
Widoczność polskich publikacji naukowych w Internecie

Arkadiusz Pulikowski

*Institut Bibliotekoznawstwa i Informacji Naukowej
Uniwersytet Śląski w Katowicach*

Abstrakt

Cel/teza: Celem przeprowadzonych badań było ustalenie, w jakim stopniu miejsce udostępniania publikacji wpływa na widoczność dokumentów w wyszukiwarkach ogólnych i specjalistycznych oraz pokazanie zmian w widoczności polskich publikacji naukowych, jakie zaszły w latach 2012–2014.

Koncepcja/metody badań: Badaniem objęto wybrane publikacje, głównie z zakresu informatologii, znajdujące się w różnych typach zasobów (m.in.: repozytoriach, bibliotekach cyfrowych, stronach internetowych czasopism). Widoczność wyselekcjonowanych dokumentów określono na podstawie możliwości ich odnalezienia w wybranych wyszukiwarkach ogólnych i specjalistycznych. Zestawienie wyników badań wykonanych z dwuletnim odstępem czasu umożliwiło pokazanie zmian w widoczności obserwowanych dokumentów.

Wyniki i wnioski: Najlepiej pod względem widoczności wypadły repozytoria i czasopisma na platformie Open Journal Systems. W porównaniu z rokiem 2012 odnotowano wyraźną poprawę widoczności polskich zasobów naukowych. Google powtórzył dobre wyniki z 2012 roku, pozostając liderem, natomiast Bing poprawił indeksowanie i znacznie zbliżył się do niego osiąganymi rezultatami.

Oryginalność/wartość poznawcza: Wyniki badań pomogą pracownikom naukowym w wyborze najkorzystniejszych pod względem widoczności miejsc publikowania. Jednocześnie pokazują również, do jakich zasobów są w stanie dotrzeć znane serwisy wyszukiwawcze. Tego typu badania nie były dotąd prowadzone w naszym kraju.

Słowa kluczowe

Czasopisma elektroniczne. Biblioteki cyfrowe. Niewidoczny Internet. Ukryty Internet. Open Access. Otwarta Nauka. Repozytoria. Serwisy społecznościowe dla naukowców. Widoczność zasobów naukowych. Wyszukiwanie informacji. Wyszukiwarki naukowe.

Otrzymano: 7.04.2015. Zrecenzowano: 01.06.2015. Zaakceptowano: 29.06.2015.

1. Zasoby otwartej nauki na tle zjawiska niewidocznego Internetu

Po latach dominacji serwisów komercyjnych coraz większą rolę w upowszechnianiu wyników badań naukowych zaczynają odgrywać publikacje o dostępie otwartym. Są one udostępniane w repozytoriach, bibliotekach cyfrowych, na stronach wydawców czasopism, na stronach instytucji naukowych, w serwisach społecznościowych dla naukowców, na stronach domowych autorów, blogach oraz w popularnych serwisach, takich jak: Scribd, Zoho, Prezi, SlideShare czy Calameo. Celem otwartego publikowania w Internecie jest przede wszystkim upowszechnianie wiedzy. Dla autorów nie mniej ważna jest promocja dorobku, która w dobie postępującej parametryzacji nauki, staje się nieodzownym wręcz

elementem aktywności zawodowej, szczególnie w naukach ścisłych, przyrodniczych i technicznych. W wielu badaniach¹ potwierdzono wzrost liczby cytowań publikacji udostępnianych w modelu Open Access. Nie bez znaczenia dla autorów jest również przyspieszenie ogłaszania wyników badań. Jest to możliwe dzięki udostępnianiu preprintów i prezentacji konferencyjnych, a także za sprawą skrócenia procesu wydawniczego czasopism ukazujących się w modelu otwartym. Zamieszczanie prac naukowych w otwartych zasobach daje także szansę na nawiązanie nowych kontaktów, szczególnie w wymiarze międzynarodowym.

Zarówno autorom publikacji, jak i instytucjom organizującym otwarty dostęp, zależy na tym, by jak najwięcej osób dotarło do udostępnianych treści. W jakim stopniu cel ten uda się osiągnąć, zależy w głównej mierze od widoczności prezentowanych zasobów. Problem widoczności zasobów pojawia się w literaturze najczęściej w kontekście zjawiska niewidocznego² Internetu (ang. *invisible web*, *hidden web*, *deep web*). Takim mianem określa się zasoby nieindeksowane przez wyszukiwarki internetowe. Widoczność zasobu jest więc zależna od tego, czy można do niego dotrzeć za pośrednictwem narzędzi wyszukiwawczych. Choć w definicjach niewidocznego Internetu pojawia się często zawężenie do wyszukiwarek głównych (DIST, 2007, 370), czy też ogólnego przeznaczenia (Sherman & Price, 2001, 57), nic nie stoi na przeszkodzie, by widoczność-niewidoczność zasobów odnieść do konkretnego serwisu wyszukiwawczego. W takim ujęciu każda wyszukiwarka tworzy własny niewidoczny Internet, zależnie od zapisanego w niej algorytmu indeksującego (Devine & Egger-Sider, 2009, 8). W ten właśnie sposób będzie rozumiana widoczność publikacji w analizowanych serwisach typu Open Access.

