2006, dok. elektr.]. Co więcej, w ramach uniwersalnych wyszukiwarek tworzone są wyspecjalizowane usługi, takie jak Google Scholar (<http://scholar.google.com/>) dedykowany informacji i komunikacji naukowej. Także same wyszukiwarki aspirują do „totalności”, na przykład na stronie WWW poświęconej informacji o firmie czytamy: „Misją Google jest uporządkowanie światowych zasobów informacji, aby stały się one powszechnie dostępne i użyteczne. (...) Google powszechnie uważa się za największą na świecie wyszukiwarę – bezpłatną i prostą w użyciu usługę, która w ułamku sekundy wyświetla trafne rezultaty wyszukiwania” [2006, dok. elektr.]. Nie powinno zatem dziwić, iż – jak pokazują badania – 72 % autorów wykorzystuje Google do poszukiwania artykułów naukowych w cyberprzestrzeni [Friend 2006, dok. elektr.].

Przedstawiona sytuacja, tj. znaczny wzrost możliwości globalnych wyszukiwarek, a w szczególności – jednej z nich, powoduje, iż część internautów utożsamia identyfikację i odnajdywanie informacji, także naukowej, z posłużeniem się serwisem Google. W języku angielskim słowa *googling* oraz *to google* stały się synonimami terminów „wyszukiwanie” i „wyszukiwać” [Brophy, Bawden 2005] [Mostafa 2005]. Być może w przyszłości ten mit się urzeczywistni, na razie jednak teza o omnipotencji Google i ewentualnie innych uniwersalnych narzędzi jest błędna. Charakteryzuje nieprofesjonalnych użytkowników Internetu. Dotychczasowe badania uczestników komunikacji naukowej prowadzone w tym kontekście nie są jeszcze wyczerpujące, poza tym – częściej za swój przedmiot miały zachowania informacyjne studentów – a nie uczonych. Tym niemniej można już pokusić się o pewne uogólnienia, mianowicie – „mit Google” w środowisku akademickim występuje w dwu wersjach. „Mocniejsza” z nich, opisana na początku tego akapitu, cechuje wyłącznie niektóre grupy studenckie. W jednej z publikacji stwierdza się wręcz: „Nawet prosta sugestia, iż poza zasięgiem Google istnieje ogromny zasób wiedzy poszerzy horyzonty studentów” [Devine, Egger-Sider 2004, s. 268]. „Słabsza” wersja analizowanego nieporozumienia polega na tym, iż użytkownicy – pomimo świadomości istnienia w środowisku WWW innych źródeł informacji, na przykład elektronicznych zasobów bibliotek akademickich, nie korzystają z nich bądź korzystają sporadycznie i niechętnie, ponieważ są przekonani, że Google niemal zawsze „wystarczy” dla zaspokojenia potrzeb informacyjnych w stopniu zadowalającym [Brophy, Bawden 2005]. I zapewne czasami wystarczy, dominacja Google na rynku narzędzi wyszukiwawczych ma solidne podstawy.

Jednakże w przypadku gromadzenia literatury na potrzeby prac licencjackich, magisterskich, nie mówiąc już o poważniejszych badaniach naukowych, nie należy ograniczać się do standardowych wyszukiwarek – a to przede wszystkim dlatego, iż żadna z nich, nawet ta najlepsza, nie tylko nie indeksuje całej zawartości informacyjnej Sieci, nie indeksuje nawet jej połowy. Przyczyny takiego stanu rzeczy są różne – mogą mieć charakter techniczny (związany ze sposobami działania i możliwościami robotów wyszukiwarek, limitami czasowymi etc.), prawny i własnościowy (ograniczenia lub brak dostępu), finansowy; mogą także wynikać z polityki firmy, między innymi – celowego wykluczenia z indeksów pewnych stron WWW (najczęściej są to strony generowane dynamicznie z baz danych w odpowiedzi na kwerendy użytkownika). Poza tym jest jeszcze kwestia jakości wyszukanej informacji, której – mimo znacznej poprawy w tym zakresie – nie są w stanie do końca zagwarantować automatyczne mechanizmy wyszukiwarek, a którą może próbować zagwarantować człowiek – na przykład tworząc *subject gateways* – tematyczne serwisy WWW o kontrolowanej jakości [Derfert-Wolf 2004, dok. elektr.].

