

Hanna Celoch
Biblioteka Główna Politechniki Lubelskiej

Bibliograficzne bazy danych szybką ścieżką w obiegu informacji naukowej

Abstrakt

Bibliograficzne bazy danych od lat są stałym elementem warsztatu pracy naukowców. Wbrew pozorom, pełnotekstowe bazy danych czy biblioteki cyfrowe nie wyparły z obiegu informacji naukowej baz bibliograficznych – wprost przeciwnie – wykorzystały ich potencjał jako swoistą trampolinę, by szybciej dotrzeć do odbiorcy. Bazy danych stały się czynnikiem dynamizującym obieg informacji. Artykuł jest próbą prześledzenia niektórych ścieżek w tym obiegu, a zwłaszcza tych najpopularniejszych, czyli wyszukiwarek internetowych i obecności w nich bibliograficznych baz danych. Podjęto również próbę określenia wpływu tej obecności na zwiększenie cytawalności polskich technicznych czasopism naukowych (na przykładzie bazy BazTech).

Słowa kluczowe

bibliograficzne bazy danych, obieg informacji, wyszukiwarki

Abstract

Bibliographic databases: the fast path to circulating scientific information. Bibliographic databases traditionally have been a regular feature of scientific work. Despite appearances, full-text databases, or digital libraries, have not displaced bibliographic databases from the circulation of scientific information. On the contrary, bibliographic databases have made use of their potential as a kind of springboard to more quickly reach the recipient. Databases have become a dynamic factor in the circulation of information. This paper is an attempt to track down some paths in this circuit, especially the most popular, that is, search engines and the presence in them of bibliographic databases. Efforts were also made to try to determine the impact of this presence, to increase the citations of Polish technical scientific journals (using the example of the BazTech base).

Keywords

bibliographic databases, circulating scientific information, search engines

Uprawianie nauki, to posługiwanie się informacją – pozyskiwanie, tworzenie i rozpowszechnianie. Uczony, prowadząc badania, musi zapoznać się z wynikami innych badaczy, a później, po opracowaniu sprawozdania z własnych badań, czyli wytworzeniu informacji, musi ją rozpowszechnić, aby inni badacze mogli ją poznać i ustosunkować się do niej. Ten obieg informacji w nauce musi następować szybko – to czas jest jednym z czynników stymulujących postęp w nauce.

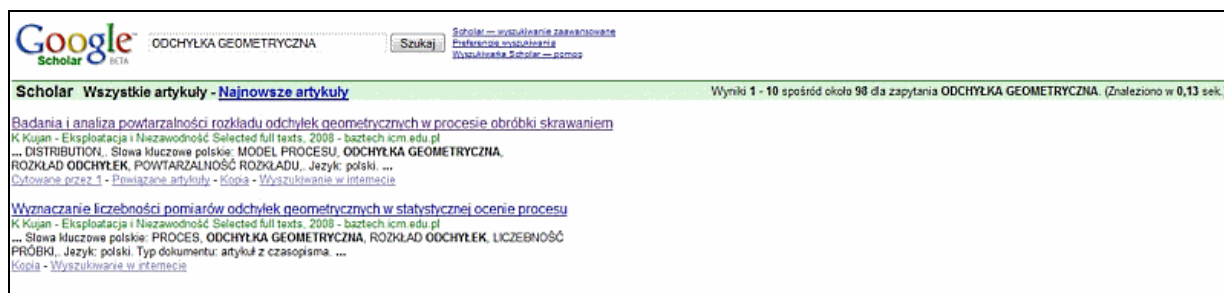
Szybkość w pozyskiwaniu informacji mogą zapewnić badaczom bibliograficzne bazy danych – to najprostszy sposób na zapoznanie się z literaturą przedmiotu, niekoniecznie dostępną w pełnym tekście. Wystarczy się zorientować, kto, kiedy i w jakim zakresie prowadził badania,

aby móc przystąpić do swoich, bez obawy, że po raz kolejny badamy coś, co już zostało dziesiątki razy zbadane. Uczony pragnie porównać wyniki swoich badań z innymi, przedyskutować na łamach pism naukowych. To bazy bibliograficzne, rejestrujące również cytowania, dają możliwość natychmiastowego dotarcia do tych artykułów. Chcemy wiedzieć, kto się odniósł do naszych badań? Musimy prześledzić bazy cytowań. Bazy bibliograficzne są tutaj nieocenione – uczeni znają je i wykorzystują. Już dawno mamy za sobą okres zeszytów „Current Contents” czy „Bibliografii Zawartości Czasopism” – wydawcy zabiegają o użytkownika, zaopatrując swe bazy w coraz to nowsze i lepsze narzędzia do wyszukiwania i analizy danych.