W zależności od tego, w jakim stopniu wyszukiwarki zaindeksują informacje o dokumencie, można wyróżnić:

- widoczność pełną (opis dokumentu i pełny tekst),
- widoczność częściową (ograniczona najczęściej do elementów opisu),
- niewidoczność.

Każdy z wymienionych typów widoczności można odnieść do pojedynczych serwisów wyszukiwawczych lub sumarycznej widoczności we wszystkich serwisach branych pod uwagę.

2. Charakterystyka badań

Pierwsze badanie przeprowadzono w listopadzie 2012 r., drugie w listopadzie 2014 r.³ W ramach wydzielonych typów źródeł badaniami objęto następujące zasoby o dostępie otwartym:

- a. repozytoria
 - AMUR – repozytorium Uniwersytetu im. Adama Mickiewicza w Poznaniu,
 - E-LIS – E-prints in Library and Information Science,
 - CEON – Repozytorium Centrum Otwartej Nauki,

¹ Przegląd ponad 30 badań, z których większość potwierdziła wzrost liczby cytowań publikacji udostępnianych w otwartym dostępie, można znaleźć w pracy Almy Swan (Swan, 2010).

² W polskojęzycznej literaturze częściej używany jest termin *ukryty Internet*. Autor preferuje jednak w tym przypadku równorzędny termin *niewidoczny Internet*, który tworzy logiczną opozycję do omawianego problemu widoczności zasobów.

³ Wyniki były prezentowane na IV i V ogólnopolskiej konferencji z cyklu „Zarządzanie Informacją w Nauce”.

- b. biblioteki cyfrowe
 - Dolnośląska Biblioteka Cyfrowa (dalej: DBC),
 - Śląska Biblioteka Cyfrowa (dalej: ŚBC),
 - Zachodniopomorska Biblioteka Cyfrowa Pomerania (dalej: ZBCP),
 - Kujawsko-Pomorska Biblioteka Cyfrowa (dalej: KPBC),
 - Bibliologiczna Biblioteka Cyfrowa (dalej: BBC),
 - Biblioteka Cyfrowa Politechniki Warszawskiej (dalej: BCPW),
 - Biblioteka Cyfrowa UMCS (dalej: UMCS),
- c. strony instytucji (materiały konferencyjne)
 - PTIN – „X Forum INT”, 2008 (dalej: PTIN),
 - Biblioteka Politechniki Poznańskiej – „Informacja dla nauki a świat zasobów cyfrowych”, 2008 (dalej: BPP),
- d. czasopisma internetowe
 - Biuletyn EBIB,
 - iNFOTEZY,
- e. serwisy społecznościowe dla naukowców
 - Academia.edu,
 - ResearchGate.net,
- f. popularne serwisy
 - SlideShare,
 - Scribd.

Dla każdego źródła wybrano przynajmniej dwa dokumenty, które były poszukiwane w opisanych niżej wyszukiwarkach według tytułu lub w jego fragmencie oraz według fraz pochodzących z pełnego tekstu. Tematyka większości publikacji obejmowała zagadnienia informatologii i bibliologii. Badania objęto polskojęzyczne dokumenty, które pojawiły się w Internecie nie później niż na początku 2012 r., tak, by serwisy wyszukiwawcze zdążyły je zaindeksować.

Do analizy widoczności wyżej wymienionych źródeł wykorzystano wyszukiwarki ogólnego przeznaczenia: Google i Bing⁴ (bing.com), wyszukiwarkę naukową Google Scholar (scholar.google.pl) oraz wyszukiwarkę repozytoriów, bibliotek cyfrowych i czasopism BASE⁵ (base-search.net). Ostatnia z wymienionych indeksuje (a dokładniej agreguje) tylko opisy dokumentów, dlatego w części wyników dotyczących pełnych tekstów pojawia się oznaczenie „nd” – nie dotyczy. Z uwagi na zakończenie na początku 2014 r. działalności wyszukiwarki Scirus, nie została ona uwzględniona. Kilka serwisów odrzucono w trakcie wstępnych badań, gdyż liczba zwracanych przez nie wyników była zbyt mała: CiteSeerX.ist.psu.edu, ScienceResearch.com, Core.ac.uk, Academic.Research.Microsoft.com, WorldWideScience.org.

