

Jak zwiększyć widoczność publikacji naukowych w Internecie z pomocą Google Scholar

Tomasz Lewandowski

Platforma Otwartej Nauki, ICM, UW

„Otwórz, aby inspirować”, Białystok, 21.11.2017

Co nas czeka?

Co nas czeka?

1. Google Scholar jako narzędzie rozpowszechniania publikacji naukowych.

Co nas czeka?

1. Google Scholar jako narzędzie rozpowszechniania publikacji naukowych.
 - a) Jak GS gromadzi dane o publikacjach.

Co nas czeka?

1. Google Scholar jako narzędzie rozpowszechniania publikacji naukowych.
 - a) Jak GS gromadzi dane o publikacjach.
 - b) Jak GS prezentuje dane.

Co nas czeka?

1. Google Scholar jako narzędzie rozpowszechniania publikacji naukowych.
 - a) Jak GS gromadzi dane o publikacjach.
 - b) Jak GS prezentuje dane.
 - c) Najważniejsze zalety i wady GS.

Co nas czeka?

1. Google Scholar jako narzędzie rozpowszechniania publikacji naukowych.
 - a) Jak GS gromadzi dane o publikacjach.
 - b) Jak GS prezentuje dane.
 - c) Najważniejsze zalety i wady GS.
2. Repozytorium akademickie jako pomoc w dostarczeniu danych do Google Scholar.

Co nas czeka?

1. Google Scholar jako narzędzie rozpowszechniania publikacji naukowych.
 - a) Jak GS gromadzi dane o publikacjach.
 - b) Jak GS prezentuje dane.
 - c) Najważniejsze zalety i wady GS.
2. Repozytorium akademickie jako pomoc w dostarczeniu danych do Google Scholar.
3. Co my sami możemy zrobić, by pomóc.

1. Google Scholar jako narzędzie rozpowszechniania publikacji naukowych

1. Google Scholar jako narzędzie rozpowszechniania publikacji naukowych

a) Konstrukcja wyszukiwarki.

Konstrukcja wyszukiwarki

Konstrukcja wyszukiwarki

Crawler – robot przeszukujący Internet

Konstrukcja wyszukiwarki

Crawler – robot przeszukujący Internet

Przeogląda strony internetowe i analizuje ich zawartość.

Konstrukcja wyszukiwarki

Crawler – robot przeszukujący Internet

Przeogląda strony internetowe i analizuje ich zawartość.

Porusza się między nimi podążając za linkami.

Konstrukcja wyszukiwarki

Crawler – robot przeszukujący Internet

Przegląda strony internetowe i analizuje ich zawartość.

Porusza się między nimi podążając za linkami.

Pieszczotliwie nazywany „pajęczkiem”.

Konstrukcja wyszukiwarki

Crawler – robot przeszukujący Internet

Przegląda strony internetowe i analizuje ich zawartość.

Porusza się między nimi podążając za linkami.

Pieszczotliwie nazywany „pajęczkiem”.



Konstrukcja wyszukiwarki

Crawler – robot przeszukujący Internet

Indeks – spis stron internetowych przygotowywanych
przez flotę crawlerów

Konstrukcja wyszukiwarki

Crawler – robot przeszukujący Internet

Indeks – spis stron internetowych przygotowywanych
przez flotę crawlerów

Ranking – dowolna hierarchia na zawartości indeksu

Konstrukcja wyszukiwarki

Google Scholar ma trzy źródła indeksowanych stron.

Konstrukcja wyszukiwarki

Google Scholar ma trzy źródła indeksowanych stron.

1. **Harvestery.**

Konstrukcja wyszukiwarki

Google Scholar ma trzy źródła indeksowanych stron.

1. **Harvestery.**

Google decyduje, dla kogo ustawia.

Konstrukcja wyszukiwarki

Google Scholar ma trzy źródła indeksowanych stron.

1. **Harvestery.**

Google decyduje, dla kogo ustawia.

2. **Konta naukowców.**

Konstrukcja wyszukiwarki

Google Scholar ma trzy źródła indeksowanych stron.

1. **Harvestery.**

Google decyduje, dla kogo ustawia.

2. **Konta naukowców.**

Warto o nie zadbać.

Konstrukcja wyszukiwarki

Google Scholar ma trzy źródła indeksowanych stron.

1. **Harvestery.**

Google decyduje, dla kogo ustawia.

2. **Konta naukowców.**

Warto o nie zadbać.

3. **Crawlery.**

Konstrukcja wyszukiwarki

Google Scholar ma trzy źródła indeksowanych stron.

1. **Harvestery.**

Google decyduje, dla kogo ustawia.

