



Aneta Drabek
Biblioteka Uniwersytetu Śląskiego

Do użytku bibliometrycznego... Niebibliograficzne wykorzystanie baz bibliograficznych



Dr Aneta Drabek - kustosz dyplomowany, absolwentka Bibliotekoznawstwa i Informatyki Uniwersytetu Śląskiego. Współtwórcza bazy Arianta - Naukowe i Branżowe Polskie Czasopisma Elektroniczne. Specjalistka w zakresie baz danych, cytowań i czasopism elektronicznych. Autorka publikacji dotyczących bibliometrycznej analizy oceny czasopism naukowych.

Streszczenie: Bibliograficzne bazy danych należą do najważniejszych elementów warsztatu każdego naukowca. Jednakże dane zgromadzone w tych źródłach dają o wiele więcej możliwości niż tylko sporządzanie kwerend tematycznych. Artykuł ukazuje przykłady zastosowania baz bibliograficznych do badań bibliometrycznych i naukometrycznych (np. identyfikacja wiodących czasopism w niektórych dziedzinach, badania struktury i rozwoju poszczególnych dziedzin, w tym analizy trendów i prognoz, a także udział poszczególnych krajów e-nauczania światowej). Praca ukazuje warunki konieczne do stworzenia polskiego indeksu cytowań.

Słowa kluczowe: bibliograficzne bazy danych, bibliometria, polski indeks cytowań

Summary: Bibliographic databases are one of the most important tools of every scientist. However, the data collected in these resources allow for a lot more than making thematic queries. The purpose of this article is to identify the potential of Polish databases for doing bibliometric research. The paper reviews some applications of various bibliographic databases in bibliometric and scientometric research, such as identifying leading journals in certain fields of knowledge, investigating the structure and development of particular fields, including a trend analysis and forecasting, as well as the study of the contribution of various countries to world science. The article presents the conditions necessary to create a Polish citation index.

Keywords: bibliographic databases, bibliometrics, Polish citation index

[Prezentacja](#)

Wprowadzenie

Bibliograficzne bazy danych na nośnikach maszynowych, powstałe w latach 60. ubiegłego wieku jako narzędzia ułatwiające wydawanie przeglądów dokumentacyjnych, bardzo szybko stały się podstawowymi źródłami informacji. Obecnie trudno wyobrazić sobie poruszanie się wśród milionów dokumentów bez tych banków wiedzy¹.

Oczywiście na tym praktyczne aspekty wykorzystania baz bibliograficznych się nie kończą, ponieważ potencjał w nich zawarty — w postaci przeróżnych danych — pozwala na przygotowywanie różnorodnych analiz opartych na badaniu stanu piśmiennictwa naukowego. Badania te mogą pomóc w stworzeniu ilościowej

¹ STEFANIAK, B. Niekonwencjonalne zastosowania bibliograficznych baz danych, cz. 1. *PTINT Praktyka i Teoria Informatyki Naukowej i Technicznej* 1995, nr 1, s. 3.



Bibliograficzne bazy danych i ich rola w rozwoju nauki

II Konferencja naukowa Konsorcjum BazTech

Poznań, 17-19 kwietnia 2013



charakterystyki struktury nauki oraz pomóc w określeniu dynamiki i tendencji jej rozwoju. Mowa tu oczywiście o analizach biblio- i naukometrycznych.

Bibliometria pozwala na ukazywanie stanu i tendencji rozwoju piśmiennictwa metodami statystycznymi na podstawie bibliografii, spisów i baz danych, „naukometria” bada natomiast wybrane obszary systemu nauki (a takim obszarem jest piśmiennictwo naukowe) przy zastosowaniu metod ilościowych².

Rozwój bibliometrii jest związany z rozwojem bibliografii, a pierwsze badania, które dzisiaj określamy jako bibliometryczne, właśnie na takich tradycyjnych spisach się opierały, nosząc wówczas nazwę „bibliografii statystycznej”. Były to proste analizy polegające na zliczaniu wydawanych publikacji pod kątem wybranych kryteriów³. Stworzenie bibliograficznych baz danych, czyli połączenie dużej ilości danych z odpowiednimi narzędziami do ich przeszukiwania, spowodowało rozkwit tego typu badań i ich coraz szersze wykorzystanie w polityce naukowej państw, ministerstw, uczelni, redakcji czasopism itp. W badaniach tych bierze się pod uwagę wszystkie elementy opisu bibliograficznego, takie jak: nazwiska autorów, tytuł publikacji, słowa kluczowe, słowa z abstraktów, afiliację (instytucja, miasto, kraj), symbole klasyfikacji tematycznej, język, datę wydania, typ publikacji, tytuł czasopisma, dane z bibliografii załącznikowej (cytowania). I choć w ostatnich latach największą karierę zrobiły właśnie cytowania oraz wskaźniki, które na podstawie tych informacji są wyliczane (*Impact Factor*, pięcioletni *Impact Factor*, *Index Hirscha*, *Real Impact Factor*, *Predicted Impact Factor* i in.), to jednak warto zauważyć, że możliwości wykorzystania baz są o wiele większe. Drzemie w nich wręcz nieograniczony potencjał analizowania piśmiennictwa naukowego i tworzenia na tej podstawie złożonych opracowań.

