

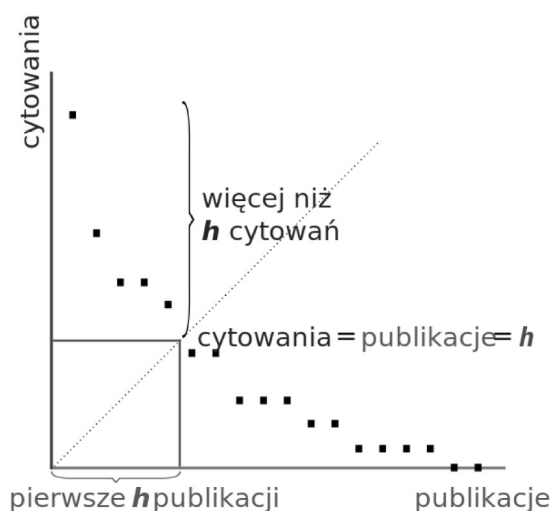
Bibliometria i więcej: kilka uwag nie na temat

STRESZCZENIE

Praca podsumowuje 25 lat doświadczeń zdobytych podczas tworzenia i używania usług internetowych na Politechnice Wroclawskiej oraz aktywnego uczestnictwa w tworzeniu i rozwoju stron WWW Instytutu, Wydziału, a wreszcie Katedry. Autor stara się sformułować niezbyt precyzyjne wnioski dotyczące ekspozycji metadanych w prezentacji publikacji naukowych na stronach WWW uczelni wyższej.

WSTĘP

W 2005 roku J.E. Hirsch opublikował (raczej dla żartu) pracę zatytułowaną *An index to quantify an individual's scientific research output* [1]. Praca ta definiowała dla kilkunastu laureatów nagrody Nobla wskaźnik mówiący o „popularności ich prac”. Popularność mierzona była ilością cytowań, a sam indeks autora h definiowany był jako liczba prac, które były cytowane h i więcej razy (graniczna praca cytowana jest h razy, a prace lepsze — więcej niż h razy, a wszystkie pozostałe — mniej).



Ryc. 1. Wskaźnik Hirscha z wykresu publikacji z malejącą liczbą cytowań.
Źródło Wikipedia [8]

Autor w pracy przedstawia dosyć szczegółową i krytyczną dyskusję porównując proponowany indeks z innymi miarami „produktywności” uczonych, takimi jak liczba publikacji, liczba cytowań, średnia liczba cytowań, liczba „znaczących prac” (gdzie za znaczące uważa się prace cytowane więcej niż zadana liczba razy) czy liczba cytowań zadanej liczby najczęściej cytowanych prac.

Ważnym wnioskiem jest stwierdzenie, że dwaj naukowcy o zbliżonym indeksie h , niezależnie od sumarycznej liczby prac czy sumarycznej liczby cytowań, mogą być traktowani jednakowo pod

względem ich „siły oddziaływania”.

Innym wnioskiem było to, że z dwu uczonych o zbliżonym „wieku akademickim” (mierzonym jako czas od pierwszej publikacji) i podobnej liczbie publikacji, „lepszy” jest ten, który ma wyższy indeks h .

Według oceny Hirscha uczonego o indeksie h na poziomie 20 (po dwudziestu latach pracy) odniósł sukces. Indeks (po dwudziestu latach pracy) na poziomie 40 oznacza wybitnego uczonego. Natomiast $h = 60$ (po dwudziestu latach pracy) definiuje prawdziwe indywidualności.

Hirsch oszacował też wartości indeksu h dla profesorów zatrudnionych na stałe (*associate professor*) — 12, uczonych z tytułem *full professor* od 18, członkowie American Physical Society to h w zakresie od 15 do 20, a dla członków Akademii Nauk USA — 45 lub więcej (poza znaczącymi wyjątkami).

W naturalny sposób indeks h powszechnie przyjęty został jako „miara uczonego” uczonego. Spowodowało to szereg dyskusji, krytyk i zaowocowało tworzeniem kolejnych mniej lub bardziej obiektywnych i dobrze zdefiniowanych wskaźników. Wszyscy natomiast zapomnieli, że Hirsch budował pewien model patrząc na uczonych, którzy odnieśli jakiś sukces, a my próbujemy pompować indeks h mając nadzieję, że stworzy to osobowości naukowe.