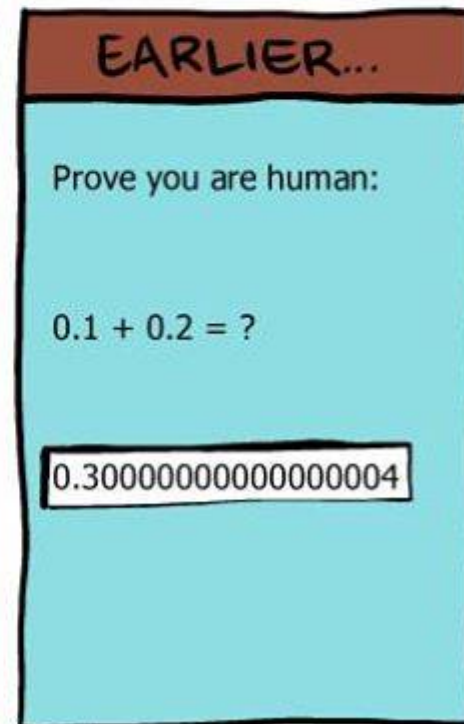
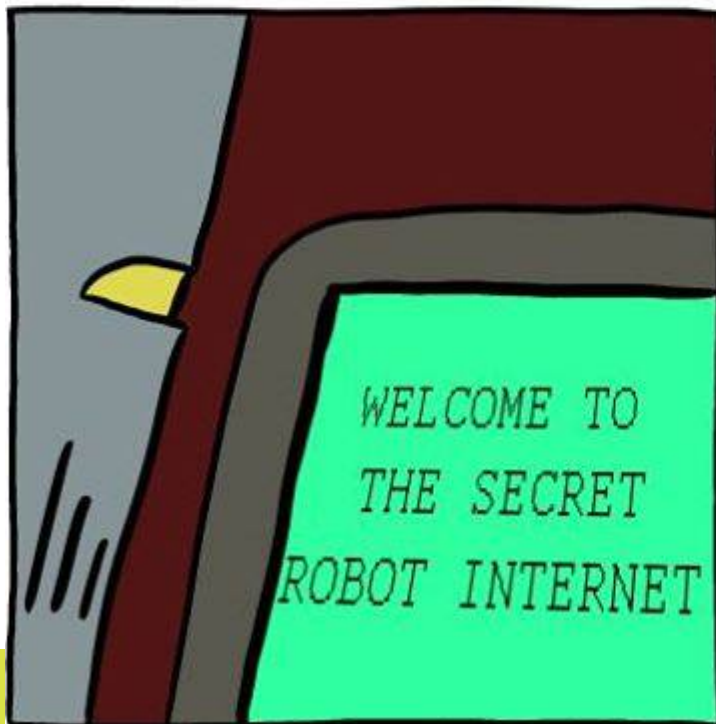
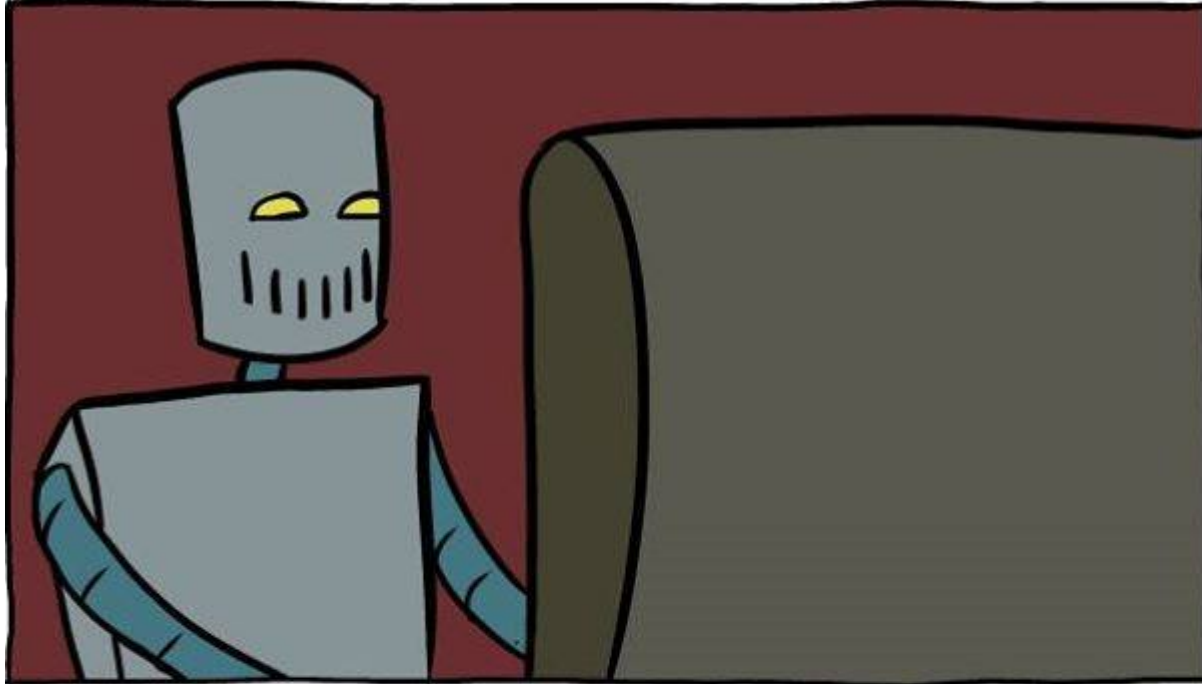
2. **Konta naukowców.**

Warto o nie zadbać.

3. **Crawlery.**

Główne źródło danych.

Internet z punktu widzenia crawlera



HELP

Which program are you using for which you want online help?

Line Mode Browser[1] Using the line mode browser, select an option with a number and press RETURN

Viola For help on Viola, just press the question-mark (?) icon near the top right-hand corner.

NeXTStep[2] To select help with this browser, use Help from the Info menu or select this option with the mouse (double click).

MidasWWW[3] For help on this browser, select this option with the mouse.

See also the W3 server user manual[4]. For other questions, browse from the W3 project page[5].