W ostatnich latach zaobserwować można coraz częstsze używanie metod bibliometrycznych (czy też naukometrycznych) w polityce naukowej. Trend ten zauważalny jest nie tylko w Polsce, lecz także w wielu innych krajach. Na popularność tych metod wpływa przede wszystkim ich obiektywny charakter, możliwość uzyskania i przetwarzania wielkich zbiorów danych oraz opracowana metodologia. Poza tym wyniki uzyskuje się stosunkowo łatwo i szybko. Możemy dzięki nim otrzymać aktualny obraz nauki, przewidywać i obserwować nowe trendy badawcze, wpływać na powstanie międzynarodowych grup naukowców prowadzących badania w określonym kierunku, badać wkład poszczególnych państw czy instytucji w rozwój nauki itp. Możliwości stosowania analiz ilościowych do monitorowania rozwoju nauki, a następnie wykorzystania ich do kształtowania polityki naukowej są bardzo duże i z pewnością to one będą dominowały w najbliższych latach⁴.

² Tamże, s. 3.

³ NOWAK, P. *Bibliometria, webometria. Podstawy. Wybrane zastosowania*. Poznań, 2006, s. 16.

⁴ DRABEK, A. Wykorzystanie bibliometrii w polityce naukowej. *Biuletyn EBIB* [on-line]. 2012, nr 3(130) [Dostęp 28.04.2013]. Dostępny w World Wide Web: http://www.nowyebib.info/images/stories/numery/130/130_drabek.pdf.



Zalety baz bibliograficznych

Dlaczego akurat bazy bibliograficzne są dobrym źródłem danych do tego typu analiz? Odpowiedź na to pytanie składa się z kilku czynników. Omówmy je po kolei.

Po pierwsze, duża ilość danych. Rozwój technologii informatycznych pozwala nie tylko na zapis milionów rekordów, lecz także dostarcza odpowiednich narzędzi umożliwiających wieloaspektowe wyszukiwania. Duży zbiór danych wpływa na jego reprezentatywność i upoważnia do wyciągnięcia uogólniających wniosków.

Po drugie, wiarygodność danych, na którą składa się rzetelność w opracowywaniu danych (najczęściej z autopsji), przygotowanie zawodowe osób wprowadzających dane (najczęściej są to bibliotekarze i pracownicy informacji), wieloetapowy proces przygotowania rekordów (wprowadzanie, korekta, autoryzacja).

Po trzecie, niezmienna struktura danych, czyli każdy rekord wprowadzany jest tak samo; wypełniane są te same pola, wybrane czasopisma indeksowane są w całości (na zasadzie *cover-to-cover*).

Po czwarte, właściwy dobór danych. O tym, czy odpowiednie rekordy pojawią się w bazie, decyduje tylko ich zawartość treściowa, a nie względy pozamerytoryczne.

Po piątę, dane te są stabilne, czyli w dowolnym momencie można powtórzyć daną analizę, a wynik powinien wyjść taki sam.

Po szóste, powszechny dostęp. Bazy bibliograficzne są zazwyczaj dostępne on-line. Z większości polskich baz danych można korzystać bezpłatnie, co także ułatwia prowadzenie badań bibliometrycznych.

Dlaczego nie Google Scholar?

Warto dodać, że polecany ostatnio (m.in. przez Narodowe Centrum Nauki) Google Scholar nie jest bazą danych, tylko wyszukiwarką opartą na światowych zasobach naukowych i edukacyjnych. Nie dotyczy jej prawie żaden (poza ostatnim) z wymienionych powyżej argumentów. Dane uzyskane w ten sposób pochodzą z materiałów niepodlegających ocenie (np. różnej jakości publikacje konferencyjne, prace seminaryjne i raporty wewnętrzne) i są pokazywane na równych prawach z publikacjami recenzowanymi. Te same informacje, zamieszczane na różnych stronach, mogą sztucznie podwyższać wskaźniki. Google Scholar obejmuje tylko część internetowych zasobów, odczuwalny jest zwłaszcza brak materiałów z zakresu nauk humanistycznych i społecznych. Niewiele można znaleźć starszych publikacji, zwłaszcza tych, co do których majątkowe prawa autorskie jeszcze nie wygasły. Dane aktualizowane są rzadziej niż bazy danych⁵.