STAN OBECNY

Nie zamierzam skupiać się na różnego rodzaju kryteriach ani dyskutować ich zalet lub wad. Chcę zwrócić uwagę, że wszystkie są łatwe do automatycznego wyznaczenia na podstawie różnego rodzaju baz danych (publikacji, cytowań, częstości występowania w Internecie czy plotkarskich gazetach) przez programy komputerowe. Wszystkie bibliometryczne wskaźniki (a takim jest też indeks h) są właśnie takie.

Bazując na tej idei zaczęły powstawać rankingi już nie uczonych, ale również instytucji naukowych, czasopism, czy nawet państw.

Jeden z dosyć popularnych rankingów Ranking Web of Universities¹ (Webometrics) został w całości skonstruowany na rezultatach które można zdobyć w sieci. Ranking oparty jest na następujących parametrach:

Wielkość

— liczba stron WWW indeksowana, przez Google, Yahoo czy Bing.

Widoczność

— liczba odwołań zewnętrznych do stron jednostki (uzyskana z Yahoo Site Explorer).

Zawartość (*Rich Files*)

— liczba plików typu .pdf, .ps, .doc, .ppt udostępniana na stronach jednostki².

Liczba publikacji

— na podstawie Google Scholar i Scimago SIR.

Pomijając kwestię wysiłku (Webometrics ocenia w ten sposób ogromną liczbę uniwersytetów) każdy może dziś skonstruować swój własny ranking, a później publikować „niezależne” oceny jednostek naukowych, laboratoriów czy uniwersytetów.

Zaczynamy żyć w świecie, w którym wszystko to co opublikujemy trafia taką lub inną drogą do sieci i pozwala propagować nasze idee. Najlepsze nawet pomysły, jeżeli nie trafiają do sieci nie będą zauważone. Coś, czego nie znajduje żadna wyszukiwarka internetowa — nie istnieje (bo dotarcie do tego jest za trudne). Generuje to bardzo poważne konsekwencje.

METADANE

Coraz częściej nasze życie przenosi się ze świata „realnego” w świat „elektroniczny”, w którym nie tylko poszukujemy informacji, robimy zakupy ale też prowadzimy „życie towarzyskie”. Internecie używamy do kształcenia się czy poszukiwania dostępu do potrzebnych w pracy naukowej

1 <http://www.webometrics.info/en>

2 Zakłada się, że pliki takie zawierają istotną zawartość (*content*).

materiałów.

Dostawcy treści (w trosce o nasz komfort) starają się zdobyć jak najwięcej informacji o pojedynczych osobach będących ich klientami by jak najlepiej zaspakajać ich oczekiwania i zatrzymać ich przy sobie. Firmy komercyjne potrafią powiązać informacje o miejscu w którym znajduje się klient, używanym systemie operacyjnym, rozdzielczości ekranu, używanej przeglądarce by stworzyć indywidualny profil pozwalający jednoznacznie zidentyfikować klienta.

Nawet jeżeli nie korzystamy z Internetu, ale idziemy do sklepu, płacimy kartą, a kasjerka prosi nas o kod pocztowy — są to informacje wystarczające do zidentyfikowania osoby.

Aby łatwiej posługiwać się tymi informacjami wprowadzono pojęcie metadanych (czyli danych o danych). Budzi dziś ono szereg kontrowersji. Organizacje i ludzie zajmujący się cyberbezpieczeństwem uświadomili nam, że nawet nie znając treści komunikacji, a wiedząc z jakimi źródłami się komunikujemy — można wyciągnąć daleko idące wnioski.