Jednak, aby baza zaistniała w sposób trwały na rynku i funkcjonowała niezależnie od tego, czy jest polecana przez pracowników informacji, czy nie, musi być dostępna w Internecie w tych miejscach, gdzie użytkownicy spodziewają się najszybciej znaleźć informację. Doświadczony użytkownik, badacz ma zwykle swoje miejsca w sieci, swoje źródła, bazy, które zna i przegląda, gdyż wie, że indeksują te źródła, które mogą go zainteresować. Ale niewyrobiony jeszcze w poszukiwaniach student czy uczeń nie trafi od razu do takich specjalistycznych źródeł. Nie rozróżnia bowiem podstawowych narzędzi wyszukiwawczych, takich jak: wyszukiwarki, katalogi tematyczne czy multiwyszukiwarki. Nie potrafi również zastosować odpowiednich strategii wyszukiwawczych, które wymagają jednak pewnego przygotowania. Taki użytkownik nastawiony jest na szybkie uzyskanie informacji, nawet kosztem jej jakości. Zacznie zawsze od źródła najprostszego, najbardziej znanego i dostępnego.

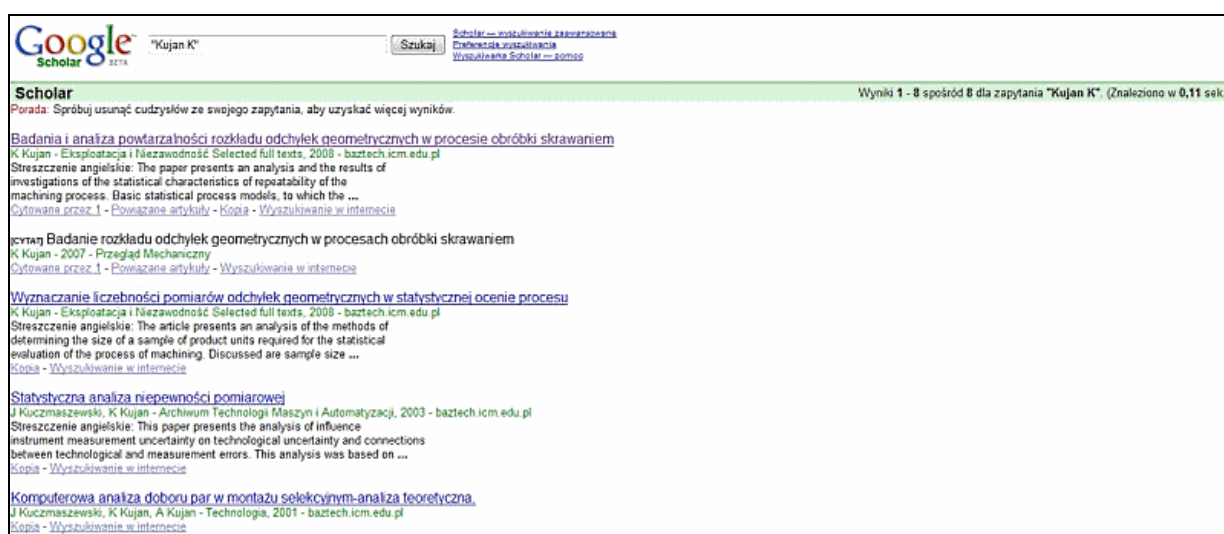
Prześledźmy drogę poszukiwań literatury przez studenta. Jest prawie pewne, że pierwszym adresem, jaki wpisze w przeglądarce internetowej, będzie <http://www.google.pl>. Prawie pewne też jest, że nie będzie szukał hasła „bazy dziedzinowe” czy „bazy bibliograficzne”, ale od razu wpisze hasło, które go interesuje i będzie oczekiwał rezultatu, który pozwoli mu od razu dotrzeć do źródła. Jeśli istnieje źródło, gdzie poszukiwana literatura jest w wersji pełnotekstowej, to oczywiście pełny sukces, można od razu je wykorzystać, ale co z literaturą przedmiotu dostępną tylko z półek w magazynach bibliotek? Co z całym „niewidocznym” – Ukrytym Internetem? Są narzędzia służące do odkrywania tych zasobów – są to katalogi, bazy dziedzinowe, portale, a jednak najczęściej student trafia do nich poprzez Google, bo tak jest mu najłatwiej. Nie musi znać adresu katalogu, nie musi umieć wyszukiwać za pomocą specjalnie skonstruowanych zapytań. Po prostu wpisuje hasło. Jeśli wynik wyszukiwania jest niezadowolający, uważa, że literatury nie ma. Do oddziałów informacji trafiają studenci w poszukiwaniu materiałów, twierdząc, że „nigdzie tego nie ma”. Nigdzie, to znaczy w Google. Naszym zadaniem jest wtedy wskazać źródła – bazy, katalogi – gdzie wyszukiwanie skieruje do źródła. Często tym źródłem będzie właśnie czasopismo stojące na półce w bibliotece, często we własnej bibliotece, tuż obok. Obecność baz danych w serwisach Google, to przejście tej roli – wynik wyszukiwania nie daje pełnego tekstu, ale pokazuje miejsce, gdzie jest dokument – konkretną bibliotekę. Serwis Google może to zrobić tylko za pośrednictwem bazy danych. Idąc dalej tym tropem, bazy danych kierują do pełnego tekstu, niezależnie od tego, gdzie ten tekst się znajduje: na stronie w sieci, w repozytorium czy na półce w bibliotece.