Zasoby niewidoczne w wynikach wyszukiwania uniwersalnych serwisów, takich jak Google, noszą nazwę *Invisible Web* (*Deep Web*, *Hidden Web*) – niewidzialny (głęboki, ukryty) Internet – w opozycji do widzialnego albo powierzchniowego Webu (*Visible Web*, *Surface Web*). Specjaliści zgodnie twierdzą, że ukryty Internet stanowi w dużej mierze zbiór wartościowej, autorytatywnej, cechującej się wysoką jakością informacji, co więcej – często bezpłatnej dla użytkownika [Devine, Egger-Sider 2004] [Pamuła-Cieślak 2004, dok. elektr.]. Natomiast co do wielkości zasobów *Invisible Web* zdania są podzielone. Powszechnie w tym zakresie cytowany autor, Michael K. Bergman w 2001 roku szacował, iż *Deep Web* jest od 400 do 550 razy większy niż *Surface Web* [2001, dok. elektr.]. Nie zgadzają się z tym Dirk Lewandowski i Philipp Mayr, twierdząc, iż jest to ocena bardzo zawyżona [2006, dok. elektr.]. Wedle innego źródła w grudniu 2005 roku istniało około 55 miliardów statycznych oraz ponad 200 miliardów dynamicznych (*database-driven*) stron WWW; 90 % z nich pozostawało poza zasięgiem globalnych wyszukiwarek i katalogów [Gil 2006, dok. elektr.]. Jakkolwiek by nie było, *Deep Web* stanowi ogromny zasób informacyjny.

Szczególnie interesująca jest jego naukowa część – *Academic Invisible Web* (AIW), która zawiera prawdopodobnie od 20 miliardów do 100 miliardów dokumentów elektronicznych różnego typu [Lewandowski, Mayr 2006, dok. elektr.]. Najważniejszym – z punktu widzenia potrzeb informacyjnych w nauce – elementem AIW są bazy danych (bibliograficzne, np. katalogi biblioteczne typu OPAC; pełnotekstowe) oraz inne kolekcje wirtualne (rządowe, statystyczne, z „surowymi” danymi – *raw data*), których zawartość („wnętrze”) nie jest indeksowana przez standardowe wyszukiwarki, ze wskazanych już przyczyn. Źródła należące do *Academic Invisible Web* niejednokrotnie wymagają wypełnienia formularza wyszukiwawczego, zarejestrowania się bądź zalogowania, co nadal pozostaje domeną człowieka. W *Deep Web* znajduje się także informacja powstająca „w czasie rzeczywistym” (np. bieżące notowania giełdowe), pliki skompresowane, w nietypowych formatach, programy komputerowe itp. – ale to wydaje się nieco mniej istotne w aspekcie źródeł do badań naukowych.

Oczywiście, jak już wspomniano, wyszukiwarki uniwersalne ustawicznie rozwijają się i włączają do swoich zasobów nowe typy źródeł – redukując w ten sposób „rozmiary” ukrytego Internetu. Trudno przewidzieć, czy w przyszłości *Invisible Web* zniknie, na razie jednak nawet tak innowacyjne serwisy jak Google są w stanie uchwycić jedynie niewielki jego procent. Z tego powodu, a także ze względu na wymogi związane z jakością informacji potrzebna jest znajomość specjalistycznych narzędzi wyszukiwawczych, takich jak wyszukiwarki naukowe (np. Scirus <http://www.scirus.com/srsapp/>), serwisy tematyczne (wortale) o kontrolowanej jakości (np. SOSIG <http://www.sosig.ac.uk/>), przeszukiwarki archiwów i repozytoriów (np. OAIster <http://oaister.umdl.umich.edu/o/oaister/>), katalogi baz

danych w *Invisible Web* (np. CompletePlanet <http://aip.completeplanet.com/>) – a nawet fachowe blogi (np. Deep Web Research <http://www.deepwebresearch.info/>). Niezbędne jest kształcenie internautów w tym zakresie. Zasadniczą rolę mają tu do spełnienia bibliotekarze i profesjonaliści informacji, zarówno jako kreatorzy dostępu do wartościowej informacji, jak i edukatorzy użytkowników.