Najczęściej cytowanym przykładem wykorzystania metadanych w marketingu jest przypadek amerykańskiej sieci handlowej Target. Zdając sobie sprawę z ważności i potencjału rynku młodych matek, opracowała ona specjalny program analizy danych (i zwyczajów zakupowych klientów) pozwalający nie tylko rozpoznać osoby w ciąży, ale również przewidzieć termin rozwiązania. Sprawa się wydała, gdy nastolatka, przy okazji zakupów w sklepie tej sieci, zaczęła dostawać zniżkowe kupony na pieluszki. Oburzyło to jej ojca, który nic o sprawie nie wiedział, a zaowocowało z jednej strony sławą statystyka, który był autorem modelu i oprogramowania ale też pewnymi kłopotami samej sieci podejrzewanej o szpiegowanie [2].

Takie same technologie używane są do identyfikowania klienta i wyświetlania mu reklam produktów których kiedyś szukał i przez różne portale internetowe do blokowania treści i oferowania płatnej prenumeraty. Sytuację ułatwia fakt, że przeciętny użytkownik sieci jest zawsze zalogowany: a to do Facebooka, a to do usług Google, a to do poczty Microsoft. Firmy mają ułatwione zadanie nie musząc stale go identyfikować.

Kolejne sensacje ujawnione przez Edwarda Snowdena (wraz z opisami programów i technik używanych przez NSA) uświadomiły specjalistom i wielu zwykłym użytkownikom Internetu, że służby specjalne zbierają, przy okazji walki z terroryzmem, bardzo wiele informacji, używanych później do podejmowania działań militarnych. Rozpoczęła się również dyskusja na ile używane modele statystyczne są pewne i jak często w wyniku podejmowanych działań giną niewinni ludzie.

Zdecydowanie metadane nie mają najlepszej prasy.

METADANE DLA NAUKOWCÓW

Jest też i jasna strona metadanych. Najważniejszą ich cechą — gdy używane poprawnie — jest możliwość powiązania autora z dziełem. Jest to bardzo ważne w czasach gdy bardzo wiele publikacji prezentowanych jest w Internecie. Co więcej dobre mechanizmy indeksujące bardzo chętnie korzystają z metadanych, gdyż znacznie łatwiej niż z analizy treści pliku PDF, DOC/DOCX wydobyć informacje na temat autora, tytułu pracy czy słów kluczowych z metadanych. Zwłaszcza pliki PDF potrafią ciągle stwarzać wiele problemów mechanizmom indeksującym. W szczególności „polskie litery” w przypadku użycia do produkcji dokumentu, fontów kiepskiej jakości stwarzają kłopoty. Co więcej, niektóre mechanizmy indeksujące nastawione są na kodowanie ASCII (w najlepszym razie Latin1). Jak na to nałożyć możliwe do popełnienia błędy podczas konwersji do PDF — ozazać się może że tekst dokumentu jest niemożliwy (lub trudny) do wydobywania z PDFa.

Co więcej układ dokumentu nie zawsze podpowiada gdzie jest wymieniony autor, który z członów to imię, a który nazwisko, gdzie tytuł, gdzie organizacja, a gdzie słowa kluczowe. Wyszkolony człowiek da sobie radę nawet z obcym tekstem, ale komputer, który ma wydobyć z jednej strony treść, a z drugiej poprawne informacje na temat autorów i ich przynależności może mieć kłopoty. Co gorsza często stosowaną metodą konwersji do PDF jest skanowanie dające w efekcie tylko obrazki.

Zazwyczaj zaczynamy od wyszukiwania w ogólnodostępnym Internecie, wpisując szukaną fra-

zę czy słowa kluczowe w wyszukiwarce. Gdy ogólne zasoby internetu nie dostarczają nic wartościowego możemy sięgnąć do bibliotek elektronicznych (zwłaszcza tych dających dostęp do pełnych tekstów). Dostajemy jakieś wyniki, przeglądamy je, a następnie po przemyśleniu (i wzbogaceniu naszych teorii) informacje te trafiają do pracy. Rzetelność naukowa wymaga jednak aby korzystając z danych zewnętrznych w jakiś sposób poinformować o tym czytelnika. Korzystamy tu ze znanego od lat mechanizmu cytowania. Staramy się przy tym w sposób jednoznaczny określić źródło inspiracji (w przypadku podawania cytatów odwołując się nawet do numeru strony).