[End]

Internet z punktu widzenia crawlera

Internet z punktu widzenia crawlera

- crawler wchodzi na stronę główną

Internet z punktu widzenia crawlera

- crawler wchodzi na stronę główną
- czyta *robots.txt* i stosuje się do niego

Internet z punktu widzenia crawlera

- crawler wchodzi na stronę główną
- czyta *robots.txt* i stosuje się do niego
- jeśli jest w repozytorium, znajduje *browse*

Internet z punktu widzenia crawlera

- crawler wchodzi na stronę główną
- czyta *robots.txt* i stosuje się do niego
- jeśli jest w repozytorium, znajduje *browse*
- jeśli to nie repo, musi skanować cały serwis

Internet z punktu widzenia crawlera

- crawler wchodzi na stronę główną
- czyta *robots.txt* i stosuje się do niego
- jeśli jest w repozytorium, znajduje *browse*
- jeśli to nie repo, musi skanować cały serwis
(zazwyczaj pomaga sobie mapą strony)

Internet z punktu widzenia crawlera

- crawler wchodzi na stronę główną
 - czyta *robots.txt* i stosuje się do niego
 - jeśli jest w repozytorium, znajduje *browse*
 - jeśli to nie repo, musi skanować cały serwis
- (porozumiewa się z serwerem tylko metodą GET)

Internet z punktu widzenia crawlera

- crawler wchodzi na stronę główną
 - czyta *robots.txt* i stosuje się do niego
 - jeśli jest w repozytorium, znajduje *browse*
 - jeśli to nie repo, musi skanować cały serwis
- (porozumiewa się z serwerem tylko metodą GET)
(wszystkie formularze, logowania itp.
wykorzystują metodę POST)

Internet z punktu widzenia crawlera

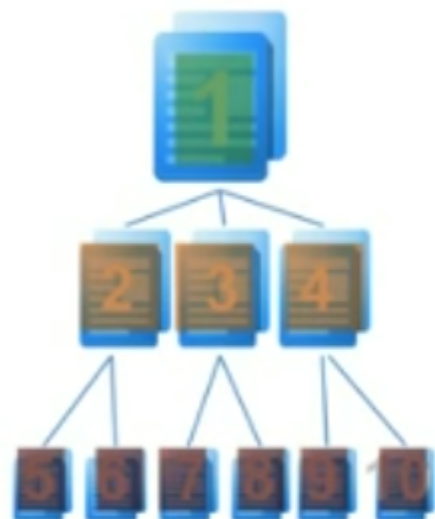
- crawler wchodzi na stronę główną
- czyta *robots.txt* i stosuje się do niego
- jeśli jest w repozytorium, znajduje *browse*
- jeśli to nie repo, musi skanować cały serwis
- artykuły to liście drzewa, po którym się wspina

Internet z punktu widzenia crawlera

- crawler wchodzi na stronę główną
- czyta *robots.txt* i stosuje się do niego
- jeśli jest w repozytorium, znajduje *browse*
- jeśli to nie repo, musi skanować cały serwis
- artykuły to liście drzewa, po którym się wspina
- kształt drzewa zależy od segregacji zasobów

Best practice: Year-month list of articles

Because crawlers crawl the entire web at the same time, they need to reach all URLs quickly & easily



Fast Browse



S..l..o..w Browse

I love it when I can
crawl quickly and
collect lots and lots
of links in only two
to three hops!



Internet z punktu widzenia crawlera

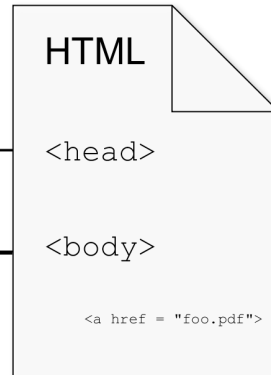
- crawler odnajduje stronę z rekordem

Internet z punktu widzenia crawlera

- crawler odnajduje stronę z rekordem
- crawler stara się odczytać metadane

1 Metadane zawarte w tagach HTML typu <meta>

```
<meta name="citation_author" content="Kowalska, T.">
```

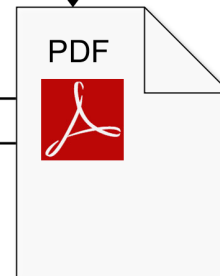


3 Metadane zawarte w nagłówku pliku PDF

```
67 6001 0 obj<</F 163 0 R/I<</Title(<author> A.)>>>>  
68 endobj  
69 6002 0 obj<</F 176 0 R/I<</Title(Kowalska , T .)>>>>  
70 endobj
```

2 Dane zawarte w rekordzie bibliograficznym
zakodowane w sekcji <body> pliku HTML

```
<tr>  
  <td> autor </td>  
  <td>  
    <a href="contr/f49eda7b ">  
      Kowalska, T.  
    </a>  
  </td>      autor  <u>Kowalska, T.</u>  
</tr>
```