Badając tę drogę poszukiwań na konkretnym przykładzie czasopisma indeksowanego przez bazę BazTech widać, że pomimo iż w sieci są dostępne pełne teksty artykułów, to najprostsza droga do nich wiedzie właśnie przez: Google Scholar (<http://scholar.google.com/>) -> BazTech -> stronę wydawcy i bibliotekę cyfrową.



Rys. 1. Wyszukiwanie w Google Scholar według wyrażenia z artykułu z czasopisma indeksowanego przez bazę BazTech
Źródło: Google Scholar.

czy autora:



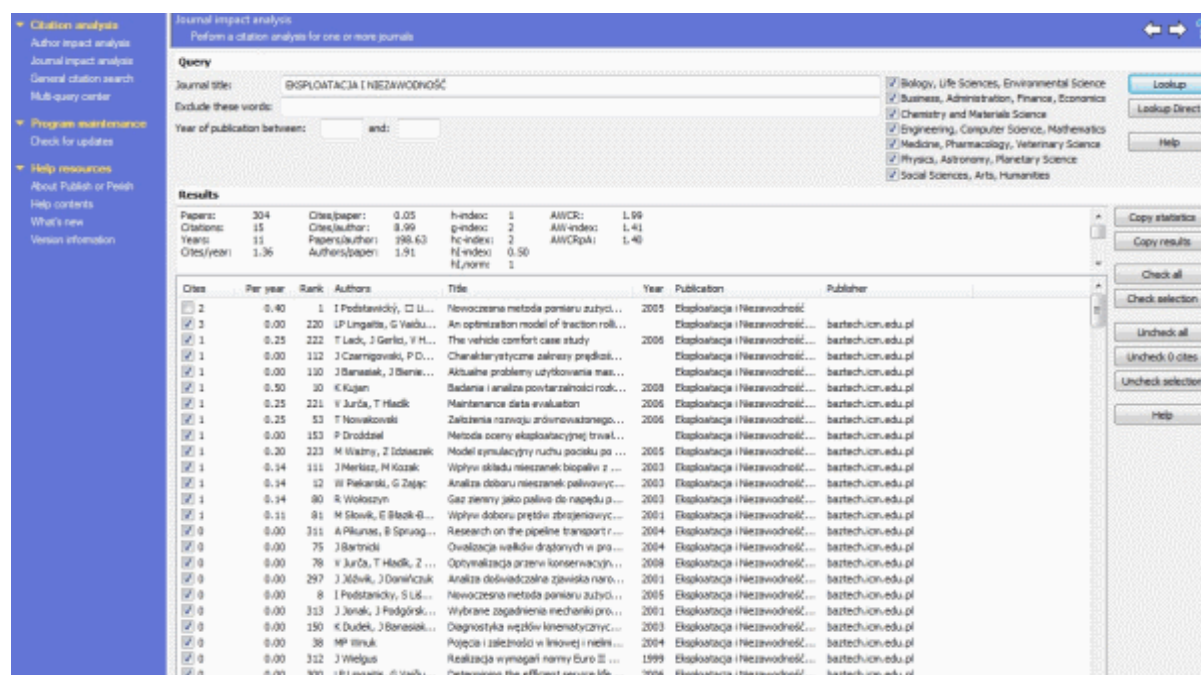
Rys. 2. Wyszukiwanie w Google Scholar według nazwiska autora artykułu z czasopisma indeksowanego przez bazę BazTech
Źródło: Google Scholar.