⁵ RACKI, G., DRABEK, A. Cytowania i wskaźnik Hirscha. Gdzie szukać, jak obliczać? *Forum Akademickie* [on-line]. 2013, nr 2 [Dostęp 28.04.2013]. Dostępny w World Wide Web: <http://forumakademickie.pl/fa/2013/02/cytowania-i-wskaznik-hirscha-gdzie-szukac-jak-obliczac/>.



Analizy bibliometryczne na podstawie baz bibliograficznych

Jeśli decydujemy się na wykorzystanie baz bibliograficznych do tego typu analiz, kluczowy staje się wybór odpowiedniego źródła, które zapewni reprezentatywność przeprowadzonych badań. Nie wszystkie bazy są takie same i nie każdej da się użyć do tych samych celów, ale wszystkie mają możliwości, nad którymi warto się zastanowić. Prześledźmy zatem najpopularniejsze typy analiz stworzonych na podstawie danych z baz.

1. Ocena bibliometryczna aktywności państwa

Pojawienie się dużych interdyscyplinarnych baz danych pozwala na ukazanie dynamiki aktywności badawczej państw. Na podstawie liczby publikacji tworzy się rankingi ukazujące wkład poszczególnych krajów i ich udział w światowej nauce. Dzięki takim zestawieniom łatwo zauważyć rosnącą rolę Chin, Korei Południowej czy Tajwanu w międzynarodowej aktywności badawczej. Analizy takie dają możliwość ukazania mocnych stron, czyli tych dyscyplin, w przypadku których dany kraj jest liderem. Ważnym elementem jest także ustalenie mapy współpracy naukowej między krajami. W Polsce wyniki badań dotyczących wkładu poszczególnych państw opublikowała Irina Marshakova-Shaikievich⁶. Autorka omówiła także wkład Polski w rozwój nauki na przełomie XX i XXI w., a także zidentyfikowała najaktywniejsze polskie ośrodki badawcze⁷. Szczegółową analizę polskiej działalności badawczo-rozwojowej przedstawili Agnieszka Olechnicka i Grzegorz Płoszaj. Badacze ukazali potencjał publikacyjny Polski w podziale na poszczególne podregiony, współpracę ośrodków naukowych oraz specjalizację dziedzinową podregionów⁸.

2. Analiza dziedziny wiedzy

Bibliograficzne bazy danych doskonale nadają się do kompleksowej oceny dziedziny wiedzy lub jednej z jej dyscyplin. W zależności od zasięgu bazy analiza taka może obejmować jeden kraj, porównawczo kilka państw lub całą naukę światową. W ramach badań udaje się ustalić tendencje rozwojowe danej dziedziny, najważniejsze ośrodki reprezentujące czołowe fronty badawcze, nakreślić mapę współpracy między poszczególnymi ośrodkami czy badaczami. Wreszcie dają one podstawę do wyróżnienia czołowych czasopism z danej dziedziny. Do analiz tego typu wykorzystuje się bazy dziedzinowe lub interdyscyplinarne. Uzyskane wyniki mogą zostać spożytkowane jako pomoc w tworzeniu polityki naukowej państwa, wzmocnieniu (dofinansowaniu) wybranych branż czy też w ustaleniu przyczyn ich złej kondycji. Przykładem takich analiz są badania nauk geologicznych przeprowadzone przez Grzegorza Rackiego⁹.

⁶ MARSHAKOVA-SHAIKEVICH, I. *Badania ilościowe nauki. Podejście bibliometryczne i webometryczne*. Poznań, 2009, s. 193–227.

⁷ Tamże, s. 228–243.

⁸ OLECHNICKA, A., PŁOSZAJ, G. *Polska nauka w sieci. Przestrzeń nauki i innowacyjności. Raport z Badań* [on-line]. Warszawa, 2008 [Dostęp 28.04.2013]. Dostępny w World Wide Web: http://www.espon.pl/files/10_2/2/Polska%20nauka%20w%20sieci.pdf.

⁹ RACKI, G. Geologia Polska w świetle Science Citation Index. *Przegląd Geologiczny* 1996, T. 44, nr 9, s. 928–930.