4 Informacje zawarte w tekście artykułu

tekst autorstwa dr T. Kowalskiej

```
<title>Asynchroniczna obsługa urządzeń wejścia-wyjścia</title>
<link rel="schema.DCTERMS" href="http://purl.org/dc/terms/">
<link rel="schema.DC" href="http://purl.org/dc/elements/1.1/">
<meta name="DC.creator" content="Nowak, Robert">
<meta name="DCTERMS.dateAccepted" content="2014-01-09T10:12:13Z" scheme="DCTERMS.W3CDTF">
<meta name="DCTERMS.available" content="2014-01-09T10:12:13Z" scheme="DCTERMS.W3CDTF">
<meta name="DCTERMS.issued" content="2010" scheme="DCTERMS.W3CDTF">
<meta name="DC.identifier" content="https://depot.ceon.pl/handle/123456789/3160" scheme="DCTERMS.URI">
<meta name="DCTERMS.abstract" content="Urządzenia wejścia - wyjścia działają znacznie wolniej niż procesor, dlatego
czasie oczekiwania na odpowiedź urządzenia warto go zwalniać. Współczesne systemy operacyjne dostarczają udogodnień
które pozwalają to osiągnąć bez angażowania niezależnych wątków." xml:lang="pl">
<meta name="DC.language" content="pl" xml:lang="pl" scheme="DCTERMS.RFC1766">
<meta name="DC.publisher" content="Software Developers Journal" xml:lang="pl">
<meta name="DC.rights" content="Creative Commons Uznanie autorstwa 3.0 Polska" xml:lang="pl_PL">
<meta name="DC.rights" content="http://creativecommons.org/licenses/by/3.0/pl/legalcode" scheme="DCTERMS.URI">
<meta name="DC.subject" content="wzorce projektowe" xml:lang="pl">
<meta name="DC.subject" content="C++" xml:lang="pl">
<meta name="DC.subject" content="programowanie" xml:lang="pl">
<meta name="DC.title" content="Asynchroniczna obsługa urządzeń wejścia-wyjścia" xml:lang="pl">
<meta name="DC.type" content="info:eu-repo/semantics/article" xml:lang="pl">
<meta name="DC.contributor" content="Politechnika Warszawska, Instytut Systemów Elektronicznych" xml:lang="pl">
<meta name="DC.description" content="Robert Nowak">
<meta name="DC.rights" content="info:eu-repo/semantics/openAccess">
<meta content="wzorce projektowe; C++; programowanie; info:eu-repo/semantics/article" name="citation_keywords">
<meta content="Asynchroniczna obsługa urządzeń wejścia-wyjścia" name="citation_title">
<meta content="Software Developers Journal" name="citation_publisher">
<meta content="pl" name="citation_language">
<meta content="Nowak, Robert" name="citation_author">
<meta content="https://depot.ceon.pl/bitstream/123456789/3160/1/sdj2010asio.pdf" name="citation_pdf_url">
<meta content="2010" name="citation_date">
<meta content="https://depot.ceon.pl/handle/123456789/3160" name="citation_abstract_html_url">
```

Internet z punktu widzenia crawlera

- crawler odnajduje stronę z rekordem
- crawler stara się odczytać metadane

GS ma swoje preferencje odnośnie do metadanych

Internet z punktu widzenia crawlera

- crawler odnajduje stronę z rekordem
- crawler stara się odczytać metadane

GS ma swoje preferencje odnośnie do metadanych

Preferuje:

Highwire Press (citation_title)

Eprints (eprints.title)

BE Press (bepress_citation_title)

<https://scholar.google.com/intl/en/scholar/inclusion.html#indexing>

Internet z punktu widzenia crawlera

- crawler odnajduje stronę z rekordem
- crawler stara się odczytać metadane

GS ma swoje preferencje odnośnie do metadanych

Odradza:

Dublin Core (dc.creator)