... to wynikiem wyszukiwania będą rekordy z bazy BazTech.

Wyszukiwanie po tytułach czasopism indeksowanych przez BazTech daje rezultaty jeszcze bardziej jednoznaczne. Prawie wszystkie kierują do bazy. Nieliczne są wskazania stron wydawcy czy bibliotek cyfrowych, nawet tych czasopism, które mają pełne teksty zarówno na stronie wydawcy, jak i w bibliotece cyfrowej. Przykładem takiego czasopisma jest „Eksplotacja i Niezawodność”. Od początku istnienia czasopismo udostępnia teksty na swojej stronie internetowej (<http://darmaz.pollub.pl>, potem [http://www.ein.org.pl/\[1\]](http://www.ein.org.pl/[1])), od ubiegłego roku w Bibliotece Cyfrowej Politechniki Lubelskiej. Czasopismo jest indeksowane przez BazTech od 1999 r., a od ubiegłego roku również przez bazę SCI Expanded.

Jeśli w Google Scholar wpisujemy jako hasło wyszukiwawcze tytuł czasopisma: Eksplotacja i Niezawodność, to na 313 znalezionych rekordów 304 to rekordy BazTech-u, czyli ponad 97%. Do szczegółowego przeglądu rezultatów użyto programu Publish or Perish (PoP) <http://www.harzing.com/pop.htm> (rys. 3.). PoP umożliwia analizę publikacji i cytowań wyszukanych w Google Scholar. Aplikacja dostarcza wielu statystyk, np.: liczba artykułów,

liczba cytowań, średnia cytowań na artykuł, średnia cytowań na rok, h-indeks Hirscha, g-indeks Egghego, AW-indeks.



Rys. 3. Analiza wyszukiwania według tytułu czasopisma: Eksploatacja i Niezawodność w Google Scholar, wykonana za pomocą programu Publish or Perish
Źródło: opracowanie własne na podstawie Google Scholar.

Podobnie wygląda wyszukiwanie innych tytułów czasopism indeksowanych przez BazTech:

Tab. 1. Analiza rezultatów wyszukiwania w BazTech-u oraz Google Scholar według wybranych tytułów czasopism

Czasopismo	Liczba rekordów w BazTech	Liczba rekordów w Google Scholar	Liczba rekordów BazTech w Google Scholar	Procent
Eksploatacja i Niezawodność	374	313	304	97
Rynek Energii	535	430	385	89,5
Hutnik. Wiadomości Hutnicze	1016	929	924	99
Logistyka	2609	2330	2150	92
Inżynieria Materiałowa	2208	2160	2150	99,5

Źródło: opracowanie własne.