3. Analiza instytucji naukowej

Wykorzystując informacje zawarte w interdyscyplinarnych bazach danych, można przeprowadzić analizę wybranej instytucji naukowej: przedstawić udział pracowników danej jednostki w nauce światowej czy krajowej, zidentyfikować najsilniej reprezentowane dziedziny, określić współpracę krajową i międzynarodową. Na podstawie tych informacji udaje się także wyróżnić najważniejsze czasopisma, w których publikują pracownicy, co może pomóc w kształtowaniu polityki gromadzenia tego typu materiałów. Wreszcie część tych materiałów może zostać wykorzystana w decyzjach związanych z awansami. Analizy takie przeprowadzane są na potrzeby uczelni, a także przeróżnych rankingów szkół wyższych. Niektóre uczelnie przygotowują — na podstawie baz bibliograficznych — bardzo obszerne zestawienia działalności naukowej¹⁰.

4. Analiza czasopism naukowych

Analizy bibliometryczne czasopism naukowych często pojawiają się w piśmiennictwie naukowym. Dotyczyć mogą jednego tytułu, porównania dwóch lub więcej tytułów z danej dziedziny czy też opisują wydawnictwa naukowe publikowane w danym kraju. Do takich badań można wykorzystać nawet niewielkie bazy bibliograficzne, które — indeksując całą zawartość danego tytułu — pozwalają na przygotowanie takiej analizy. Rzetelną oceną czasopism naukowych zainteresowanych jest kilka grup odbiorców: bibliotekarze traktują je jako pomoc w selekcji prenumerowanych periodyków; naukowcy muszą się orientować w najważniejszych tytułach ze swojej specjalności; redakcje czasopism często same zlecają lub wykonują podobne zestawienia i na jej podstawie kształtują profil swojego pisma; wreszcie ministerstwa czy inne organizacje finansujące naukę zainteresowane są takimi danymi, które następnie mogą wykorzystać do rzetelnej oceny i przyznania (lub nie) dofinansowania. Do analizy jednego tytułu bierze się pod uwagę trzy rodzaje cech, które następnie składają się na jego ocenę. Pierwsze z nich dotyczą autorów: kwalifikacje, afiliacje, współpraca autorska, płeć. Drugi element to informacje dotyczące zawartości treściowej artykułów, czyli słowa kluczowe, słowa z abstraktów, symbole klasyfikacyjne. Trzecia z cech to cytowania. Analiza tych danych pozwala nie tylko na stworzenie dokładnej charakterystyki danego tytułu, lecz także ułatwia prace radom naukowym¹¹. Badania polegające na porównywaniu ze sobą dwóch lub więcej tytułów i stworzone na tej podstawie rankingi mogą być przydatne jako jeden z elementów oceny czasopism.

¹⁰ KOMPERDA, A., URBAŃCZYK, B. Analiza naukometryczna dorobku publikacyjnego uczelni — kryteria, statystyki, raporty. *PTINT Praktyka i Teoria Informacji Naukowej i Technicznej* 2010, nr 3, s. 46–54; BIRSKA, R., BUSSE-TURCZYŃSKA, E. Analizy bibliometryczne dla środowiska naukowego generowane na podstawie bazy danych „Bibliografia publikacji pracowników Uniwersytetu Medycznego w Lublinie”. *Forum Bibliotek Medycznych* [on-line]. 2009, nr 2, s. 523–535 [Dostęp 28.04.2013]. Dostępny w World Wide Web: http://www.bg.umed.lodz.pl/index.php?option=com_content&task=category§ionid=9&id=33&Itemid=59.

¹¹ DRABEK, A. *Bibliometryczna analiza czasopism naukowych w dziedzinie nauk społecznych*. Toruń, 2010, s. 48–49; DRABEK, A., TOMASZCZYK, J. Czasopismo „Praktyka i Teoria Informacji Naukowej i Technicznej” w świetle danych bazy CYTBIN. W: PIETRUCH-REIZES, D. (red.) *Zarządzanie informacją w nauce*. Katowice, 2008, s. 365–375.



5. Analiza dorobku jednego uczonego lub zespołu badawczego

Bazy bibliograficzne są doskonałym narzędziem pozwalającym na ukazanie dorobku jednego uczonego lub jakiejś wybranej grupy. Przygotowane zestawienie wyszczególnia zazwyczaj liczbę publikacji w podziale na poszczególne typy (książki, artykuły, materiały konferencyjne itp.), liczbę cytowań, a także wskaźniki służące ocenie (*Impact Factor*, punktacja Ministerstwa Nauki i Szkolnictwa Wyższego). Analizy takie przygotowuje się na potrzeby awansów, okresowej weryfikacji pracowników naukowych, konkursów na stanowiska czy też innych instytucji lub organizacji (nagrody, członkostwo w towarzystwach naukowych). Są one również elementem wniosków składanych np. do Narodowego Centrum Nauki czy innych instytucji finansujących naukę¹².