Przeprowadzone badania miały dwa zasadnicze cele:

- określenie stopnia widoczności publikacji naukowych umieszczanych w wybranych zasobach otwartych,
- określenie skuteczności serwisów wyszukiwawczych w poszukiwaniach naukowych.

⁴ Bing to wyszukiwarka firmy Microsoft, wcześniej znana pod nazwą MSN Search i Live Search. To jeden z nielicznych realnych konkurentów Google.

⁵ BASE – Bielefeld Academic Search Engine. Wyszukiwarka rozwijana jest na Uniwersytecie w Bielefeld (Niemcy).

Realizacja pierwszego celu pozwoli autorom tekstów na wybór korzystniejszych pod względem widoczności serwisów. Jednocześnie wskaże instytucjom, zarządzającym tymi zasobami, ewentualną potrzebę optymalizacji. Z kolei realizacja drugiego celu uświadomi użytkownikom wyszukiwarek, do jakich dokumentów rzeczywiście można dotrzeć za pomocą narzędzi, z których na co dzień korzystają. Przedmiotem analizy nie były elementy interfejsu wykorzystywanych w badaniu wyszukiwarek, a jedynie możliwość dotarcia za ich pośrednictwem do wybranego wcześniej dokumentu.

3. Wyniki badań

Wyniki przeprowadzonych badań przedstawione zostały w poniższych zestawieniach tabelarycznych. W kolumnach umieszczono źródła analizowanych zasobów. Cyframi oznaczono dokumenty, których opisy znajdują się w załączniku. Osobno badana była widoczność tytułu (T) i pełnego tekstu (F) publikacji. Poprawę widoczności zasobów, w porównaniu z danymi uzyskanymi w 2012 r., zaznaczono kolorem niebieskim, obniżenie widoczności – czerwonym. Kolor zielony został użyty dla oznaczenia dokumentów widocznych już w 2012 r., w celu lepszego zobrazowania zasobów niewidocznych (bez koloru, lub sporadycznie kolorze czerwonym) na tle widocznych. Wszystkie tabele są opracowaniami własnymi autora.

Tabela 1 prezentuje zestawienie widoczności publikacji pochodzących z trzech repozytoriów: dziedzicznego – E-LIS, instytucjonalnego – AMUR i adresowanego do całego polskiego środowiska naukowego – CEON. Wszystkie korzystają z otwartego oprogramowania DSpace, wspierającego wiele standardów, w tym protokoły OAI-PMH i OAI-ORE, zapewniające możliwość przesyłania i agregowania (na potrzeby takich serwisów jak BASE) metadanych obiektów cyfrowych. Ogólnie repozytoria wypadły bardzo dobrze pod względem widoczności zdeponowanych zasobów. W porównaniu do 2012 r. wyraźnie poprawiła się widoczność repozytoriów w Google Scholar. Lepiej radzi sobie również wyszukiwarka Bing, jednak Google wciąż pozostaje najbardziej wszechstronnym narzędziem pod względem dostępu do publikacji.

Tabela 1. Widoczność publikacji w repozytoriach

	E-LIS				AMUR				CEON					
	1.		2.		3.		4.		5.		6.		7.	
	T	F	T	F	T	F	T	F	T	F	T	F	T	F
Google	+	+	+	+	+	+	+	+	+	+	+	+	+	-
Bing	+	+	+	-	+	+	-	-	+	+	-	-	+	+
Google Scholar	+	+	-	-	+	+	+	+	+	+	+	+	+	+
BASE	+	nd	+	nd	+	nd	+	nd	+	nd	+	nd	+	nd

W tabelach 2 i 3 przedstawiono widoczność publikacji udostępnianych w bibliotekach cyfrowych z podziałem na książki i artykuły z czasopism. W obu przypadkach rzuca się w oczy całkowita niewidoczność w Google Scholar. Z pewnością nie wynika to z braku obsługi

standardów, gdyż pod tym względem biblioteki cyfrowe korzystające z oprogramowania dLibra (wszystkie ujęte w badaniu), wypadają nie gorzej niż repozytoria. Można jedynie przypuszczać, że Google Scholar nie indeksuje zawartości bibliotek cyfrowych z uwagi na to, że nie jest w stanie odróżnić zasobów o charakterze naukowym od pozostałych dostępnych w bibliotekach cyfrowych. Powodów zapewne jest więcej.