Przykłady można by mnożyć, i zawsze liczba przekierowań do bazy BazTech będzie największa, bez względu na to, czy będziemy szukać za pomocą słów kluczowych, autora czy tytułu czasopisma. Statystyki wykorzystania bazy również wskazują na to, że czytelnicy trafiają do BazTech-u głównie z serwisów Google (rys. 4.).

Pierwsze 20 z ogólnej liczby 1328 odnośników			
#	Wywołania		Odnośniki
1	237766	42.36%	http://baztech.icm.edu.pl/ *
2	90486	16.12%	http://www.google.pl/search
3	11146	1.99%	http://www.google.com/search
4	7174	1.28%	http://scholar.google.pl/scholar
5	981	0.17%	http://www.bg.polsl.pl/index.php
6	969	0.17%	http://szukaj.onet.pl/query.html
7	909	0.16%	http://scholar.google.com/scholar
8	837	0.15%	http://www.bg.pw.edu.pl/listwa.html
9	597	0.11%	http://www.bg.agh.edu.pl/PL/index.php
10	438	0.08%	http://biblioteka.pollub.pl/
11	422	0.08%	http://baztech.icm.edu.pl
12	408	0.07%	http://www.google.de/search
13	348	0.06%	http://www.seo-linki.pl
14	328	0.06%	http://www.google.co.uk/search
15	324	0.06%	http://www.bg.po.opole.pl/index.php
16	321	0.06%	http://www.prz.rzeszow.pl/biblio/bazy.htm
17	315	0.06%	http://www.ml.put.poznan.pl/pl/1_2_1.html
18	299	0.05%	http://cachas-follando-o9d.blogspot.com/
19	253	0.05%	http://www.google.com/custom
20	250	0.04%	http://209.85.129.132/search

Rys. 4. Statystyka odwiedzin BazTech-u według witryn i serwisów, w których wyszukano rekordy z bazy

Źródło: opracowanie własne na podstawie bazy BazTech.

Widać wyraźnie przewagę serwisów nad innymi stronami. Każda z bibliotek Konsorcjum BazTech ma na swej stronie ikonę albo bezpośredni link do bazy. Promujemy bazę, gdzie tylko się da, a jednak odsetek wejść z naszych stron oscyluje w granicach 0,16-0,06% – to jest naprawdę znikoma liczba w porównaniu z Google.

Wydaje się to paradoksalne: tworzymy z wysiłkiem bazy i katalogi, a użytkownicy niekiedy nie znają ich adresu i nie mają świadomości, z czego korzystają. Często się zdarza, że użytkownik, prosząc o pomoc w wydrukowaniu artykułu, mówi: znalazłem to w Google. Otóż nie w Google, tylko np. w bibliotece cyfrowej, gdzie skrzętnie digitalizujemy nasze zasoby, ale nie mamy aż tyle siły przebicia, żeby nasze katalogi i bazy przebiły popularnością serwisy Google. Skoro nie możemy odwrócić tej kolejności, zadbajmy o to, aby nasze bazy i katalogi były obecne w wyszukiwarkach internetowych. Nie traktujmy tych wyszukiwarek jak konkurencji, zresztą nie mamy nawet szans na konkurencję. Wykorzystajmy ich popularność i nowe możliwości do pracy zarówno przez nas samych, jak i naszych czytelników. Zadbajmy, aby na szkoleniach dla naszych studentów i pracowników

naukowych uczyć, jak łatwiej i efektywniej znajdować odpowiedź, korzystając zarówno z wyszukiwarek internetowych, jak i baz danych dostępnych w bibliotece. Serwisy Google nas nie zastąpią – nadal będzie istnieć zakres działań, który będzie domeną tylko i wyłącznie bibliotekarzy.