6. Analiza cytowań

Cytowania w ostatnich latach stały się jednym z podstawowych wskaźników używanych do oceny (ludzi, jednostek, państw). Wykorzystuje się zarówno liczbę cytowań, jak i te współczynniki, które wykorzystują dane o cytowaniach. Metoda ta ma swoich zwolenników i przeciwników, a dyskusje na jej temat od lat poruszane są na łamach czasopism. Jednakże wszystko wskazuje na to, że ani Ministerstwo Nauki i Szkolnictwa Wyższego, ani podległe mu urzędy nie zrezygnują z tego czynnika stanowiącego część oceny parametrycznej. Bazy bibliograficzne — o ile zawierają te dane — są z pewnością najlepszym źródłem informacji o cytowaniach. Dane te wykorzystać można zarówno bezpośrednio, jak i do tworzenia własnych (lokalnych) wskaźników.

Polskie bazy cytowań

Analizy cytowań jeszcze do niedawna przeprowadzane były wyłącznie przy użyciu bazy Web of Science (Thomson Reuters). W ostatnich latach pojawiły się nowe źródła informacji, które również zawierają dane pochodzące z bibliografii załącznikowej. Ta potrzeba zrodziła się zwłaszcza w kontekście nauk społecznych i humanistycznych, które mają słabą reprezentację w Web of Science (ok. 30% czasopism indeksowanych to czasopisma z tego obszaru badawczego), a przede wszystkim narodowego charakteru tych nauk, które są ściśle związane z konkretnym obszarem językowym. Luki tej nie wypełniła również baza Scopus (Elsevier). Wykorzystywanym źródłem danych stała się wyszukiwarka Google Scholar, ale biorąc pod wymienione powyżej wady, do wyników uzyskanych w ten sposób należy podchodzić z dużą ostrożnością.

W związku z tymi potrzebami twórcy niektórych baz bibliograficznych także zaczęli dołączać informacje pobrane z bibliografii. Wydaje się, iż staną się one elementem obowiązkowym każdego rekordu, a ich wyszukiwanie tych danych może nawet zacząć dominować nad poszukiwaniami tradycyjnymi (bibliograficznymi). Również w Polsce mamy do czynienia z bazami zawierającymi cytowania.

¹² MAMORSKA, K., SIEK, B. O zależnościach między bibliometrią a bibliografią na przykładzie bibliografii GUMed. *Forum Bibliotek Medycznych* [on-line]. 2011, nr 1(7) [Dostęp 28.04.2013], s. 557–563. Dostępny w World Wide Web: <http://www.bg.umed.lodz.pl/images/stories/wydawnictwa/forum/ForumNr7.pdf>.



Najstarszym polskim indeksem cytowań był Toruński Indeks Cytowań Socjologii Polskiej opracowany w Instytucie Socjologii Uniwersytetu Mikołaja Kopernika w Toruniu w latach 90. ubiegłego wieku. Indeks ten uwzględniał 4 polskie czasopisma socjologiczne¹³.

Kolejna inicjatywa utworzenia bazy bibliograficznej z indeksem cytowań powstała na Uniwersytecie Śląskim, gdzie stworzono bazę Polska Literatura Humanistyczna — Arton. Zawierała ona dane z 26 polskich czasopism filologicznych¹⁴. Obecnie baza jest zawieszona.

W 2002 r. polska bibliologia wzbogaciła się o bazę CYTBIN¹⁵. Baza rejestruje 6 polskich periodyków z zakresu bibliotekoznawstwa i informacji naukowej¹⁶.

Tworzona przez biblioteki uczelni ekonomicznych BazEkon¹⁷ to obecnie największy polski indeks cytowań indeksujący informacje z ponad 350 czasopism i serii wydawniczych. Oprócz standardowych danych, takich jak: autorzy, tytuł, abstrakt, słowa kluczowe, źródło, każdy rekord zawiera opisy z bibliografii załącznikowej i przypisów. Dodatkowo baza ma specjalną wyszukiwarkę¹⁸, dzięki której wykonać można analizę cytowań autora lub czasopisma. Wyniki pojawiają się w postaci bardzo czytelnej tabeli, a także podsumowania następujących wskaźników: liczba cytowań, indeks Hirscha, liczba cytowanych publikacji, średnia liczba cytowań publikacji, średnia liczba cytowań na rok, liczba publikacji cytowanych więcej niż raz, zakres czasowy.

W Krakowie w 2005 r. powstał Indeks Cytowań Historiografii Mediów Polskich, w którym zarejestrowano 24 627 dokumentów i 63 811 cytowań¹⁹. Analiza zawartych tam danych pozwoliła na wniosek, że w stosunku do historii zastosować można te same prawa bibliometryczne jak w przypadku nauk przyrodniczych²⁰.