Proces tworzenia bazy bibliograficznej jest złożony i wyczerpujący. Gdy mnie tego uczono (w szkole średniej) sprowadzało się to do tworzenia fiszek, czyli notatek na sztywnym papierze zawierających przede wszystkim autora i tytuł dzieła, rok wydania, liczbę stron. Reszta karty służyła do wynotowywania interesujących cytatów czy umieszczania uwag i wskazówek. Pisząc pracę zamiast sięgać do źródła (które mogło być akurat niedostępne w bibliotece) można było posiłkować się fiszkami i własną pamięcią. Fiszki stanowiły podstawę tworzenia bibliografii.

Trudno powiedzieć jak wygląda ten proces dziś, w dobie komputerów. Ciągle wiele osób, które znam drukują pozyskane w wersji elektronicznej (ważne) źródła i gromadzą je w tematycznych skoroszytach. Czasami ponawia się wyszukiwanie (ale ani Internet nie jest statyczny, ani mechanizmy wyszukiwania takie nie są, więc nie można mieć pewności, że za każdym razem dostaniemy ten sam zestaw wyników. Nie potrafię powiedzieć jak wiele osób korzysta z jakiejś wersji „elektronicznych fiszek” (czyli menedżera bibliografii³).

Tak, czy inaczej tworzenie bibliografii sprowadza się najczęściej do pracowitego przepisywania z papieru lub obrazu na ekranie podstawowych informacji: autor, tytuł, czasopismo, tom (rocznik), numer,... Jest to dosyć pracochłonne (i potrafi prowadzić do pomyłek polegających na przykład na tym, że te same dzieło w bibliografii pojawia się dwa razy).

Wykorzystanie dobrego menedżera bibliografii potrafi pracę nieco ułatwić, ale i tak najtrudniejszą (i najbardziej żmudną czynnością) jest wypełnienie go informacjami. I tu właśnie przydają się metadane. Wszystkie dobre czasopisma i porządne wydawnictwa naukowe, katalogi biblioteczne, a także niektóre dobre księgarnie udostępniają metadane. Dobrze widać to na rysunku 2 pokazującym stronę Dolnośląskiej Biblioteki Cyfrowej opisującej podręcznik *Mechanika płynów* [3]. Dane zawarte w okienku to podstawowe informacje pozwalające „wypełnić fiszkę” (dodać rekord do bazy danych menedżera bibliografii). Oprócz tego, po lewej stronie, można metadane zaimportować w jednym z najczęściej używanych formatów: BibTeX⁴, RIS⁵, RDF⁶ i OAI-PMH⁷.

Dobre⁸ menedżery bibliografii potrafią samodzielnie zaimportować metadane bezpośrednio ze strony WWW prezentującej publikację.

Popatrzmy na to z punktu widzenia autora pracy, który umieszcza w internecie swoje opracowanie: zależy mu na cytowaniu, zależy mu na tym, aby praca została poprawnie zindeksowana przez popularne wyszukiwarki (a zwłaszcza wyszukiwarki publikacji: Google Scholar⁹ czy Microsoft Academic¹⁰) i żeby ewentualne cytowania były zliczane przez różne automaty. Google Scholar może zauważać również cytowania do materiałów umieszczonych na stronach WWW [4], [5].

Patrząc z punktu widzenia „użytkownika” tych informacji (czyli czytelnika opublikowanej w Internecie pracy) liczy się merytoryczna wartość pracy. Wydaje się też zasadnym przypuszczenie,

3 https://en.wikipedia.org/wiki/Comparison_of_reference_management_software

4 <http://www.bibtex.org/>

5 [https://en.m.wikipedia.org/wiki/RIS_\(file_format\)](https://en.m.wikipedia.org/wiki/RIS_(file_format))

6 <https://www.w3.org/RDF/>

7 <https://www.openarchives.org/pmh/>

8 Możliwość samodzielnego importu metadanych ze stron WWW jest częścią definicji dobrego menedżera bibliografii. Do dobrych należą: Zotero (<https://www.zotero.org/>), Mendeley (<https://www.mendeley.com/>), RefWorks (<http://www.proquest.com/products-services/research-tools/refworks.html>) i jeszcze parę innych.

