



Politechnika Wroclawska

Bibliometria i więcej: kilka uwag nie na temat

Wojciech Myszka

Katedra Mechaniki i Inżynierii Materiałowej
Politechnika Wroclawska

Polanica-Zdrój, 20 października 2016

XXIII Krajowa Konferencja Komputerowe
Wspomaganie Badań Naukowych KOWBAN'2016



Biblioteki





Biblioteki

- ▶ Najpierw były biblioteki.
- ▶ Do bibliotek zaczęli zjeżdżać się ludzie żeby studiować księgi.
- ▶ Po pewnym czasie wokół bibliotek pojawiły się grupy ludzi objaśniające trudniejsze fragmenty ksiąg. I tak powstały uczelnie.



Bibliometria

- ▶ Gdy książki się zdemokratyzowały (dzięki Gutenbergowi) zaczęły pojawiać się biblioteki prywatne.
- ▶ Po jakimś czasie naukowcy zaczęli prowadzić badania „bibliograficzne” szukając wpływów literatury znalezionej w bibliotece na twórczość jej właściciela.
- ▶ Na początku lat 60 zeszłego stulecia pojawiła się współczesna bibliometria.
- ▶ Prawdziwym hitem w czasach „przedinternetowych” był Science Citation Index pozwalający badać naukową „przeszłość” i „przyszłość” artykułu.



Bibliometria





Indeks h

- ▶ Królem w tamtych czasach była „liczba cytowań”.



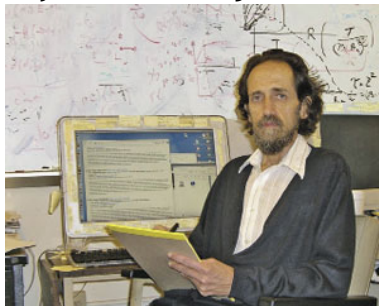
Indeks h

- ▶ Królem w tamtych czasach była „liczba cytowań”.
- ▶ Ale dopiero pojawienie się, z jednej strony komputerów, a z drugiej pomysłu Jorge Hirscha doprowadziło system na skraj absurdu.



Indeks h

- ▶ Królem w tamtych czasach była „liczba cytowań”.
- ▶ Ale dopiero pojawienie się, z jednej strony komputerów, a z drugiej pomysłu Jorge Hirscha doprowadziło system na skraj absurdu.



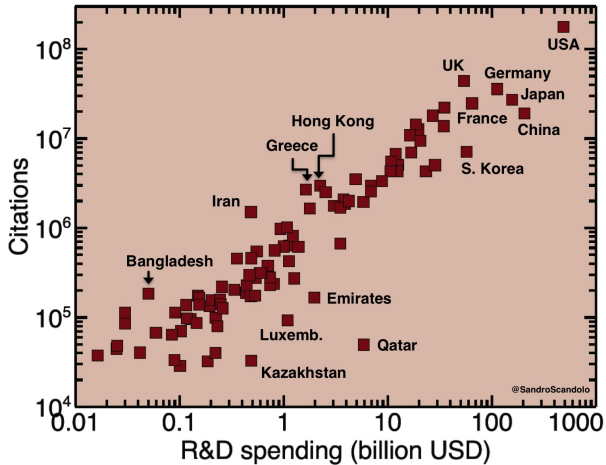


Indeks h

- ▶ Królem w tamtych czasach była „liczba cytowań”.
- ▶ Ale dopiero pojawienie się, z jednej strony komputerów, a z drugiej pomysłu Jorge Hirscha doprowadziło system na skraj absurdu.
- ▶ Z drugiej strony nie można tak w czambuł potępiać bibliometrii – jest to bardzo użyteczne narzędzie do różnego rodzaju syntetycznych ocen.



Indeks h





Indeks h

- ▶ Królem w tamtych czasach była „liczba cytowań”.
- ▶ Ale dopiero pojawienie się, z jednej strony komputerów, a z drugiej pomysłu Jorge Hirscha doprowadziło system na skraj absurdu.
- ▶ Z drugiej strony nie można tak w czambuł potępiać bibliometrii – jest to bardzo użyteczne narzędzie do różnego rodzaju syntetycznych ocen.
- ▶ Zapominamy natomiast, że narzędzia statystyczne w swym charakterze działają dobrze w warunkach wielu prób.