¹³ WINCŁAWSKA, B. M. Indeks cytowań socjologii polskiej (Założenia ideowe i omówienie pierwszych wyników). *Zagadnienia Naukoznawstwa* 1995, nr 3–4, s. 243–246.

¹⁴ WAGA, M., DRABEK, A. Arton — baza cytowań polskiej literatury humanistycznej (stan prac nad bazą). *Zagadnienia Naukoznawstwa* 2002, T. 38, z. 151–152, s. 83–89.

¹⁵ CYTBIN [on-line]. Katowice: Instytut Bibliotekoznawstwa i Informacji Naukowej UŚ, 2000–2013 [Dostęp 05.04.2013]. Dostępny w World Wide Web: <http://www1.bg.us.edu.pl/bazy/cytbin/>.

¹⁶ STEFANIAK, B. i in. Bibliograficzna baza danych — CYTBIN. *PTINT Praktyka i Teoria Informacji Naukowej i Technicznej* 2005, nr 4, s. 3–9.

¹⁷ BazEkon [on-line]. Kraków: Biblioteka Główna Uniwersytetu Ekonomicznego, 2010 [Dostęp 28.04.2013]. Dostępny w World Wide Web: http://kangur.uek.krakow.pl/bazy_ae/bazekon/nowy/.

¹⁸ Cytowania w BazEkon [on-line]. 2012 [Dostęp 28.04.2013]. Dostępny w World Wide Web: http://kangur.uek.krakow.pl/bazy_ae/bazekon/cytowania/.

¹⁹ KOLASA, W. M. Retrospektywny indeks cytowań w humanistyce. Koncepcja, metoda, zastosowanie. *Przegląd Biblioteczny* 2011, z. 4, s. 466–486.

²⁰ KOLASA, W. M. Historia mediów polskich w świetle wskaźników bibliometrycznych. *Zeszyty Prasoznawcze* 2011, s. 3–4, s. 14.



Kolejną polską bazą, która rejestruje bibliografie załącznikowe, jest BazTech. Spośród ponad 250 tys. artykułów 105 tys. rekordów zostało uzupełnione o te ważne dane²¹.

Wykorzystanie cytowań do oceny parametrycznej zaczyna mieć coraz większe znaczenie. Nic więc dziwnego, że twórcy bibliograficznych baz danych podejmują decyzję o wzbogaceniu źródła o te informacje oraz o specjalne narzędzia, które ułatwiają wyszukiwania i przygotowują gotowe raporty zawierające różnorodne wskaźniki. W tym kontekście warto zastanowić się, czy powinien powstać ogólnopolski indeks cytowań.

Polski indeks cytowań

Wiele osób, które zgłasza się do bibliotek z prośbą o przygotowanie wykazu cytowań, po jego otrzymaniu nie kryje rozczarowania. Uczucie to związane jest przede wszystkim z niedostateczną — ich zdaniem — liczbą cytowań. A przecież przeglądając różne publikacje, natykają się na powołania do własnych prac. Będąc specjalistami w jakiejś dziedzinie, są w stanie podać nazwiska osób, które na pewno z ich publikacji korzystały. Czemu więc informacji tych nie ma w zestawieniu? Przede wszystkim dlatego, że nie istnieje wskaźnik pod nazwą „liczba cytowań”. Nie istnieje dlatego, że nie ma bazy, która zawierałaby całe światowe piśmiennictwo. To, co wyszukujemy, pochodzi więc ze źródła, które stanowi zaledwie skromną reprezentację międzynarodowej literatury. Cała reszta pozostaje jako zapis w papierowej wersji książki czy czasopisma i jej odnalezienie jest prawie niemożliwe. Problem ten jest zwłaszcza nasilony w przypadku nauk humanistycznych i społecznych. Najbardziej znany światowy indeks cytowań — Web of Science — jest zdominowany przez nauki ścisłe, które stanowią blisko 90% publikacji. Nauki społeczne to ok. 8% prac, a humanistyczne niewiele więcej niż 2%²². Jeśli chodzi o polskie czasopisma, to w przypadku tych obszarów badawczych, mamy ich w tej bazie symboliczną liczbę — 8 tytułów z nauk humanistycznych i 9 z nauk społecznych. Większość z tych periodyków indeksowana jest w bazie od niedawna. Jeśli dodamy do tego fakt, że większość prac humanistycznych i spora część społecznych powstaje w językach narodowych, to otrzymujemy odpowiedź na pytanie: „dlaczego tak mało?”. Podobnie rzecz się ma z drugą międzynarodową bazą, czyli Scopusem. Sytuację ratuje trochę Google Scholar, wyławiając cytowania z prac umieszczonych w sieci, ale wyszukiwarka ta też ma swoje ograniczenia i wyniki uzyskane za jej pomocą należy potraktować z odpowiednim umiarem. Może więc po prostu stworzyć własny polski indeks cytowań? Pomysł oczywiście nie jest nowy. Bazy takie powstają w różnych krajach: Hiszpanii, Chinach, Tajwanie czy Indiach²³.