Najczęstszym pytaniem, prawie zawsze pierwszym na spotkaniach prezentujących bibliograficzne bazy danych, jest: Czy baza umożliwi dostęp do pełnych tekstów? Trudno już dzisiaj znaleźć bazę, która nie linkowałaby pełnych tekstów. Większość rekordów w tych bazach ma opcję „fulltext” – oczywiście skorzystać z tego ułatwienia będą mogli tylko ci użytkownicy, których macierzyste biblioteki mają wykupiony dostęp do baz pełnotekstowych, ale przekierowania będą następowały również do pełnych tekstów baz open access. Tutaj również zaciera się dla przeciętnego czytelnika granica między bazami – czy jeszcze korzysta z bibliograficznej, czy już z pełnotekstowej, bo niekiedy przyciski „fulltext” łagodnie przeprowadzą czytelnika z jednej bazy do drugiej, niekiedy bez pośrednich ekranów, a korzystający nie zawsze obserwuje pasek adresu. Mniej wyrobiony czytelnik będzie nam udowadniał, że w bazie np. „ISI Web of Knowledge” są pełne teksty, bo korzystał z nich w bibliotece. Czasem ma do tych baz ustawiony dostęp również w domu, więc bez problemu w ten sposób korzysta. Śmiało więc można stwierdzić, że statystyki wykorzystania baz pełnotekstowych w dużej mierze generują również bazy dziedzinowe – bo rzadko zdarza się, że serwisy Google wskażą bazy pełnotekstowe, głównie z tego powodu, że niektórzy wydawcy nie zezwalają na indeksowanie swoich czasopism.

Trudno nam jednak porównywać się do czołowych wydawców. Nasze czasopisma, zwłaszcza techniczne, nie mają zbyt wielkich wskaźników cytowań. Musimy je najpierw rozpowszechnić, aby zwiększyć ich cytowalność. Serwis Google Scholar pozwala śledzić również cytowania. Dwa czasopisma, o których wspomniano wcześniej – „Eksploracja i Niezawodność” oraz „Rynek Energii”, weszły w ubiegłym roku na listę filadelfijską. Porównanie cytowań obu czasopism w obu źródłach wygląda następująco:

„Rynek Energii” – cytowania w Google Scholar:

Query					
Author(s):					
Publication:	Rynek Energii				
All of the words:					
Any of the words:					
None of the words:					
The phrase:					
Year of publication between:	1998	and:	2009		
Results					
Papers:	403	Cites/paper:	0.13	h-index:	2
Citations:	51	Cites/author:	40.95	g-index:	3
Years:	12	Papers/author:	333.12	hc-index:	2
Cites/year:	4.25	Authors/paper:	1.43	hI-index:	2.00
				hI,norm:	2
				AWCR:	8.31
				AW-index:	2.88
				AWCRpA:	6.76

Rys. 5. Analiza wyszukiwania według tytułu czasopisma „Rynek Energii” w Google Scholar, wykonana za pomocą programu Publish or Perish
Źródło: opracowanie własne.

Cytowania w bazie „ISI Web of Knowledge”:

- Results found: 119 Sum of the Times Cited: 3
- Average Citations per Item: 0.03 h-index: 1

Baza ISI zawiera wszystkie artykuły z roczników, ale notuje tylko trzy cytowania, a Google Scholar, pomimo że znajduje mniej artykułów w konkretnych rocznikach, to notuje znacznie więcej cytowań.

„Eksplotacja i Niezawodność”:

The screenshot shows the 'Query' section of the Publish or Perish software. The 'Publication' field is filled with 'Eksplotacja i Niezawodność'. The 'Year of publication between' is set from 1998 to 2009. The 'Results' section displays a table of metrics:

Results					
Papers:	271	Cites/paper:	0.07	h-index:	2
Citations:	18	Cites/author:	12.50	g-index:	2
Years:	11	Papers/author:	180.77	hc-index:	2
Cites/year:	1.64	Authors/paper:	1.86	hI-index:	1.00
				hI,norm:	1
				AWCR:	4.18
				AW-index:	2.04
				AWCRpA:	2.90

Rys. 6. Analiza wyszukiwania według tytułu czasopisma „Rynek Energii” w Google Scholar, wykonana za pomocą programu Publish or Perish
Źródło: opracowanie własne.