9 <https://scholar.google.pl/>

10 <http://academic.research.microsoft.com/>

The image shows a screenshot of the Dolnośląska Biblioteka Cyfrowa (dLibra) website. The page displays metadata for a book titled "Mechanika płynów" by Henryk Szewczyk. The left sidebar contains a menu with options like "Opis", "Informacje", "Treść", "Treść (nowe okno)", "Pobierz", "Podobne wydania", and "Opcje wyświetlania". The "Pobierz" button is highlighted with a red box. The main content area shows the book's title, author, publisher (Wrocław), and other details. The right side of the page features a book cover image and a "Pokaż treść" button.

Ryc. 2. Metadane na stronie dLibra (Dolnośląska Biblioteka Cyfrowa); po lewej stronie, w obwódce zaznaczono możliwość ich pobierania

że mając do dyspozycji kilka prac o podobnej wartości merytorycznej odgrzywać mogą również względy praktyczne, na przykład łatwość cytowania (pomyślmy o wszystkich akcentach i znakach spoza podstawowego zestawu ASCII). Ten warunek zazwyczaj będzie spełniony w przypadku prac publikowanych w dobrych czasopismach, gorzej z różnego rodzaju raportami czy „preprintami” prac wysłanych do publikacji. Zazwyczaj Autor nie wie/pamięta o konieczności uzupełnienia metadanych.

GDZIE UMIESZCZAĆ METADANE

Jeżeli publikujemy pracę na stronach WWW zadbajmy żeby każda praca miała swoją indywidualną „metryczkę”. Tam powinny być zawarte metadane w postaci czytelnej dla człowieka. Dodatkowo warto zadbać, żeby kod HTML zawierał dodatkowe informacje pozwalające wszystkim przystosowanym do tego programom i wyszukiwarkom internetowych wydobyć podstawowe metadane publikacji (tytuł, autor, rok i miejsce publikacji). Służą do tego, w najprostszym przypadku protokoły COinS¹¹ albo Dublin Core, możemy też skorzystać z DOI¹².

Zarówno pliki PDF czy DOC/DOCX albo ODT pozwalają na zapis metadanych wraz z dokumentem. Warto z tej możliwości korzystać, gdyż większość wyszukiwarek potrafi do nich dotrzeć, a dzięki standaryzacji (zarówno Word jak i LibreOffice korzystają ze standardu Dublin Core¹³). Nie zawsze dane te będą poprawnie wypełnione więc warto w ustawieniach właściwości dokumentu oraz w informacjach o autorze wszystko odpowiednio zmodyfikować. Ponieważ metadane nie mają najlepszej opinii, najłatwiej zdobyć informację jak się ich pozbyć. Jak zajrzeć na strony wsparcia Microsoft znaleźć można taki passus [6]: *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*

11 <http://ocoins.info/>

12 <https://www.doi.org/>

13 <http://dublincore.org/>

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. Nie jest to, oczywiście, najlepsza rada dla tych którzy chcą w Internecie publikować.

Przy odrobinie szczęścia (i po zastosowaniu właściwego oprogramowanie) metadane powinny wyeksportować się do pliku PDF.

Pliki PDF również mają odpowiedni mechanizm do przechowywania metadanych. Dane przechowywane są, podobnie jak w przypadku plików ODT/DOCX w formacie XML. Istnieją programy pozwalające na dodawanie i modyfikację metadanych w plikach PDF. Warto pamiętać o wypełnianiu tych pól gdyż PDF to dziś najpopularniejszy format publikacyjny. Dobre menedżery bibliografii potrafią odczytać te metadane.

Nie jest to głównym tematem tych rozważań, ale publikując ilustracje warto wiedzieć, że najpopularniejsze formaty graficzne (JPG czy PNG) również mogą (a może nawet powinny) zawierać metadane wskazujące na autora i prawa do użytkowania.