Tabela 2. Widoczność książek w bibliotekach cyfrowych

	PDF						DJVU		HTML	
	ŚBC				BCPW		UMCS		DBC	
	1.		2.		3.		4.		5.	
	T	F	T	F	T	F	T	F	T	F
Google	+	-	+	+	+	+	+	+	+	+
Bing	+	-	+	+	+	-	+	+	+	+
Google Scholar	-	-	-	-	-	-	-	-	-	-
BASE	+	nd	+	nd	+	nd	+	nd	+	nd

Wyniki zgromadzone w obu tabelach pozwalają stwierdzić, że formaty PDF, DJVU i HTML nie różnią się w sposób istotny widocznością w serwisach wyszukiwawczych indeksujących zasoby bibliotek cyfrowych. Dokładniejsze rozróżnienie wymagałoby znaczącego zwiększenia wielkości próby badanych dokumentów. Podobnie jak w przypadku repozytoriów, Bing poprawił swoje wyniki w porównaniu z 2012 r. Wyszukiwarka BASE nie jest obecna w tabeli 3, gdyż przedmiotem analizy były tu artykuły zawarte w dokumentach zawierających cały numer czasopisma. Metadane są w takim przypadku tworzone dla całego zeszytu, co uniemożliwia dotarcie do zawartych w nim publikacji przez agregatory takie jak BASE. Problem ten dotyczy również prac zbiorowych, w szczególności materiałów konferencyjnych.

Tabela 3. Widoczność artykułów z czasopism w bibliotekach cyfrowych

	PDF				DJVU			
	ZBC		ŚBC		KPBC		BBC	
	1.	2.	3.	4.	5.	6.	7.	8.
Google	+	+	+	+	+	+	+	+
Bing	+	+	-	-	+	+	+	+
Google Scholar	-	-	-	-	-	-	-	-

Widoczność artykułów z czasopism dostępnych na stronach wydawców prezentują tabele 4 i 5. W pierwszej z nich zestawiono publikacje udostępniane w dwóch różnych systemach zarządzania treścią (CMS) – WordPress dla Biuletynu EBIB i Open Journal Systems dla czasopisma iNFOTEZY. Mimo poprawy indeksowania w wyszukiwarkach Google i Bing dla artykułów z Biuletynu EBIB w 2014 r., pozostała zasadnicza różnica jeśli chodzi o obecność w wyszukiwarkach Google Scholar i BASE.

Tabela 4. Widoczność artykułów z czasopism na stronach internetowych wydawców

	EBIB WordPress				iNFOTEZY			
	1.		2.		3.		4.	
	T	F	T	F	T	F	T	F
Google	+	-	+	+	+	+	+	+
Bing	+	+	+	+	+	+	+	+
Google Scholar	-	-	-	-	+	+	+	+
BASE	-	nd	-	nd	+	nd	+	nd

Obie publikacje z Biuletynu EBIB pochodzą z 2011 r. Począwszy od 2012 r. czasopismo zaczęło korzystać z Open Journal Systems (OJS), tego samego systemu zarządzania treścią, który pozwolił iNFOTEZ-om osiągnąć prawie pełną widoczność we wszystkich serwisach wyszukiwawczych już w 2012 r.

Tabela 5. Widoczność artykułów Biuletynu EBIB w systemie OJS

	1.		2.		3.		4.	
	T	F	T	F	T	F	T	F
Google	+	+	+	+	+	+	+	+
Bing	+	+	+	-	+	+	+	-
Google Scholar	+	+	+	+	+	+	+	+
BASE	-	nd	-	nd	+	nd	+	nd

W tabeli 5 przedstawiono wpływ wdrożenia nowego systemu na widoczność artykułów. Google Scholar „zobaczył” publikacje z Biuletynu EBIB w OJS, lecz inaczej niż w przypadku iNFOTEZ, artykuły nie uwidoczniły się dla wyszukiwarki BASE. Co ciekawe, publikacje o numerach 3 i 4 są odnajdywane w BASE, ale tylko dlatego, że ich autorki umieściły swoje teksty równolegle w repozytoriach instytucjonalnych. Wśród indeksowanych obecnie przez BASE 3450 źródeł z całego świata, 92 pochodzą z Polski. Są to biblioteki cyfrowe, repozytoria i nieliczne wciąż czasopisma. Nie ma wśród nich Biuletynu EBIB, zapewne dlatego, że nie został zarejestrowany w agregatorze. Służy do tego specjalny formularz⁶. Czasopisma w agregatorach metadanych są obecne od niedawna. W naszym rodzimym serwisie Federacji Bibliotek Cyfrowych (fbc.pionier.net.pl) pojawiają się dopiero pierwsze periodyki (np. *Computer Science Journal*). Z czasem na pewno zaczną ich przybywać, a to za sprawą rosnącej popularności systemu OJS. Podobnie jak repozytoria i biblioteki cyfrowe, obsługuje on wiele standardów, w tym wspomniany już OAI-PMH, dzięki któremu artykuły dostępne na platformie OJS mogą być indeksowane przez agregatory.