Indeks h

- ▶ Królem w tamtych czasach była „liczba cytowań”.
- ▶ Ale dopiero pojawienie się, z jednej strony komputerów, a z drugiej pomysłu Jorge Hirscha doprowadziło system na skraj absurdu.
- ▶ Z drugiej strony nie można tak w czambuł potępiać bibliometrii – jest to bardzo użyteczne narzędzie do różnego rodzaju syntetycznych ocen.
- ▶ Zapominamy natomiast, że narzędzia statystyczne w swym charakterze działają dobrze w warunkach wielu prób.
- ▶ Co nie przeszkadza w budowaniu śmiesznych modeli przewidujących naukową karierę.



Indeks h

METRICS

Predict your future h-index

These are approximate equations for predicting the h -index of neuroscientists in the future. They are probably reasonably

precise for life scientists, but likely to be less meaningful for the other sciences. Try it for yourself online at go.nature.com/z4rroc.

- **Predicting next year** ($R^2=0.92$):

$$h_{+1} = 0.76 + 0.37\sqrt{n} + 0.97h - 0.07y + 0.02j + 0.03q$$

- **Predicting 5 years into the future** ($R^2=0.67$):

$$h_{+5} = 4 + 1.58\sqrt{n} + 0.86h - 0.35y + 0.06j + 0.2q$$

- **Predicting 10 years into the future** ($R^2=0.48$):

$$h_{+10} = 8.73 + 1.33\sqrt{n} + 0.48h - 0.41y + 0.52j + 0.82q$$

Key: n , number of articles written; h , current h -index; y , years since publishing first article; j , number of distinct journals published in; q , number of articles in *Nature*, *Science*, *Nature Neuroscience*, *Proceedings of the National Academy of Sciences* and *Neuron*.



Indeks h

- ▶ Królem w tamtych czasach była „liczba cytowań”.
- ▶ Ale dopiero pojawienie się, z jednej strony komputerów, a z drugiej pomysłu Jorge Hirscha doprowadziło system na skraj absurdu.
- ▶ Z drugiej strony nie można tak w czambuł potępiać bibliometrii – jest to bardzo użyteczne narzędzie do różnego rodzaju syntetycznych ocen.
- ▶ Zapominamy natomiast, że narzędzia statystyczne w swym charakterze działają dobrze w warunkach wielu prób.
- ▶ Co nie przeszkadza w budowaniu śmiesznych modeli przewidujących naukową karierę.
- ▶ Ale nie o tym chciałem. . .

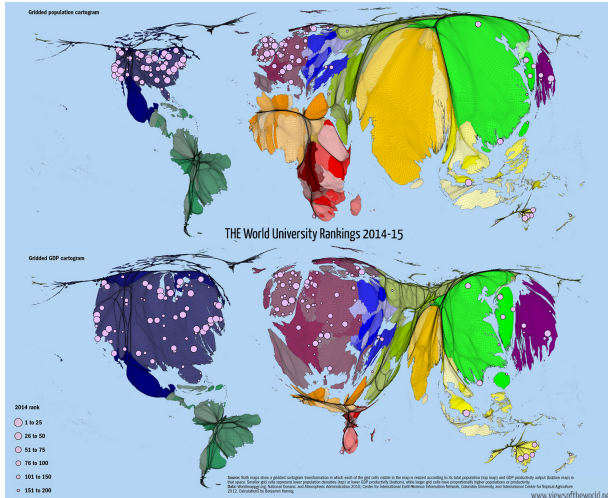


Rankingi





Rankingi





Rankingi

Każdy ranking zawiera ocene bibliometryczną uczelni!



Metadane

- ▶ Metadane to dane o danych.
- ▶ W przypadku pracy naukowej mogą zawierać informacje o tym:
 - ▶ autorze,
 - ▶ typie publikacji (artykuł. książka, referat konferencyjny, . . .),
 - ▶ miejscu publikacji,
 - ▶ źródle,
 - ▶ słowach kluczowych,
 - ▶ streszczeniu,
 - ▶ cytowanych pracach.