MENEDŻERY BIBLIOGRAFII — ZOTERO

Jak mi się wydaje menedżery bibliografii nie są najczęściej używanymi programami. Świadczy o tym nawet brak odpowiedniego hasła w polskiej Wikipedii. W jej angielskiej wersji jest dosyć obszerne hasło *Reference management software*¹⁴ oraz bardzo dobre porównanie cech różnych menedżerów bibliografii: *Comparison of reference management software*¹⁵ natomiast brakuje takich odpowiedników po polsku, a szkoda. Niewiele informacji na temat użycia takiego oprogramowania można znaleźć na polskich, akademickich stronach WWW.

Praca nie jest poświęcona temu tematowi, ale warto wspomnieć o najważniejszych cechach takiego oprogramowania:

1. Potrafią importować metadane ze stron WWW.
2. Najlepsze pozwalają przechowywać dane w chmurze (co pozwala mieć dostęp do nich z kilku komputerów).
3. Integrują się z popularnymi procesorami tekstu (Microsoft Office, LibreOffice, Google Documents).
4. Pozwalają wybrać styl cytowania.

Jednym z bardziej popularnych menedżerów bibliografii jest wspomniane już Zotero. Jest to program, który integruje się z najpopularniejszymi przeglądarkami WWW (Firefox, Google Chrome, Opera, Internet Explorer, Safari), wyposażony jest w rozszerzenia współpracujące z Microsoft Office i LibreOffice, pozwalające na łatwe przeszukiwanie bazy danych, wstawianie cytowań i generowanie bibliografii w odpowiednim formacie. W prosty sposób może być używane z Dokumentami Google.

Jedną z ciekawszych cech Zotero jest to, że może być używany jako zwykły menadżer zakładek, który potrafi przechować obraz strony (*snapshot*) z chwili jej odwiedzania, co pozwala powrócić do niej nawet wtedy gdy przestanie już istnieć.

Inne programy tego typu mają podobne możliwości, natomiast Zotero powstaje w środowisku akademickim jako projekt Open Source. Autorzy planują rozwój programu tak, aby mógł łatwo się integrować z repozytoriami instytucjonalnymi.

PODSUMOWANIE

Przedstawione przemyślenia prowadzą mnie do następujących wniosków:

1. Jesteśmy i będziemy oceniani na podstawie publikacji i liczby cytowań.
2. Analizy te będą przeprowadzały komputery.
3. Cały system informacyjny nowoczesnej uczelni powinien to uwzględniać.
4. Czy tego chcemy, czy nie — jesteśmy fragmentem nowoczesnego świata, gdzie nasi

14 https://en.wikipedia.org/wiki/Reference_management_software

15 https://en.wikipedia.org/wiki/Comparison_of_reference_management_software

klienci (tak studenci jak i potencjalni współpracownicy z przemysłu i nauki) podejmując swoje decyzje kierują się nie tyle dogłębną analizą ale wrażeniem odebranych z kontaktu z nami. Ten kontakt, dziś, jest najczęściej wirtualny — odbywa się przez Internet.

Ad. 1. Celem uczonego jest prowadzić dyskusję z innymi. Pomijając wykłady, seminaria i konferencje (które stanowią istotny, choć trudny do udokumentowania obszar dyskusji), publikacje naukowe są istotą dialogu, w którym ktoś pisze, kto inny czyta i jeżeli dyskusję nawiązuje — pracę cytuje ustosunkowując się do niej.

Urzędnicy odpowiedzialni za naukę nie są w stanie ocenić pracy naukowców. Można, co prawda, powoływać gremia recenzentów, ale wypracowują one swoje decyzje w długotrwałym procesie. Uwzględniając zaś fakt, że decyzje finansowe podejmowane są w cyklu rocznym — korzystanie z analiz komputerowych jest nieuniknione. Nie są one doskonałe, ale — w pewnym sensie — identyczne dla wszystkich, zatem „obiektywne”.

Za pewny ideał można uznać brytyjski system ocen nauki¹⁶ [7]. Ale jak popatrzeć na planowane wydatki (ponad 200 milionów Funtów Brytyjskich) to zapół do pójścia w ich ślady zamiera.