Formułowanie zapytań wyszukiwawczych (kwerend) do systemów informacyjnych również uległo przemianom pod wpływem Internetu. Jak twierdzą Nigel Ford i współautorzy: „wyszukiwanie informacji w Web różni się pod wieloma ważnymi względami od podejmowanego w bardziej tradycyjnych środowiskach wyszukiwawczych” [Ford i in. 2002, s. 30]. Generalnie cechuje je – w większym niż kiedyś stopniu – tzw. zasada najmniejszego wysiłku (*the principle of least effort*) [Griffiths, Brophy 2005]. W rezultacie użytkownicy gotowi są obniżyć standardy jakości informacji, na rzecz łatwości i szybkości jej odnalezienia. Są niecierpliwi i „niewyrozumiali” dla serwisów trudniejszych w obsłudze, rezygnują w obliczu przeszkód, raczej „skaczą” między stronami WWW niż zagłębiają się w zawartość serwisów [Nicholas i in. 2003]. Jak wykazały liczne badania, większość użytkowników informacji elektronicznej z reguły (jeżeli nie są do tego zmuszeni) nie używa bardziej zaawansowanych technik wyszukiwawczych, nie formułuje kompleksowych pytań i nie ma ochoty wchodzić w głębsze interakcje z systemem ani wykorzystywać wszystkich jego możliwości, niezależnie od tego czy jest to wyszukiwarka internetowa, czy biblioteczny OPAC. Między innymi – tylko kilka procent kwerend zawiera operatory boolowskie, zdecydowana większość sesji wyszukiwawczych w Web składa się z dwu pytań złożonych średnio z dwu terminów, a więcej niż połowa użytkowników przegląda jedynie pierwsze dziesięć (a nawet mniej) „wydanych” przez wyszukiwarki dokumentów [Ford i in. 2002] [Griffiths, Brophy 2005] [Spink i in. 2006].

Poza tym, zdaniem niektórych uczonych użytkownicy informacji doby Internetu stają się „informacyjnymi graczami”, tzn. wyszukiwanie informacji, niezależnie od celu, zawiera – albo musi zawierać – elementy zabawy, gry, konkurencji [Nicholas i in. 2003]. Dodatkowo, coraz bardziej widoczne jest zachowanie polegające – w uproszczeniu – na wyszukiwaniu kilku informacji jednocześnie (*multitasking information behavior & information task switching*) [Spink i in. 2006].

„Bezproblemowość” wyszukiwania jest drugim składnikiem „mitu Google”. Użytkownicy „z góry” postrzegają odnajdywanie informacji przy pomocy Google i innych uniwersalnych serwisów jako łatwe, szybkie i przyjemne; natomiast wyszukiwanie poprzez wyspecjalizowane narzędzia, np. biblioteczne – jako trudne, powolne i wymagające zbyt dużego (w relacji do przewidywanych rezultatów) wysiłku intelektualnego – i w efekcie rezygnują z tych ostatnich. Niejednokrotnie jest to, oczywiście, postępowanie uzasadnione – ale niestety nie zawsze, zwłaszcza w odniesieniu do źródeł naukowych. *Notabene* właściciele globalnych wyszukiwarek doskonale zdają sobie sprawę z preferencji użytkowników i sukcesywnie „przystosowują” je do zachowań użytkowników informacji.