I analogiczne wyniki z bazy ISI:

- Results found: 58 Sum of the Times Cited: 1
- Average Citations per Item: 0.02 h-index: 1

Nie są to imponujące wyniki, zważywszy na liczbę tych cytowań, ale oba czasopisma od niedawna są indeksowane w bazie ISI, więc brak obszernego materiału porównawczego; nasze czasopisma techniczne dopiero wchodzą do światowych baz danych – do tej pory tylko nieliczne z nich były tam indeksowane. Przykładem czasopisma, które jest indeksowane od początku istnienia bazy BazTech, jest czasopismo „Polimery” – wyniki wyszukiwania i porównań będą tutaj zupełnie inne – tytuł indeksowany jest również przez kilka światowych bibliograficznych baz danych, na czele z bazą „ISI Web of Knowledge”.

Coraz częściej też trafiamy na cytowania bezpośrednio do rekordów bazy, a nie do samego źródła, co świadczy o tym, że sama baza staje się źródłem, pomimo że nie jest bazą pełnotekstową. Poniżej kilka przykładów:

As you've probably experienced, programming languages are always a good topic to sit around and chew the FAT on. (That's "FAT" as in the "FAT programming language"; see baztech.icm.edu.pl/baztech/cgi-bin/btgetdoc.cgi?BPS2-0003-0033). And a good chew we had at the SD West 2008 panel on *Beautiful Code*, a book edited by O'Reilly's Andy Oram and Dr. Dobb's contributing editor Greg Wilson. You

Rys. 7. Opis artykułu w BazTech-u jako cytowanie w tekście, w bazie pełnotekstowej
 Źródło: ERICKSON, J. Was George Konstanza a Computer Programmer? W: *nxtbook media* [on-line]. [Dostęp 9 maja 2009]. Dostępny w World Wide Web:
<http://www.nxtbook.com/nxtbooks/cmp/ddj0508/index.php?startid=6>).

As you've probably experienced, **programming languages** are always a good topic to sit around and chew the FAT on. (That's "FAT" as in the "FAT programming language"; see baztech.icm.edu.pl/baztech/cgi-bin/btgetdoc.cgi?BPS2-0003-0033). And a good chew we had at the SD West 2008 panel on *Beautiful Code*, a book edited by O'Reilly's Andy Oram

Rys. 8. Opis artykułu w BazTech-u jako cytowanie w tekście, w bazie pełnotekstowej
 Źródło: ERICKSON, J. Was George Konstanza a Computer Programmer? W: *Dr Dobb's* [on-line]. [Dostęp 9 maja 2009]. Dostępny w World Wide Web:
<http://www.ddj.com/architect/207400254>.

[Saxon Garden - Wikipedia, the free encyclopedia](#)
 The Saxon Garden was originally the site of Warsaw ... <http://baztech.icm.edu.pl/baztech/cgi-bin/btgetdoc.cgi?BSW1-0032-0003>. Retrieved on 2008-02-13. ...
en.wikipedia.org/wiki/Saxon_Garden
 More pages from en.wikipedia.org

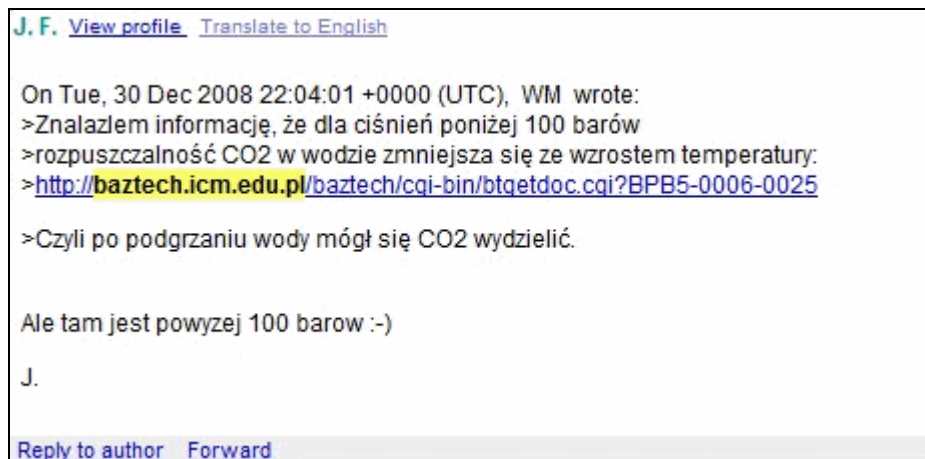
Rys. 9. Opis artykułu w BazTech-u jako cytowanie w Wikipedii
 Źródło: opracowanie własne na podstawie Wikipedii.