Metadane

- ▶ Metadane to dane o danych.
- ▶ W przypadku rozmowy telefonicznej:
 - ▶ dzwoniącym
 - ▶ odbiorcy,
 - ▶ czasie zainicjowania połączenia,
 - ▶ czasie trwania rozmowy,
 - ▶ położeniu geograficznym nadawcy i odbiorcy



Metadane

- ▶ Metadane to dane o danych.
- ▶ W przypadku zdjęcia (pliku jpg):
 - ▶ marce aparatu,
 - ▶ warunkach naświetlania zdjęcia (czas, przysłona),
 - ▶ zastosowanej optyce,
 - ▶ nazwisku Autora,
 - ▶ miejscu zrobienia zdjęcia,
 - ▶ ...



Metadane

- ▶ Metadane to dane o danych.
- ▶ W przypadku pliku z muzyką (MP3):
 - ▶ tytuł,
 - ▶ autor,
 - ▶ numer ścieżki,
 - ▶ nazwa albumu,
 - ▶ rok wydania,
 - ▶ okładka albumu,
 - ▶ ...



Metadane

- ▶ Metadane to dane o danych.
- ▶ W przypadku transakcji w markecie:
 - ▶ data transakcji,
 - ▶ spis zakupionych towarów,
 - ▶ ostatnie cztery cyfry karty płatniczej,
 - ▶ kod pocztowy.

(ostatnie cztery cyfry karty płatniczej plus kod pocztowy, w praktyce, jednoznacznie identyfikują nabywcę).



Metadane: dobre czy złe?

Edward Snowden w pierwszej połowie 2013 roku zaczął ujawniać dokumenty pokazujące skalę inwigilacji dokonywanej przez służby Stanów Zjednoczonych.





Metadane: dobre czy złe?

Na stronach pomocy technicznej firmy Microsoft bardzo łatwo znaleźć informacje o usuwaniu metadanych z dokumentów tworzonych przez pakiet biurowy Office.

 Microsoft

 Sign in

 Office

Support

Remove hidden data and personal information by inspecting documents

Applies To: Word 2016, Word 2013, Word 2010, Office 2010, Word Starter, [More...](#)

If you plan to share an electronic copy of a Microsoft Word document with clients or colleagues, it is a good idea to review the document for hidden data or personal information that might be stored in the document itself or in the document properties (metadata). Because this hidden information can reveal details about your organization or about the document itself that you might not want to share publicly, you might want to remove this hidden information before you share the document with other people.

This article describes how the Document Inspector feature in Word can help you find and remove



Metadane: dobre czy złe?

W sieci bardzo wiele przykładów pokazujących do czego może doprowadzić znajomość metadanych.

The Incredible Story Of How Target Exposed A Teen Girl's Pregnancy



Gus Lubin

Feb. 16, 2012, 10:27 AM

59,702

17



FACEBOOK



LINKEDIN



TWITTER



Target broke through to a new level of customer tracking with the help of statistical genius Andrew Pole, according to a [New York Times Magazine cover story](#) by Charles Duhigg.

Pole identified 25 products that when purchased together indicate a woman is likely pregnant. The value of this information was that Target





Metadane: dobre czy złe?

To jest strona tytułowa artykułu dostępnego gdzieś w sieci.
A wszystkie informacje zawarte na niej to metadane.

Ann. Inst. Statist. Math.
Vol. 40, No. 4, 821-827 (1988)

OPTIMUM EXPERIMENTAL DESIGN FOR A REGRESSION ON A HYPERCUBE-GENERALIZATION OF HOEL'S RESULT

EWARYST RAFAJŁOWICZ¹ AND WOJCIECH MYSZKA²

¹*Institute of Engineering Cybernetics, Technical University of Wrocław,
Wybrzeże St. Wyspiańskiego 27, 50-370 Wrocław, Poland*

²*Institute of Material Science and Mechanical Engineering, Technical University of Wrocław,
Wybrzeże St. Wyspiańskiego 27, 50-370 Wrocław, Poland*

(Received September 24, 1986; revised September 16, 1987)

Abstract. In the note Hoel's result (1965, *Ann. Math. Statist.*, **36**, 1097–1106) is generalized to a large family of experimental design optimality criterions. Sufficient conditions for optimality criterion are given, which ensure existence of the optimum experimental design measure which is a product of design measures on lower dimensional domains.