Ad. 2. Zatem jedno co możemy zrobić (poza publikowaniem coraz lepszych prac) to zadbać o ich właściwą widoczność i dostępność w Internecie. Możemy ranking Webometrics (o którym wspomniałem w rozdziale [sec:stan]) uznać za wielką pomyłkę, ale nie są tu istotne nasze odczucia: znalezienie się blisko czołówki tego rankingu jest znacznie lepsze niż znalezienie się gdzieś w ogonie. Pomijając wszystko, ranking ten mierzy dostępność **treści** oraz chęć innych do odwoływania się do tych treści (takie internetowe cytowanie). Jeżeli nasze strony WWW będą zawierały tylko i wyłącznie informacje porządkowe i administracyjne — nikt z zewnątrz nie będzie do nich podlinkowywał. Wreszcie wszystko co publikujemy powinno wychodzić z naszej domeny. A o domenę internetową należy dbać tak samo jak o godło czy logo.

Zatem wszelkie repozytoria uczelniane powinny być umieszczone pod uczelnianymi adresami, a adresy te nie powinny zmieniać się. Podobnie z nazwą instytucji. Za nazwą Politechnika Wrocławska jest 70 lat tradycji. I na rynku krajowym ona coś znaczy. A nazwa angielska? Technical University of Wrocław, Wrocław University of Technology, Wrocław University of Science & Technology... Z pewnego punktu widzenia jest to drobiazg. Wystarczy minimalna modyfikacja oprogramowania. Ale to nie nasze oprogramowanie dokonuje analiz, a dokonujący analiz wcale nie śpieszą się ze zmianami.

Podobnie z wewnętrzną strukturą organizacyjną. Dziś już nie można powołać jednostki nadając jej tylko polską nazwę. Trzeba przedyskutować i przedstawić dwie (albo może i trzy nazwy): polską, angielską i niemiecką. Z punktu widzenia przyszłych studentów to może i ukraińską.

Publikując pracę podpisujemy ją własnym nazwiskiem ale i nazwą instytucji. Czasami również nazwą jednostki organizacyjnej. Jak daleko możemy się posunąć? Zazwyczaj mamy do dyspozycji tylko dwa stopnie: uczelnia i jednostka organizacyjna. Nawet jeżeli czasami istnieje możliwość prowadzenia tego podziału głębiej — bibliograficzne bazy danych niechętnie udostępniają takie informacje. Z punktu widzenia uczelni ważne jest żeby nazwisko autora było poprawnie kojarzone z nią. Skoro jednak chcemy korzystać z baz bibliograficznych do celów zarządzania — potrzebne są na nieco dokładniejsze informacje. Ograniczenia wprowadzają również redaktorzy czasopism i organizatorzy konferencji. Wydaje się więc ważne aby konsekwentnie na zewnątrz posługiwać się nazwą uczelni i wydziału (w sposób jednolity tłumaczoną na języki obce).

Starajmy się wreszcie ułatwiać życie wszystkim, którzy mogą zechcieć nas zacytować: zwolnijmy ich ze żmudnego przepisywania naszych nazwisk — nich mają możliwość skorzystania z metadanych.

Ad. 3. W pionierskich latach Internetu (mam tu na myśli wczesne lata dziewięćdziesiąte) każdy kto potrafił, tak jak potrafił budował swoją stronę domową czy serwis WWW swojej jednostki. Pewną przewagę nad innymi dawało posiadanie własnych stron WWW.

Dziś strony WWW posiada każdy, więc nie daje to nikomu żadnej przewagi. Serwisy są bar-

16 <http://www.ref.ac.uk/>

dziej albo mniej wymyślne, ale „wymiślność” serwisu nic nie mówi o randze uczelni. O tym może powiedzieć coś zawartość informacyjna. Dzisiaj znakomita większość stron WWW Politechniki Wrocławskiej udostępnia serwis o charakterze informacyjnym (Co? Gdzie? Kiedy?) zawiera niezbędny zestaw informacji regulujących działalność (statut, regulamin studiów, i dostępne wyłącznie wewnątrz zarządzenia). Do tego dochodzi niezbędny zestaw informacji porządkowych dla studentów. Nie wspominam tu o treściach „wewnętrznych” dostępnych po zalogowaniu, biorę pod uwagę tylko to co jest dostępne dla każdego z zewnątrz.