■ Saxon Palace

Notes and references

1. ^{a b} (Polish) "Ogród Saski". *ePrzewodnik*. Retrieved on 2008-02-10.
2. ^a (English) *nathansvilla.com* accessed 19 February 2008
3. ^{a b c d e f g h i j k} (Polish) Stefan Kieniewicz, ed (1984). *Warszawa w latach 1526-1795*. Warsaw. ISBN 83-01033-23-1.
4. ^a (English) "Events". Welcome to Warsaw. Retrieved on 2008-02-10.
5. ^{a b c d e f g} (Polish) P. Giergoń. "Ogród Saski". *sztuka.net*. Retrieved on 2008-02-13.
6. ^a (Polish) A. Franta. "O Placu Piłsudskiego, tożsamości i ładzie". *icm.edu.pl*. Retrieved on 2008-02-13.
7. ^{a b} (Polish) "Ministerstwo Spraw Zagranicznych". *um.warszawa.pl*. Retrieved on 2008-02-13.
8. ^a (Polish) "Polski Teatr w XVII i XVIII w.". <http://baztech.icm.edu.pl/baztech/cgi-bin/btgetdoc.cgi?BSW1-0032-0003>
9. ^{a b} (Polish) "Blue (Zamoyski) Palace". *um.warszawa.pl*. Retrieved on 2008-02-13.
10. ^{a b} (English) "St. Anthony of Padua and Reformed Franciscan Monastery Church". *um.warszawa.pl*. Retrieved on 2008-02-13.
11. ^a (Polish) "Ogród Saski". *State Archive of the Capital City of Warsaw*. Retrieved on 2008-02-13.
12. ^a (English) Katarzyna Murawska (October 2000). "To the Glory of the Unknown". *polishnews.com*. Retrieved on 2008-02-14.
13. ^a (English) "The Tomb of the Unknown Soldier". *eGuide*. Retrieved on 2008-02-14.
14. ^a (Polish) "Fontanna". *warszawa1939.pl*. Retrieved on 2008-02-14.
15. ^{a b} (Polish) "Wodociąg wodociągu Marconiego". *warszawa1939.pl*. Retrieved on 2008-02-14.

Rys. 10. Opis artykułu w BazTech-u jako cytowanie w Wikipedii
 Źródło: opracowanie własne na podstawie Wikipedii.



Rys. 11. Cytowania i odsyłacze artykułów w BazTech-u na listach dyskusyjnych
Źródło: *Google grupy dyskusyjne* [on-line]. [Dostęp 9 maja 2009]. Dostępny w World Wide Web: <http://groups.google.pl/group/pl.sci.fizyka/>.

Poza serwisami Google, rekordy BazTech-u są widoczne w serwisie Yahoo (<http://yahoo.com>), Lycos (<http://www.lycos.com/>), Hotbot (<http://www.hotbot.com/>), MSN i wielu innych nie mniej popularnych wyszukiwarkach. Sprawia to, że dostęp do naszych polskich czasopism technicznych staje się łatwiejszy i coraz częściej są one cytowane, co jest nie bez znaczenia dla wzrostu wskaźników, a co za tym idzie, wyższej punktacji w rankingach oraz ich popularności.

Przypis

[1] Odesłania do stron internetowych przedstawiają wersję aktualną w dniu 9 maja 2009 r.

Bibliograficzne bazy danych szybką ścieżką w obiegu informacji naukowej / Hanna Celoch // W:
Bibliograficzne bazy danych : kierunki rozwoju i możliwości współpracy. Bydgoszcz, 27-29 maja 2009. -
[Warszawa] : Stowarzyszenie Bibliotekarzy Polskich, K[omisja] W[ydawnictw] E[lektronicznych], Redakcja
"Elektronicznej Biblioteki", 2009. - (EBIB Materiały konferencyjne nr 19). - ISBN: 83-921757-5-1. -
Tryb dostępu : <http://www.ebib.info/publikacje/matkonf/mat19/celoch.php>