<https://scholar.google.com/intl/en/scholar/inclusion.html#indexing>

2. Zalety i wady Google Scholar

2. Zalety i wady Google Scholar

2.1. Zalety

2.1. Zalety Google Scholar

a) Ogromny corpus tekstów

2.1. Zalety Google Scholar

a) Ogromny corpus tekstów

E. Orduna-Malea i in., *About the Size of Google Scholar (...)* (2014)

<https://arxiv.org/abs/1407.6239>

2.1. Zalety Google Scholar

a) Ogromny corpus tekstów

E. Orduna-Malea i in., *About the Size of Google Scholar (...)* (2014)

<https://arxiv.org/abs/1407.6239>

160 milionów tekstów naukowych

2.1. Zalety Google Scholar

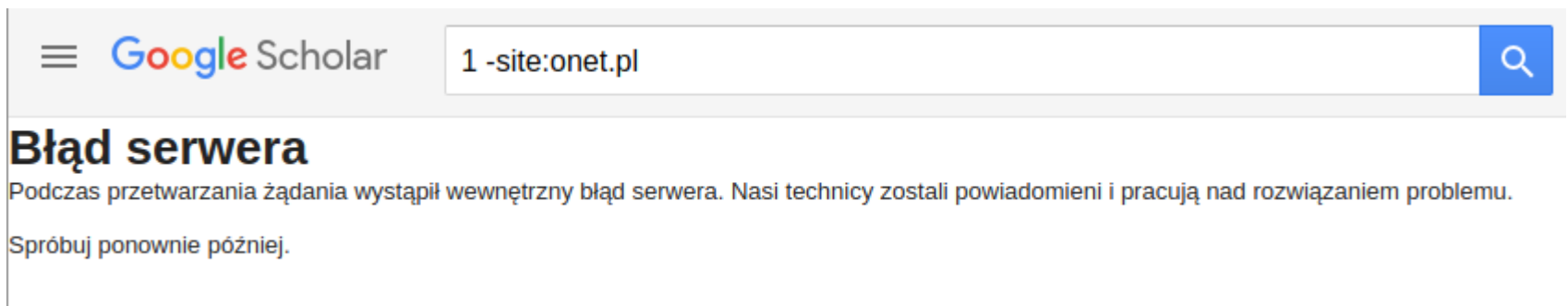
a) Ogromny corpus tekstów

E. Orduna-Malea i in., *About the Size of Google Scholar (...)* (2014)

<https://arxiv.org/abs/1407.6239>

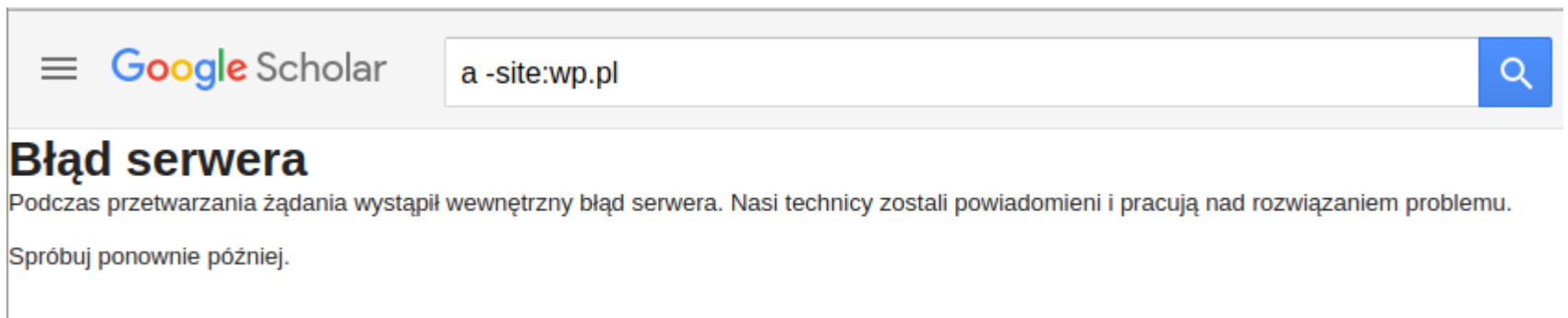
160 milionów tekstów naukowych (+/- 10%)

2.1. Zalety Google Scholar



The screenshot shows the Google Scholar interface. On the left, there is a menu icon and the text "Google Scholar". In the center, a search bar contains the query "1 -site:onet.pl". To the right of the search bar is a blue search button with a magnifying glass icon. Below the search bar, the text "Błąd serwera" is displayed in a large, bold font. Underneath this, a message reads: "Podczas przetwarzania żądania wystąpił wewnętrzny błąd serwera. Nasi technicy zostali powiadomieni i pracują nad rozwiązaniem problemu." At the bottom of the error message, it says "Spróbuj ponownie później."

2.1. Zalety Google Scholar



The screenshot shows the Google Scholar search interface. At the top left, there is a hamburger menu icon followed by the text "Google Scholar". To the right of this is a search input field containing the text "a -site:wp.pl". A blue search button with a magnifying glass icon is located to the right of the search field. Below the search bar, the text "Błąd serwera" is displayed in a large, bold font. Underneath this, a smaller line of text reads: "Podczas przetwarzania żądania wystąpił wewnętrzny błąd serwera. Nasi technicy zostali powiadomieni i pracują nad rozwiązaniem problemu." At the bottom of the error message, it says "Spróbuj ponownie później."

2.1. Zalety Google Scholar

a) Ogromny corpus tekstów

E. Orduna-Malea i in., *About the Size of Google Scholar (...)* (2014)

<https://arxiv.org/abs/1407.6239>

160 milionów tekstów naukowych (+/- 10%) w 2014

2.1. Zalety Google Scholar

a) Ogromny corpus tekstów

E. Orduna-Malea i in., *About the Size of Google Scholar (...)* (2014)

<https://arxiv.org/abs/1407.6239>

160 milionów tekstów naukowych (+/- 10%) w 2014
+ 3 mln w 2015, + 6 mln w 2016 i + 3 mln w 2017