Kolejnym typem zasobów, którego widoczność została poddana analizie, są serwisy społecznościowe dla naukowców. Nie były one badane w 2012 r., stąd dane zawarte w tabeli

⁶ Dostępny pod adresem <http://base-search.net/about/en/suggest.php>

6 uwzględniają tylko najnowsze informacje z 2014 r. Wybrane serwisy to dwa największe tego typu – Academia.edu i ResearchGate.net. Mogą one się pochwalić odpowiednio 20 i 6 milionami użytkowników. Z przedstawionych w tabeli danych wynika, że oba serwisy są indeksowane jedynie przez wyszukiwarki ogólnego przeznaczenia. Na pocieszenie można stwierdzić, że jest to indeksowanie pełne.

Tabela 6. Widoczność publikacji na stronach serwisów społecznościowych dla naukowców

	Academia.edu				ResearchGate.net			
	1.		2.		3.		4.	
	T	F	T	F	T	F	T	F
Google	+	+	+	+	+	+	+	+
Bing	+	+	+	+	+	+	+	+
Google Scholar	-	-	-	-	-	-	-	-
BASE	-	Nd	-	Nd	-	nd	-	nd

Tabela 7. Widoczność materiałów konferencyjnych na stronach instytucji

	PTIN				BPP			
	1.		2.		3.		4.	
	T	F	T	F	T	F	T	F
Google	+	+	+	+	+	+	+	+
Bing	+	+	+	+	-	-	-	-
Google Scholar	-	-	-	-	-	-	+	+

Wiele publikacji naukowych jest dostępnych na stronach internetowych instytucji i organizacji, które przyczyniły się do ich powstania. Typowym przykładem są materiały konferencyjne (prezentacje, artykuły), zamieszczane na stronie konferencji. Widoczność tego typu publikacji prezentuje tabela 7. Google wyświetlił pełny opis dokumentów, natomiast Bing – raz lepiej, a raz gorzej. Google Scholar zaindeksował jeden dokument z 4 badanych. Co ciekawe, mimo iż dokumenty 3 i 4 pochodziły z tej samej konferencji, tylko jeden z nich został zaindeksowany. Brak wsparcia dla standardów wymiany metadanych uniemożliwił przeprowadzenie badania w wyszukiwarce BASE. Warto w tym miejscu wspomnieć o możliwości wykorzystania do tworzenia stron WWW konferencji otwartego oprogramowania Open Conference Systems, powstającego w ramach programu Public Knowledge Project – tego samego, który rozwija Open Journal Systems.

Ostatnim analizowanym miejscem udostępniania zasobów naukowych są popularne serwisy Scribd i SlideShare⁷. Pliki udostępniane w Scribd stanowią wyzwanie dla wyszukiwarek. Najwyraźniej zastosowane w tym serwisie rozwiązania techniczne utrudniają

⁷ Niedostępna w drugim badaniu prezentację nr 3 – „Dzielenie się wiedzą” autorstwa Sabiny Cisek zastąpiła inna, tej samej autorki – „Użytkownicy informacji”, również z 2009 r.

indeksowanie. Tę obserwację potwierdzono na kilku innych dodatkowo sprawdzonych dokumentach. Takich niedogodności nie stwarza SlideShare, którego zasoby są dobrze widoczne zarówno w Google, jak i w Bing. Wyszukiwarka Google Scholar nie indeksuje i nie przeszukuje obu serwisów.

Tabela 8. Widoczność publikacji w serwisach SlideShare i Scribd

	Scribd		SlideShare	
	1.	2.	3.	4.
Google	+	+	+	+
Bing	-	-	+	+
Google Scholar	-	-	-	-

4. Google Scholar

Wyszukiwarka Google Scholar wymaga osobnego omówienia, gdyż jest szczególnie ważnym narzędziem dla pracowników nauki. Po zamknięciu Scirusa jest to właściwie jedyna powszechnie znana i wykorzystywana wyszukiwarka naukowa. Warto być w niej widocznym. Na jej podstawie wyliczane są cytowania w Google Scholar Citations oraz w nakładce Publish or Perish. Wyniki z Google Scholar pojawiają się często w poszukiwaniach o charakterze naukowym, jako sugestia klasycznego Google. Z wyszukiwarki tej warto korzystać w pracach badawczych. Dla większości anglojęzycznych czasopism z zakresu bibliotekoznawstwa i nauki o informacji Google Scholar indeksuje ponad 95% artykułów (Lewandowski, 2010, 259).