²¹ BazTech — Baza danych o zawartości polskich czasopism technicznych [on-line]. Konsorcjum BazTech [Dostęp 28.04.2013]. Dostępny w World Wide Web: <http://baztech.icm.edu.pl/>.

²² MARSHAKOVA-SHAIKEVICH, I. dz.cyt., s. 194.

²³ Cyt. za: KOLASA W. M. *Retrospektywny...* dz.cyt., s. 468.



Bibliograficzne bazy danych i ich rola w rozwoju nauki

II Konferencja naukowa Konsorcjum BazTech

Poznań, 17-19 kwietnia 2013



Rozpatrzmy w takim razie, jakie warunki musiałyby zostać spełnione, aby taka baza powstała i mogła być wiarygodnym źródłem cytowań i przygotowywanych na ich podstawie ocen.

1. Dobór materiałów do bazy

Podstawowa kwestia, jaką należą rozstrzygnąć, to dobór publikacji do indeksowania tak, aby powstała baza była reprezentatywna dla wszystkich dyscyplin. Dla czasopism sprawa jest o tyle prosta, że można wykorzystać listę ocenioną przez MNiSW. Dużo trudniej sprawa przedstawia się w przypadku wydawnictw zwartych. Należy pamiętać, że w naukach humanistycznych połowa cytowań pochodzi z książek. W związku z tym trzeba rozstrzygnąć, które książki wziąć pod uwagę, a które pominąć.

2. Wielkość bazy

Wszystkim zależy, żeby powstała baza jak najszybciej mogła spełniać swojej funkcje, a więc liczba indeksowanych rocznie materiałów powinna być jak największa. Gdyby tylko przyjąć, że w każdym z 2000 czasopism (czyli mniej więcej tyle, ile jest na liście B) rocznie publikuje się 25 artykułów, a w każdym z nich jest 15 opisów bibliograficznych, to daje ok. 750 tys. opisów rocznie. Do tego trzeba doliczyć opisy z książek. Oczywiście część danych można pozyskać od wydawców lub samych autorów, niemniej pozostaje problem z przyjęciem i uporządkowaniem tak dużej liczby rekordów.

3. Wykonawca

Przygotowanie multidyscyplinarnej bazy wymaga odpowiedniego zaplecza w postaci dostępu do materiałów. Trudno sobie wyobrazić, aby mogła się tego podjąć jedna instytucja, zwłaszcza że nawet Biblioteka Narodowa nie ma w swoich zbiorach wszystkich wydawanych w Polsce prac naukowych.

4. Problemy z opisem bibliograficznym

Dużym i poważnym problemem jest kwestia ujednoczenia wprowadzanych opisów oraz ich deduplikacja. W związku z tym, że w polskie czasopisma zazwyczaj stosują własny standard opisu danych, mamy do czynienia z bardzo wieloma wariantami tego samego opisu. Można przyjąć zasadę, że niczego nie poprawiamy i zdajemy się na sprawność wyszukiwarki, ale z doświadczeń osób tworzących podobne bazy wynika, że dość szybko dochodzi się do wniosku, że taka korekta jest konieczna. Na pewno dużym ułatwieniem jest skanowanie materiałów lub pozyskiwanie wersji elektronicznych, tak aby można było kopiować całe opisy lub ich fragmenty. Ale należy też pamiętać, że w naukach humanistycznych opisy bibliograficzne podawane są w przypisach dolnych, często wplecione w komentarz i trudne do „wyłuskania”. W przypadku takiego stylu cytowania pozostaje przede wszystkim praca manualna.

5. Program

Przygotowanie polskiego indeksu wymaga stworzenia odpowiedniego programu, udostępniającego nie tylko prostą wyszukiwarkę, lecz także generującego odpowiednie zestawienia i wskaźniki. I nie chodzi tu tylko o odpowiedź na pytanie:



„ile mam cytowań”, ale przede wszystkim o możliwości przygotowania raportów uwzględniających wiele danych i kryteriów.

Podsumowując, stworzenie podobnej bazy danych na pewno jest bardzo oczekiwane przez polskie środowisko naukowe, zwłaszcza reprezentujące nauki społeczne i humanistyczne. Jednakże jej stworzenie wymagałoby szerokich konsultacji na wielu płaszczyznach (naukowcy, bibliotekarze, informatycy, przedstawiciele ministerstwa). Należałoby także rozpatrzyć, czy warto robić bazę interdyscyplinarną, czy lepiej skupić się tylko na niektórych dziedzinach, tych najslabiej indeksowanych w bazach międzynarodowych.