2.1. Zalety Google Scholar

a) Ogromny corpus tekstów

E. Orduna-Malea i in., *About the Size of Google Scholar (...)* (2014)

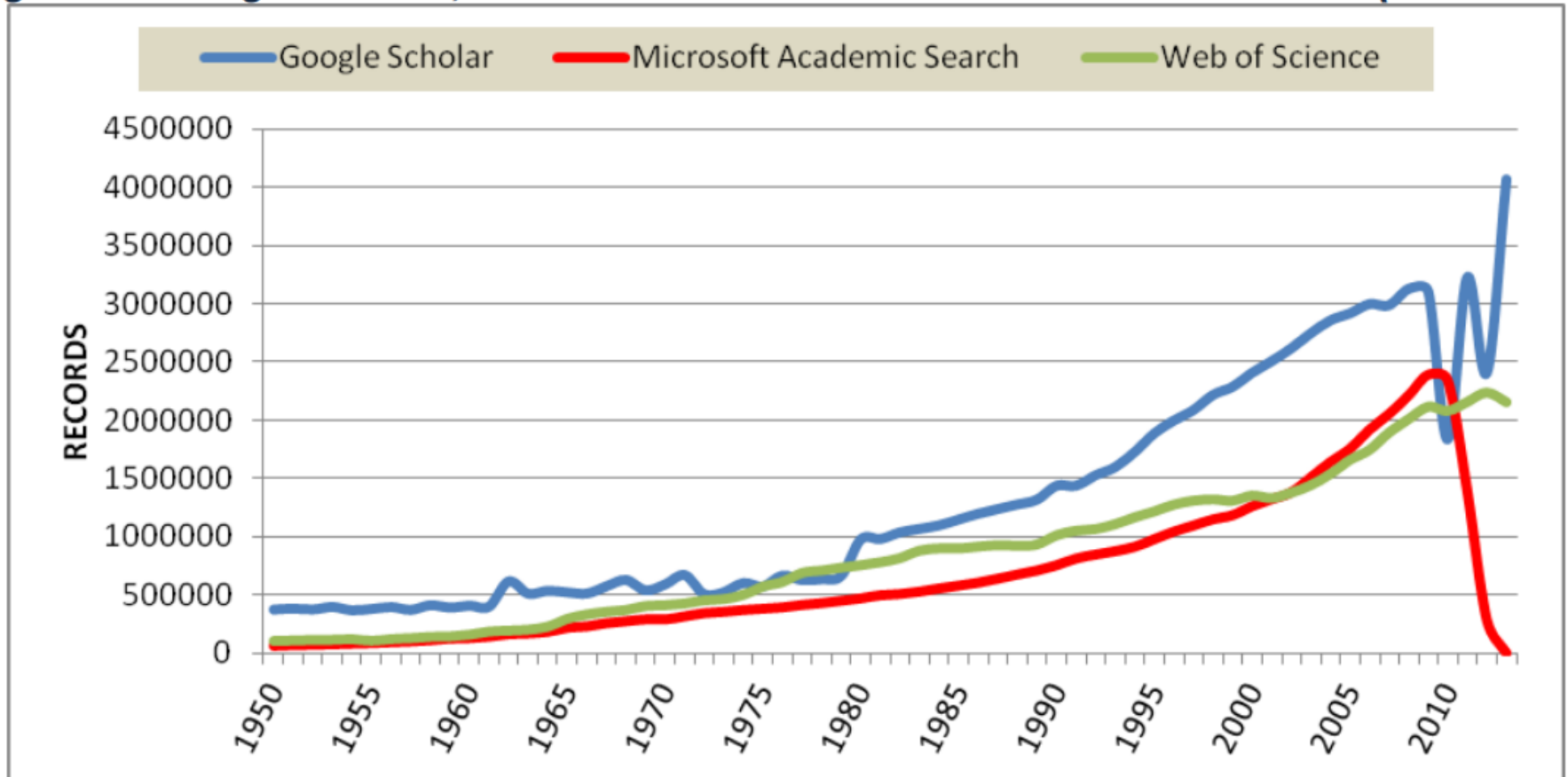
<https://arxiv.org/abs/1407.6239>

160 milionów tekstów naukowych (+/- 10%) w 2014
+ 3 mln w 2015, + 6 mln w 2016 i + 3 mln w 2017
ok. 170 - 180 milionów tekstów naukowych

2.1. Zalety Google Scholar

- a) ok. 175 mln tekstów naukowych
- b) nadreprezentacja najnowszych tekstów

Figure 10. Google Scholar, Microsoft Academic Search and Web of Science (1950–2013)



Źródło: E. Orduna-Malea i in., *About the Size of Google Scholar (...)*

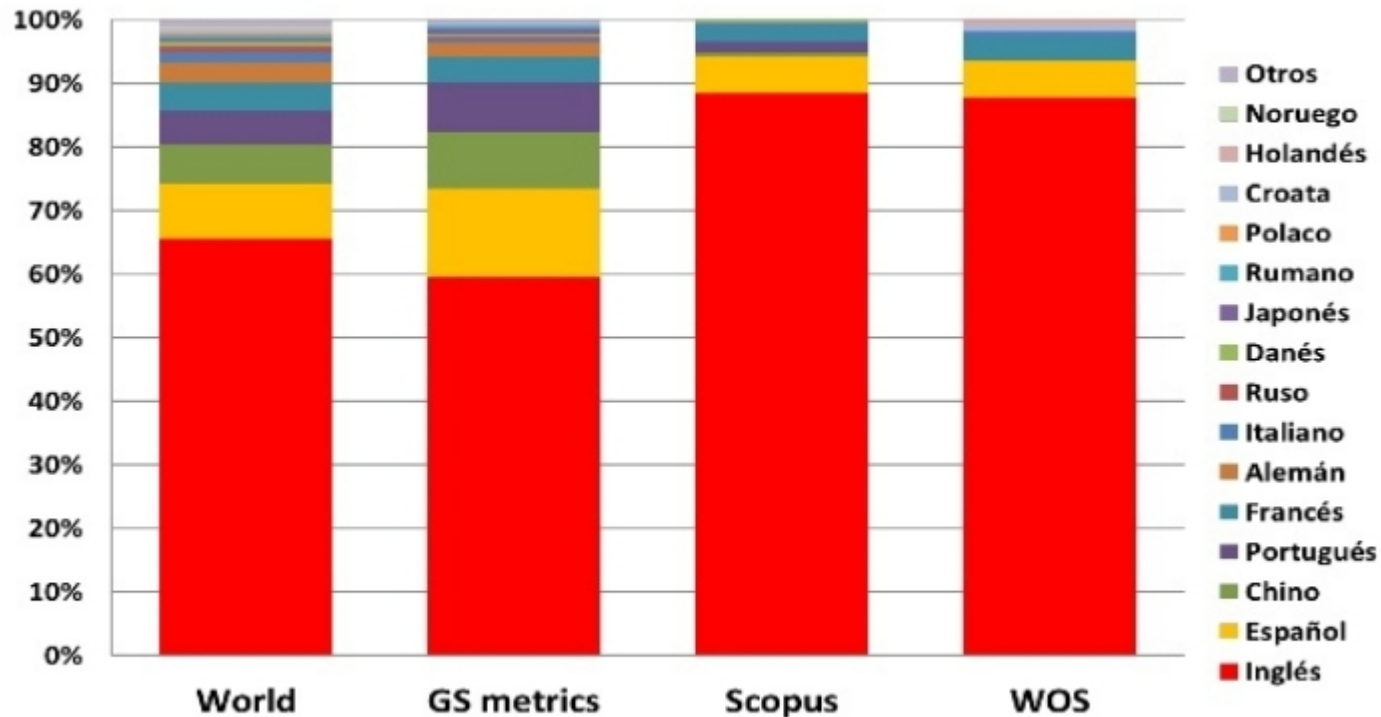
2.1. Zalety Google Scholar

- a) ok. 175 mln tekstów naukowych
- b) nadreprezentacja najnowszych tekstów
- c) brak nadreprezentacji tekstów w języku angielskim

2.1. Zalety Google Scholar

- a) ok. 175 mln tekstów naukowych
- b) nadreprezentacja najnowszych tekstów
- c) brak nadreprezentacji tekstów w języku angielskim
(inaczej niż WoS i Scopus)