Na podstawie przeprowadzonych badań można stwierdzić, że Google Scholar:

a. indeksuje:

- repozytoria (znaczną poprawę w porównaniu do 2012 r.),
- czasopisma na stronach wydawców w OJS,
- niektóre publikacje naukowe na stronach instytucji,

b. nie indeksuje:

- bibliotek cyfrowych,
- serwisów SlideShare i Scribd,
- serwisów społecznościowych dla naukowców.

W dokumencie *Inclusion Guidelines for Webmasters*, na stronach pomocy Google Scholar⁸, można znaleźć wskazówki nt. dodawania publikacji do indeksu wyszukiwarki. Sugerowane oprogramowanie dla repozytoriów to Eprints, Digital Commons i DSpace. Z kolei dla wydawców czasopism zaleca się korzystanie z Open Journal Systems, Atypion, Highwire Press, MetaPress. Informacje te są zgodne z wynikami uzyskanymi w badaniu. Indywidualnym autorom sugeruje się natomiast stosowanie:

- pełnych tekstów w formacie PDF,
- tytułu publikacji pisanego dużą czcionką na górze dokumentu,

⁸ <http://scholar.google.com/intl/pl/scholar/inclusion.html>

- poniżej w osobnym wierszu powinny znaleźć się nazwiska autorów publikacji,
 - na końcu dokumentu References/Bibliography,
 - resztą mają zająć się roboty Google Scholar.

W repozytoriach i czasopismach OJS, wspierających metadane, nie trzeba przejmować się zalecaną przez Google Scholar kolejnością opisów w plikach PDF. Natomiast w innych przypadkach trzeba tych zaleceń przestrzegać, gdyż Google Scholar może nieprawidłowo zaindeksować dokument. W przypadku prac zbiorowych, aby zostały one prawidłowo rozpoznane, należy podzielić je na osobne pliki z poszczególnymi artykułami. W przeciwnym przypadku Google Scholar rozpozna tylko pierwszą publikację, a pozostałe pominie. Więcej informacji nt. dodawania publikacji do Google Scholar można znaleźć w obszernym poradniku autorstwa Emanuela Kulczyckiego (Kulczycki, 2013).

5. Wnioski

Z całą pewnością można stwierdzić, że nie jest obojętne to, gdzie publikujemy w sieci. Najlepiej pod względem widoczności wypadły repozytoria i czasopisma na platformie Open Journal Systems. Widoczność zasobów jest w dużym stopniu zależna od standardów wspieranych przez platformę, na której są udostępniane, jednak w przypadku Google Scholar korzystanie ze standardów nie zawsze gwarantuje widoczność w tej wyszukiwarce (przykład bibliotek cyfrowych). Implementacja standardów daje jednak dodatkowe korzyści związane z wyszukiwaniem w obrębie wydzielonych pól opisu dokumentu (np. po autorze, tytule). Jeśli publikacja znajduje się w mało widocznym miejscu sieci i nie występują ograniczenia natury prawnej, to warto dodatkowo umieścić taki dokument w jednym z wielu dostępnych repozytoriów.

Porównanie wyników dla 2012 i 2014 r. pozwala stwierdzić wyraźną poprawę widoczności polskich zasobów naukowych. W sporadycznych przypadkach dokumenty widoczne w 2012 r. były niewidoczne w 2014 r.

Patrząc na wyniki z perspektywy użytkowników poszukujących informacji, trzeba zauważyć wyraźną poprawę indeksowania w Bing. Wyszukiwarka ta zbliżyła się skutecznością do Google, który z kolei powtórzył dobre wyniki z 2012 r. W przeważającej większości przypadków Google był w stanie odnaleźć zarówno opis, jak i pełny tekst zadanego dokumentu. Nie można jednak zapominać o szumie informacyjnym jaki generuje ta wyszukiwarka, co znacząco ogranicza komfort prowadzonych poszukiwań. Nie zraża to najwyraźniej studentów i pracowników naukowych, którzy właśnie od Google rozpoczynają poszukiwanie informacji (Haglund & Olsson, 2008, 56). Serwisy przeznaczone dla nauki należy postrzegać jako bardziej przejrzyste, ale jednocześnie albo bardziej „kapryśne”, jak w przypadku Google Scholar, albo ograniczone tylko do opisów dokumentów, jak to ma miejsce w BASE. Chcąc zwiększyć zasięg poszukiwań, trzeba korzystać równoległe z kilku serwisów wyszukiwawczych.