Zakończenie

Internet niewątpliwie zmienia oblicze współczesnej nauki i edukacji. Opisane w artykule mity dotyczące bezpłatnego dostępu do zasobów naukowych, publikowania w WWW oraz wyszukiwania informacji w Sieci pojawiły się nie bez przyczyny. U ich podstaw leżą rzeczywiste osiągnięcia techniczne i organizacyjne związane z rozwojem Internetu, a także – cechy ludzkiej natury – nadzieja, że „będzie lepiej” oraz niechęć do podejmowania „zbędnego” wysiłku. Być może wskazane nieporozumienia nie są szczególnie szkodliwe z punktu widzenia przeciętnego użytkownika, wykorzystującego Internet w życiu codziennym, poszukującego w nim „prostych” informacji, rozrywki czy kontaktów międzyludzkich.

Inaczej rzecz się ma w odniesieniu do naukowców, nauczycieli i studentów, którzy – z oczywistych względów – powinni być kompetentnymi użytkownikami oraz współtwórcami Sieci. Identyfikacja mitów oraz korekta nie do końca poprawnych wyobrażeń i schematów myślowych przyczynia się do bardziej efektywnego wykorzystania potencjału Internetu – a w efekcie do poprawy jakości badań naukowych i (samo)kształcenia. Ogromną rolę w tym zakresie ma do odegrania informacja naukowa i bibliotekoznawstwo, zarówno w obszarze poznawczym, tj. badając Internet jako medium komunikacji naukowej i zasób informacyjny, jak i w dziedzinie kształcenia profesjonalistów informacji – ekspertów i przewodników w cyberprzestrzeni.

Wykorzystana literatura

- Bergman, Michael K. (2001) [dok. elektr.]. The Deep Web: Surfacing Hidden Value. *The Journal of Electronic Publishing* Vol. 7 Issue 1. <http://www.press.umich.edu/jep/07-01/bergman.html> [odczyt 10.05.2006]
- Björk, Bo-Christer (2005) [dok. elektr.]. Open access to scientific publications – an analysis of the barriers to change? *EBIB* nr 2 (63). <http://ebib.oss.wroc.pl/2005/63/bjork.php> [odczyt 07.05.2006]
- Björk, Bo-Christer; Ziga Turk (2000) [dok. elektr.]. How scientists retrieve publications: an empirical study of how the Internet is overtaking paper media. *The Journal of Electronic Publishing* Vol. 6 Issue 2. <http://www.press.umich.edu/jep/06-02/bjork.html> [odczyt 07.09.2005]
- Brophy, Jan; David Bawden (2005). Is Google enough? Comparison of an internet search engine with academic library resources. *Aslib Proceedings: New Information Perspectives* Vol. 57 No. 6, pp. 498-512.
- Cronin, Blaise (2003). Scholarly communication and epistemic cultures. <http://www.arl.org/scomm/disciplines/Cronin.pdf> [odczyt 07.05.2006]
- Derfert-Wolf, Lidia (2004) [dok. elektr.]. Serwisy tematyczne o kontrolowanej jakości w Internecie. *EBIB* nr 6 (57). <http://ebib.oss.wroc.pl/2004/57/derfert.php> [odczyt 10.05.2006]
- Devine, Jane; Francine Egger-Sider (2004). Beyond Google: The Invisible Web in the Academic Library. *The Journal of Academic Librarianship* Vol. 30 No. 4, pp. 265-269.
- Ford, Nigel; David Miller; Nicola Moss (2002). Web search strategies and retrieval effectiveness: an empirical study. *Journal of Documentation* Vol. 58 No. 1, pp. 30-48.
- Friend, Frederick J. (2006) [dok. elektr.]. Google Scholar: Potentially Good for Users of Academic Information. *The Journal of Electronic Publishing* Vol. 9 No. 1. <http://hdl.handle.net/2027/spo.3336451.0009.105> [odczyt 19.05.2006]