Communications Journals (2012)



Delgado López-Cózar, E.; Repiso Caballero, R. El impacto de las revistas de Comunicación: comparando Google Scholar Metrics, Web of Science y Scopus. Comunicar, en prensa

2.1. Zalety Google Scholar

- a) ok. 175 mln tekstów naukowych
- b) nadreprezentacja najnowszych tekstów
- c) brak nadreprezentacji tekstów w języku angielskim
(inaczej niż WoS i Scopus)
- d) działa tak, jak w teorii agregator działać powinien

2.1. Zalety Google Scholar

- a) ok. 175 mln tekstów naukowych
- b) nadreprezentacja najnowszych tekstów
- c) brak nadreprezentacji tekstów w języku angielskim
(inaczej niż WoS i Scopus)
- d) działa tak, jak w teorii agregator działać powinien

2.1. Zalety Google Scholar

d) działa tak, jak w teorii agregator działać powinien

2.1. Zalety Google Scholar

- d) działa tak, jak w teorii agregator działań powinien
 - integruje wszystkie dostępne rekordy danego tekstu

2.1. Zalety Google Scholar

- d) działa tak, jak w teorii agregator działań powinien
 - integruje wszystkie dostępne rekordy danego tekstu
(o ile może rozpoznać tożsamość)

2.1. Zalety Google Scholar

- d) działa tak, jak w teorii agregator działań powinien
- integruje wszystkie dostępne rekordy danego tekstu
(o ile może rozpoznać tożsamość)
 - przeprowadza deduplikację autorów

2.1. Zalety Google Scholar

- d) działa tak, jak w teorii agregator działań powinien
- integruje wszystkie dostępne rekordy danego tekstu
(o ile może rozpoznać tożsamość)
 - przeprowadza deduplikację autorów
 - automatyczna detekcja cytowań

2.1. Zalety Google Scholar

- d) działa tak, jak w teorii agregator działań powinien
- integruje wszystkie dostępne rekordy danego tekstu
(o ile może rozpoznać tożsamość)
 - przeprowadza deduplikację autorów
 - automatyczna detekcja cytowań
 - wylicza h-index, prezentuje statystyki

2.1. Zalety Google Scholar

- a) ok. 175 mln tekstów naukowych
- b) nadreprezentacja najnowszych tekstów
- c) brak nadreprezentacji tekstów w języku angielskim
(inaczej niż WoS i Scopus)
- d) działa tak, jak w teorii agregator działać powinien
- e) powszechnie używany

2.1. Zalety Google Scholar

e) powszechnie używany

A. Martin-Martin i in., *Back to the past: on the shoulders of an academic search engine giant.*

Scientometrics 107.3 (2016): 1477-1487

GS może być odpowiedzialny za trend częstszego cytowania starszych artykułów.

2.1. Zalety Google Scholar

- a) ok. 175 mln tekstów naukowych
- b) nadreprezentacja najnowszych tekstów
- c) brak nadreprezentacji tekstów w języku angielskim
(inaczej niż WoS i Scopus)
- d) działa tak, jak w teorii agregator działać powinien
- e) powszechnie używany
- f) za darmo

2. Zalety i wady Google Scholar

2.2. Wady

2.2. Wady Google Scholar

a) kłopoty z inkluzją artykułów

2.2. Wady Google Scholar

a) kłopoty z inkluzją artykułów

- niekiedy zbyt trudno dostać się „na indeks”

2.2. Wady Google Scholar

a) kłopoty z inkluzją artykułów

- niekiedy zbyt trudno dostać się „na indeks”
- niekiedy zbyt łatwo dostać się „na indeks”

2.2. Wady Google Scholar

a) kłopoty z inkluzją artykułów

- niekiedy zbyt trudno dostać się „na indeks”

- niekiedy zbyt łatwo dostać się „na indeks”

b) całkowity brak ludzkiego nadzoru

2.2. Wady Google Scholar

[ryc 9 – author:Vietnam]

2.2. Wady Google Scholar

- a) kłopoty z inkluzją artykułów
 - niekiedy zbyt trudno dostać się „na indeks”
 - niekiedy zbyt łatwo dostać się „na indeks”
- b) całkowity brak ludzkiego nadzoru
- c) brak wyszukiwania semantycznego -
tylko słowa kluczowe

2.2. Wady Google Scholar

- a) kłopoty z inkluzją artykułów
 - niekiedy zbyt trudno dostać się „na indeks”
 - niekiedy zbyt łatwo dostać się „na indeks”
- b) całkowity brak ludzkiego nadzoru
- c) brak wyszukiwania semantycznego -
tylko słowa kluczowe
- d) za darmo...

Repozytoria akademickie a GS

Repozytoria akademickie a GS

Repozytorium akademickie: usługa internetowa pozwalająca na długoterminowe przechowywanie publikacji naukowych i/lub danych badawczych oraz dostęp do zdeponowanych zasobów za pomocą WWW.

Repozytoria akademickie a GS

Z punktu widzenia crawlera wyszukiwarki akademickiej repozytorium wyróżnia się względnie dużą ilością zasobów zgromadzonych w jednym miejscu.

Repozytoria akademickie a GS

Z punktu widzenia crawlera wyszukiwarki akademickiej repozytorium wyróżnia się względnie dużą ilością zasobów zgromadzonych w jednym miejscu.

Ponieważ crawler dysponuje ograniczoną ilością zasobów (mocy obliczeniowej i czasu), strony o słabej wydajności lub złej architekturze mogą efektywnie nie udostępniać części zdeponowanych zasobów.