Oczekiwanie, że pojawi się grupa blogerów naukowych pisząca ciekawie na tematy ich interesujące nie ma szans na spełnienie. Zdarzają się dobrzy bloggerzy naukowci, ale jest ich bardzo mało.

Wydaje się, że wszystkie jednostki Uczelni powinny prezentować się (w dosyć jednolity sposób) w Internecie przedstawiając po polsku i po angielsku główne kierunki działalności i najważniejsze osiągnięcia. Każdy pracownik powinien mieć co najmniej „wizytówkę” zawierającą dane teleadresowe i listę najważniejszych/wszystkich publikacji. Lista ta powinna być automatycznie uaktualniana na podstawie publikacji rejestrowanych przez służby biblioteczne uczelni. Ze stron jednostki powinna istnieć również możliwość „ściągnięcia” pełnych tekstów prac. (Zakładam oczywiście, że takie spisy są odpowiednio wyposażone w metadane.)

Jeżeli chodzi o informacje o dydaktyce — oprócz kalendarza dydaktycznego (terminy i miejsca zajęć, terminy konsultacji) strona powinna zawierać również podstawowe informacje na temat prowadzonych kursów (minimum „karta kursu”).

Powinna istnieć możliwość wypełniania przez pracownika swojej witryny dodatkowymi informacjami.

Ad. 4. Informacje o charakterze naukowym (czy dydaktycznym) nie muszą interesować wszystkich. Na uczonych oprócz obowiązków rozwoju nauki i kształcenia studentów spoczywa również obowiązek popularyzowania nauki. To właśnie może być najlepszą treścią o charakterze marketingowym przyciągającą tak studentów jak i klientów z małego i średniego biznesu. Muszą oni zobaczyć normalnych, otwartych na innych ludzi.

BIBLIOGRAFIA

- [1] J. E. Hirsch, “An index to quantify an individual’s scientific research output,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 46, pp. 16569–16572, Nov. 2005.
- [2] C. Duhigg, “How Companies Learn Your Secrets,” *The New York Times*, Feb-2012.
- [3] H. Szewczyk and K. Jeżowiecka-Kabsch, *Mechanika płynów*. Oficyna Wydawnicza Politechniki Wrocławskiej, 2001.
- [4] W. Myszka, “Ma to sens...,” *Katedra Mechaniki i Inżynierii Materiałowej*, 08-Jul-2014. [Online]. Available: <https://kmim.wm.pwr.edu.pl/blog/2014/07/08/ma-to-sens/>. [Accessed: 18-Sep-2016].
- [5] W. Myszka, “Zabawy z metadanymi,” Dec-2015. [Online]. Available: <https://kmim.wm.pwr.edu.pl/myszka/2015/12/05/zabawy-z-metadanymi/>. [Accessed: 17-Sep-2016].
- [6] “Remove hidden data and personal information by inspecting documents — Office Support,” 2016. [Online]. Available: <https://support.office.com/en-us/article/Remove-hidden-data-and-personal-information-by-inspecting-documents-356b7b5d-77af-44fe-a07f-9aa4d085966f>. [Accessed: 18-Sep-2016].
- [7] “Research Excellence Framework review,” Department for Business, Energy & Industrial Strategy, 2016.
- [8] “Wskaźnik Hirscha,” Jan-2016. [Online]. Available: https://pl.wikipedia.org/w/index.php?title=Wska%C5%BAnik_Hirscha&oldid=44787483. [Accessed: 18-Sep-2016].

SUMMARY

BIBLIOMETRICS AND MORE: A FEW OFF TOPIC COMMENTS

The work sums up 25 years of experience gained during the development and use of Internet services at the Wrocław University of Science and Technology, and active participation in the creation and development of the Web site of the Institute, the Department, and finally Department. The author attempts to formulate not very precise conclusions regarding exposure metadata in the presentation of scientific publications on the web university.