- Gil, Paul (2006) [dok. elektr.]. What is "The Invisible Web".
<http://netforbeginners.about.com/cs/secondaryweb1/a/secondaryweb.htm> [odczyt 10.05.2006]
- Google (2006) [dok. elektr.]. Informacje o firmie.
<http://www.google.pl/intl/pl/corporate/index.html> [odczyt 23.05.2006]
- Griffiths, Jillian R.; Peter Brophy (2005). Student Searching Behavior and the Web: Use of Academic Resources and Google. *Library Trends* Vol. 53 No. 4, pp. 539-554.
- Harnad, Stevan (2005) [dok. elektr.]. Fast-Forward on the Green Road to Open Access: The Case Against Mixing Up Green and Gold. *Ariadne* Issue 42.
<http://www.ariadne.ac.uk/issue42/harnad/> [odczyt 07.05.2006]
- Harnad, Stevan i in. (2004) [dok. elektr.]. The Access/Impact Problem and the Green and Gold Roads to Open Access. *Serials Review* Vol. 30 Issue 4.
<http://www.ecs.soton.ac.uk/~harnad/Temp/impact.html> [odczyt 07.05.2006]
- Kling, Rob; A. J. Swygart-Hobaugh (2002) [dok. elektr.]. The Internet and the velocity of scholarly publishing. <http://www.slis.indiana.edu/CSI/WP/WP02-12B.html> [odczyt 07.05.2006]
- Kling, Rob; Ewa Callahan (2003) [dok. elektr.]. Electronic journals, the Internet, and scholarly communication. *Annual Review of Information Science and Technology* Vol. 37, pp. 127-177. <http://www.slis.indiana.edu/CSI/WP/wp01-04B.html> [odczyt 15.12.2004]
- Lewandowski, Dirk; Philipp Mayr (2006) [dok. elektr.]. Exploring the Academic Invisible Web. http://eprints.rclis.org/archive/00006071/01/LHT_Preprint.pdf [odczyt 10.05.2006]
- Mogge, D. (1999). Seven years of tracking electronic publishing: the ARL Directory of electronic journals, newsletters and academic discussion lists. *Library Hi Tech* Vol. 17 No. 1, pp. 17-25.
- Mostafa, Javed (2005). Seeking Better Web Searches. *Scientific American* Vol. 292 Issue 2, pp. 66-73.
- Nentwich, Michael (2001). (Re-) De-commodification in academic knowledge distribution? *Science Studies* Vol. 14 No. 2, pp. 21-42.
- Nicholas, David; Tom Dobrowolski; Richard Withey; Chris Russell; Paul Huntington; Peter Williams (2003). Digital information consumers, players and purchasers: information seeking behaviour in the new digital interactive environment. *Aslib Proceedings: New Information Perspectives* Vol. 55 Issue 1/2, pp. 23-31.
- Niedźwiedzka, Barbara (2005) [dok. elektr.]. Bezpлатne publikacje naukowe i bezpлатne archiwa naukowe w dziedzinie medycyny. *EBIB* nr 2 (63).
<http://ebib.oss.wroc.pl/2005/63/niedzwiadzka.php> [odczyt 15.01.2006]

Notess, Greg R. (1999-2006) [dok. elektr.]. Search Engine Showdown.
<http://www.searchengineshowdown.com/> [odczyt 20.05.2006]

Owen, John Mackenzie (2002). The new dissemination of knowledge: digital libraries and institutional roles in scholarly publishing. *Journal of Economic Methodology* Vol. 9 Issue 3, pp. 275-288.

Pamuła-Cieślak, Natalia (2004) [dok. elektr.]. Ukryty Internet – jeśli nie wyszukiwarka, to co? *EBIB* nr 7 (58). <http://ebib.oss.wroc.pl/2004/58/pamula.php> [odczyt 18.05.2006]

Spink, Amanda; Minsoo Park; Bernard J. Jansen; Jan Pedersen (2006) [dok. elektr.]. Multitasking during Web search sessions. *Information Processing and Management* Vol. 42, pp. 264-275.