Bibliografia

- Derfert-Wolf, L. (2007). *Odkrywanie niewidzialnych zasobów sieci*. II seminarium z cyklu „Infobroker : Wyszukiwanie i przetwarzanie cyfrowych informacji”. Warszawa 17 kwietnia 2007. [06.04.2015], <http://eprints.rclis.org/8862>

- Devine, J.; Egger-Sider, F. (2009). *Going Beyond Google. The Invisible Web in Learning & Teaching*. London.
- DIST (2007). *Dictionary of Information Science and Technology*. Hershey, London, Melbourne, Singapore.
- Haglund, L.; Olsson, P. (2008). The Impact on University Libraries of Changes in Information Behavior Among Academic Researchers: A Multiple Case Study. *The Journal of Academic Librarianship*. Vol. 34 (1), 52–29.
- Kulczycki, E. (2013). *Jak dodać prace do Google Scholar i zwiększyć liczbę cytowań oraz indeks Hirscha. Poradnik dla początkujących*. [06.04.2015], <http://ekulczycki.pl/poradnik>
- Lewandowski, D. (2010). Google Scholar as a tool for discovering journal articles in library and information science. *Online Information Review*. Nr 34 (2), s. 250–262.
- Ortega, J. (2014). *Academic Search Engines*. Oxford.
- Rozkosz, E.A. (2014). Open Journal Systems – wartość dodana czasopisma. *Biuletyn EBIB*. Nr 149 [06.04.2015], <http://open.ebib.pl/ojs/index.php/ebib/article/view/246>
- Sherman, Ch.; Price G. (2001). *The Invisible Web. Uncovering Information Sources Search Engines Can't See*. Medford, New Jersey.
- Swan, A. (2010). *The Open Access citation advantage: Studies and results to date*. ePrints Soton [06.04.2015], <http://eprints.soton.ac.uk/268516>

Załącznik

Repozytoria

E-LIS

1. Derfert-Wolf, L.: *Odkrywanie niewidzialnych zasobów sieci*. 2007.
2. Skórka, S.: *Systemy nawigacji w przestrzeni słuchowej. Analiza porównawcza*. 2011.
3. Cisek, S.: *Nauka o informacji na świecie: badania metanaukowe*. 2008.

AMUR

4. Sidor, M.W.: *Elektroniczny system wspomagający zarządzanie zasobami w bibliotece Wyższej Szkoły Biznesu – National-Louis University w Nowym Sączu*. 2010.
5. Chachlikowska, A.: *Badania wykorzystania przez polskie biblioteki naukowe środków europejskich, grantów ministerialnych i samorządowych oraz dotacji sponsorów w latach 2000–2008*. 2009.

CEON

6. Giętkowski, T.: *Zmiany lesistości Borów Tucholskich w latach 1938 – 2000*. 2009.
7. Klimczuk, A.: *Żyjemy razem, czyli metody budowania kapitału społecznego*. 2009.

Biblioteki cyfrowe – artykuły

Zachodniopomorska BC Pomerania: Bibliotekarz Zachodnio-Pomorski

1. 2008, nr 1–2
2. 2010, nr 1

Śląska BC: Bibliotheca Nostra. Śląski Kwartalnik Naukowy

3. 2009, nr 2
4. 2011, nr 1

Kujawsko-Pomorska BC: Poradnik Bibliotekarza

5. 2007, nr 5
6. 2009, nr 3

Bibliologiczna BC: Zagadnienia Informacji Naukowej

7. 2006, nr 2
8. 2008, nr 2

Biblioteki cyfrowe – książki

Śląska BC

1. Tomaszczyk, J.: *Angielsko-polski słownik informacji naukowej i bibliotekoznawstwa*. 2009.
2. Roszkowski, M.: *Język informacyjno-wyszukiwawczy jako narzędzie organizacji informacji w dziedzinowych systemach hipertekstowych*. 2009.

Biblioteka Cyfrowa Politechniki Warszawskiej

3. Płoszajski, G.: *Standardy w procesie digitalizacji obiektów dziedzictwa kulturowego*. 2008.

Biblioteka Cyfrowa UMCS

4. Osiński, Z.: *Biblioteka, książka, informacja i Internet*. 2010.

Dolnośląska BC

5. Leśniewski, D.: *Digitalizacja zasobów bibliotecznych*. 2002.

Strony wydawców

Biuletyn EBIB, nr 1/2012

1. Derfert-Wolf, L.: *Archiwizacja Internetu – wprowadzenie i przegląd wybranych inicjatyw*.
2. Nalewajska, L.: *Archiwizowanie stron internetowych w krajach nordyckich*.

iNFOTEZY, nr 1/2011

3. Uchańska, A.: *Strategie przedsiębiorstwa prasowego w XXI wieku*.
4. Zapała, M.: *Boom komiksowy. Polski rynek historii obrazkowych w latach 2000–2003*.