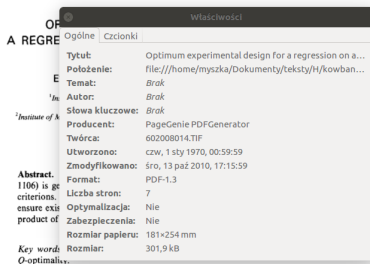
Key words and phrases: Experimental design, regression, D -, L -, A -, Q -optimality.



Metadane: dobre czy złe?

Każdy dobry program do przeglądania plików PDF pozwala przejrzeć specjalne pola w pliku PDF przeznaczone na przechowywanie metadanych pliku.

Ann. Inst. Statist. Math.
Vol. 40, No. 4, 821-827 (1988)



Czemu z nich nie korzystać?



Metadane: dobre czy złe?

- ▶ Nie wiem, czy metadane są dobre, czy złe.



Metadane: dobre czy złe?

- ▶ Nie wiem, czy metadane są dobre, czy złe.
- ▶ Wiem, że jesteśmy wszyscy od nich uzależnieni:



Metadane: dobre czy złe?

- ▶ Nie wiem, czy metadane są dobre, czy złe.
- ▶ Wiem, że jesteśmy wszyscy od nich uzależnieni:
 - ▶ z jednej strony automaty indeksujące na ich podstawie obliczają różnego rodzaju wskaźniki,



Metadane: dobre czy złe?

- ▶ Nie wiem, czy metadane są dobre, czy złe.
- ▶ Wiem, że jesteśmy wszyscy od nich uzależnieni:
 - ▶ z jednej strony automaty indeksujące na ich podstawie obliczają różnego rodzaju wskaźniki,
 - ▶ z drugiej – osoba która chce nas zacytować musi te informacje zamieścić w swojej pracy.



Metadane: dobre czy złe?

- ▶ Nie wiem, czy metadane są dobre, czy złe.
- ▶ Wiem, że jesteśmy wszyscy od nich uzależnieni:
 - ▶ z jednej strony automaty indeksujące na ich podstawie obliczają różnego rodzaju wskaźniki,
 - ▶ z drugiej – osoba która chce nas zacytować musi te informacje zamieścić w swojej pracy.
- ▶ Wynika z tego, że powinniśmy o metadane dbać (żeby były precyzyjne i towarzyszyły naszym pracom) i eksponować je jak najszerszej.



Ekspozycja metadanych

1. Strony WWW

- ▶ Standardy ekspozycji metadanych:
 - ▶ *Embedded metadata* (w szczególności Dublin Core).
 - ▶ OpenURL COinS (*ContextObjects in Spans*).
- ▶ Budując serwis informacyjny uczelni warto się z tym zapoznać i dobrze rozumieć.

2. Pliki PDF

3. Pliki procesorów tekstu

4. DOI (*digital object identifier* – cyfrowy identyfikator dokumentu elektronicznego)



Repozytoria I

1. Gdzie prezentować publikacje pracowników?
 - ▶ „amatorsko” na ich stronach osobistych?
czy
 - ▶ „profesjonalnie” w repozytoriach?
2. Repozytoria:
 - ▶ tworzyć własne?
czy
 - ▶ korzystać ze standardów (indeksowanych przez największe wyszukiwarki)?
 - ▶ Pod jakim adresem (uczelniany, środowiskowy, ogólnokrajowy)?



Repozytoria II

3. DOI i zarządzanie nimi:
 - ▶ korzystać z providerów?
 - ▶ zostać providerem?
 - ▶ krajowy?
 - ▶ środowiskowy?
4. Jak traktować „profile naukowców” na serwisach zewnętrznych (ResearchGate czy Google Scholar)?



Cytowanie

Po drugiej stronie jest cytowanie prac

- ▶ Nie mamy wielkiego wpływu jak są cytowane nasze prace.
- ▶ Ale możemy dbać o dobre obyczaje.
- ▶ Cytujmy poprawnie i domagajmy się tego od innych (to do recenzentów prac).
- ▶ Podstawowy zestaw metadanych (autor, tytuł, słowa kluczowe, abstrakt, data i miejsce wydania) pozwalają „zauważyć” publikację.
- ▶ Metadane rozszerzone o cytowane prace pozwalają ułatwić wyszukiwarkom liczenie indeksów cytowań.



I na koniec

citavi[®]



MENDELEY



RefWorks

citeulike

zotero





I na koniec

Korzystajmy z menedżerów cytowań!