Repozytoria akademickie a GS

Z punktu widzenia crawlera wyszukiwarki akademickiej repozytorium wyróżnia się względnie dużą ilością zasobów zgromadzonych w jednym miejscu.

Jeśli jednak strona jest dobrze zaprojektowana, Google może rozpoznać repozytorium jako cenny zasób.

Repozytoria akademickie a GS

Z punktu widzenia crawlera wyszukiwarki akademickiej repozytorium wyróżnia się względnie dużą ilością zasobów zgromadzonych w jednym miejscu.

Jeśli jednak strona jest dobrze zaprojektowana, Google może rozpoznać repozytorium jako cenny zasób. GS prowadzi na bieżąco „wewnętrzny” ranking stron i preferuje te miejsca w Sieci, z których można szybko wydobyć wysokiej jakości dane.

Repozytoria akademickie a GS

Istnieje szereg rekomendacji Google na temat tego, jak uczynić repozytorium akademickie bardziej widocznym poprzez dopasowanie go do wymagań crawlerów.

Repozytoria akademickie a GS

Istnieje szereg rekomendacji Google na temat tego, jak uczynić repozytorium akademickie bardziej widocznym poprzez dopasowanie go do wymagań crawlerów.

D. Dapra, A. Acharya *Indexing Repositories: Pitfalls and Best Practices*

Repozytoria akademickie a GS

Istnieje szereg rekomendacji Google na temat tego, jak uczynić repozytorium akademickie bardziej widocznym poprzez dopasowanie go do wymagań crawlerów.

D. Dapra, A. Acharya *Indexing Repositories: Pitfalls and Best Practices*

Google Scholar Inclusion Guide <https://google.scholar.com/intl/en/scholar/inclusion.html>

Repozytoria akademickie a GS

Istnieje szereg rekomendacji Google na temat tego, jak uczynić repozytorium akademickie bardziej widocznym poprzez dopasowanie go do wymagań crawlerów.

D. Dapra, A. Acharya *Indexing Repositories: Pitfalls and Best Practices*

Google Scholar Inclusion Guide <https://google.scholar.com/intl/en/scholar/inclusion.html>

Jak poprawić widoczność rekordów w Google Scholar (otwartanauka.pl)

Repozytoria akademickie a GS

Istnieje szereg rekomendacji Google na temat tego, jak uczynić repozytorium akademickie bardziej widocznym poprzez dopasowanie go do wymagań crawlerów.

D. Dapra, A. Acharya *Indexing Repositories: Pitfalls and Best Practices*

Google Scholar Inclusion Guide <https://google.scholar.com/intl/en/scholar/inclusion.html>

Jak poprawić widoczność rekordów w Google Scholar (otwartanauka.pl)

Co można zrobić osobiście?

Co można zrobić osobiście?

- a) Założyć konto w Google Scholar

Co można zrobić osobiście?

- a) Założyć konto w Google Scholar
 - możliwość wskazywania tekstów crawlerom

Co można zrobić osobiście?

- a) Założyć konto w Google Scholar
 - możliwość wskazywania tekstów crawlerom
 - widać rezultat przetworzenia danych przez GS

Co można zrobić osobiście?

- a) Założyć konto w Google Scholar
 - możliwość wskazywania tekstów crawlerom
 - widać rezultat przetworzenia danych przez GS
- b) Odpowiednio przygotowywać deponowane pdfy

Co można zrobić osobiście?

- a) Założyć konto w Google Scholar
 - możliwość wskazywania tekstów crawlerom
 - widać rezultat przetworzenia danych przez GS
- b) Odpowiednio przygotowywać deponowane pdfy
 - metadane (autor/autorzy, afiliacje itp.)

Co można zrobić osobiście?

- a) Założyć konto w Google Scholar
 - możliwość wskazywania tekstów crawlerom
 - widać rezultat przetworzenia danych przez GS
- b) Odpowiednio przygotowywać deponowane pdfy
 - metadane (autor/autorzy, afiliacje itp.)

Wiele komercyjnych programów (Adobe, InDesign)

Co można zrobić osobiście?

- a) Założyć konto w Google Scholar
 - możliwość wskazywania tekstów crawlerom
 - widać rezultat przetworzenia danych przez GS
- b) Odpowiednio przygotowywać deponowane pdfy
 - metadane (autor/autorzy, afiliacje itp.)

Wiele komercyjnych programów (Adobe, InDesign)

Również wiele programów darmowych.

Co można zrobić osobiście?

- a) Założyć konto w Google Scholar
 - możliwość wskazywania tekstów crawlerom
 - widać rezultat przetworzenia danych przez GS
- b) Odpowiednio przygotowywać deponowane pdfy
 - metadane (autor/autorzy, afiliacje itp.)

Wiele komercyjnych programów (Adobe, InDesign)

Również wiele programów darmowych.

(Autometadata, PDF Metadata Editor, Pdftk [ubuntu])

Co można zrobić osobiście?

- a) Założyć konto w Google Scholar
 - możliwość wskazywania tekstów crawlerom
 - widać rezultat przetworzenia danych przez GS
- b) Odpowiednio przygotowywać deponowane pdfy
 - metadane
- c) Rozpowszechniać publikacje z pomocą social media

Co można zrobić osobiście?

- a) Założyć konto w Google Scholar
 - możliwość wskazywania tekstów crawlerom
 - widać rezultat przetworzenia danych przez GS
- b) Odpowiednio przygotowywać deponowane pdfy
 - metadane
- c) Rozpowszechniać publikacje z pomocą social media
 - większość social media jest dość wysoko w rankingach Google. Zawiadamiamy więc nie tylko kolegów i koleżanki, ale także i roboty.

Dziękuję za uwagę.

t.lewandowski@icm.edu.pl