Biuletyn EBIB w systemie OJS

Biuletyn EBIB, nr 8/2013

1. Nowak, P.: *Funkcje współczesnych czasopism naukowych – pomiędzy wyobrażeniami a rzeczywistością*.
2. Rychlik, M.: *Epoka cyfrowa i jej nowe wskaźniki altmetryczne*.

Biuletyn EBIB, nr 4/2014

3. Drabek, A.: *Tytuł czasopisma i jego funkcja we współczesnym obiegu nauki*.
4. Rozkosz, E.A.: *Open Journal Systems – wartość dodana czasopisma*.

Serwisy społecznościowe dla naukowców

Academia.edu

1. Werla, M.: *Wdrażanie regionalnych bibliotek cyfrowych w sieci PIONIER w oparciu o środowisko dLibra*.
2. Wojciechowska, M.: *Ocena skuteczności działalności biblioteki*.

ResearchGate.net

3. Osiński, Z.: *Architektura informacji polskich internetowych serwisów edukacyjnych*.
4. Cisek, S.: *Metoda analizy i krytyki piśmiennictwa w nauce o informacji i bibliotekoznawstwie w XXI wieku*.

Strony instytucji

PTIN: X Krajowe Forum Informacji Naukowej i Technicznej, 2009

1. Gajos, M.: *Innowacja geoinformacyjna*.
2. Sapa, R.: *Warsztat naukowca a problem formatu informacji bibliograficznej generowanej przez systemy informacyjne*.

Biblioteka Politechniki Poznańskiej (BPP): Informacja dla nauki a świat zasobów cyfrowych, 2008

3. Wiewiórowski, W.: *Zagrożenia związane z zarządzaniem informacją prawną i prawniczą w środowisku elektronicznym*.
4. Gaziński, R.: *Świat informacji na nośnikach elektronicznych a humanista na przykładzie warsztatu historyka*.

Serwisy SlideShare i Scribd

Scribd

1. *Między Regałami*. Pismo Studentów Informatyki i Bibliotekoznawstwa. 2009.
2. Jaskowska, B.: *Broker informacji – zawód prz(e)y szłości*. 2011 (prezentacja).

SlideShare

3. Cisek, S.: *Dzielenie się wiedzą*. 2009.
 4. Skórka, S.: *Architektura informacji. Dziedzina wiedzy czy rzemiosło?* 2007.
-

Visibility of Polish Scholarly Publications on the Internet

Abstract

Purpose/thesis: The aim of the study is to determine to what extent the Internet location of scholarly publications influence the visibility of those documents within general and specialized search engines. The additional objective is to identify changes that have occurred in the visibility of Polish scholarly papers between 2012 and 2014.

Approach/methods: The study involved the analysis of selected publications on information science, located in repositories, digital libraries and on the websites of scientific journals and institutions. The visibility of selected documents was determined by the possibility of finding them in some general and specialized search engines. The comparison of the results from 2012 and 2014 resulted in conclusions on the changes in the visibility of the documents analyzed. The background of the research is the „invisible web” issue.

Results and conclusions: As regards the visibility the repositories and electronic papers available on the Open Journal Systems platform produced the best results. Compared to 2012 a considerable improvement in the visibility of Polish scholarly resources was observed. Google repeated excellent results from 2012, while Bing improved indexing and got closer to the leader.

Originality/value: The results indicate the best locations for publishing scholarly papers on the Internet in terms of their visibility to potential users. This will help researchers choose the optimal location for their publications.

Keywords

Electronic journals. Digital libraries. Invisible Web. Open Access. Open Science. Repositories. Academic social networks. Visibility of academic resources. Information retrieval. Scholarly search engines.

Dr ARKADIUSZ PULIKOWSKI jest pracownikiem naukowym Zakładu Zarządzania Informacją w Instytucie Bibliotekoznawstwa i Informatyki Uniwersytetu Śląskiego. W swoich badaniach skupia się na zagadnieniach związanych z wyszukiwaniem, przetwarzaniem i udostępnianiem informacji. Wraz z dr Anetą Drabek tworzy bazę Arianta – Naukowe i branżowe polskie czasopisma elektroniczne.

Kontakt z autorem:

arkadiusz.pulikowski@us.edu.pl

Instytut Bibliotekoznawstwa i Informatyki

Uniwersytet Śląski

Plac Sejmu Śląskiego 1

40-032 Katowice