Wnioski

Podstawowym celem działań każdego uczonego jest nie tylko przeprowadzenie badań, lecz także wprowadzenie ich wyników, koncepcji, rozwiązań metodycznych itp. do międzynarodowego obiegu informacji. Sukces w działalności badawczej polega na tym, że publikowane prace są czytane i twórczo wykorzystywane przez innych naukowców. Dopiero wówczas następuje prawdziwe włączenie do krwioobiegu nauki. Liczba publikacji naukowych na całym świecie wzrasta tak szybko, że nawet osoby reprezentujące bardzo wąską specjalność nie są w stanie być na bieżąco z całą literaturą na dany temat. Rolę swoistego łącznika między uczonym a wynikami właśnie pełnią bibliograficzne bazy danych. W ostatnich latach znaczenie to niepomniernie wzrosło. Oprócz swojej podstawowej funkcji — wyszukiwania literatury na temat — bazy stanowią doskonałą podstawę do przygotowania analiz bibliometrycznych. Mamy do czynienia z coraz bardziej efektywnym wykorzystaniem baz danych jako bezcennego źródła wiedzy dla decydentów o stanie i perspektywach rozwoju prowadzonych badań, zarówno w kontekście krajowym, jak i międzynarodowym. Wszelkie rejestry bibliograficzne wiążą się bowiem — z mniejszym czy większym — selekcjonowaniem prac naukowych oraz zapewniają ich dystrybucję. Dane te ujawniają zakres obecności w światowym obiegu informacji.

W Polsce tworzy się wiele wartościowych źródeł tego typu. Powstają one w oparciu o najlepsze wzorce i tworzone są przez fachowców. Wszystkie przedstawione powyżej argumenty przemawiają za wykorzystaniem tych istniejących źródeł, a także za tym, żeby te banki danych były wspierane, rozwijane i doceniane w polityce naukowej państwa. Doceniając ich potencjalną rolę w tego typu analizach, należy wręcz modelować bazy w taki sposób, aby można było uzyskać jak najlepsze rezultaty.

Bibliografia:

1. *BazEkon* [on-line]. Kraków: Biblioteka Główna Uniwersytetu Ekonomicznego, 2010 [Dostęp 28.04.2013]. Dostępny w World Wide Web: http://kangur.uek.krakow.pl/bazy_ae/bazekon/nowy/.



2. *BazTech* — *Baza danych o zawartości polskich czasopism technicznych* [on-line]. Konsorcjum BazTech [Dostęp 28.04.2013]. Dostępny w World Wide Web: <http://baztech.icm.edu.pl/>.
3. BIRSKA, R., BUSSE-TURCZYŃSKA, E. Analizy bibliometryczne dla środowiska naukowego generowane na podstawie bazy danych „Bibliografia publikacji pracowników Uniwersytetu Medycznego w Lublinie”. *Forum Bibliotek Medycznych* [on-line]. 2009, nr 2 [Dostęp 28.04.2013], s. 523–535. Dostępny w World Wide Web: http://www.bg.umed.lodz.pl/index.php?option=com_content&task=category§ionid=9&id=33&Itemid=59.
4. *CYTBIN* [on-line]. Katowice: Instytut Bibliotekoznawstwa i Informacji Naukowej UŚ, 2000–2013 [Dostęp 28.04.2013]. Dostępny w World Wide Web: <http://www1.bg.us.edu.pl/bazy/cytbin/>.
5. *Cytowania w BazEkon* [on-line]. 2012 [Dostęp 28.04.2013]. Dostępny w World Wide Web: http://kangur.uek.krakow.pl/bazy_ae/bazekon/cytowania/.
6. DRABEK, A., TOMASZCZYK, J. Czasopismo „Praktyka i Teoria Informacji Naukowej i Technicznej” w świetle danych bazy CYTBIN. W: PIETRUCH-REIZES, D. (red.) *Zarządzanie informacją w nauce*. Katowice: Wydaw. UŚ, 2008, s. 365–375.
7. DRABEK, A. *Bibliometryczna analiza czasopism naukowych w dziedzinie nauk społecznych*. Toruń: Dom Wydaw. Duet, 2010. ISBN 978-83-61185-78-9.
8. DRABEK, A. Wykorzystanie bibliometrii w polityce naukowej. *Biuletyn EBIB* [on-line]. 2012, nr 3(130) [Dostęp 28.04.2013]. Dostępny w World Wide Web: http://www.nowybib.info/images/stories/numery/130/130_drabek.pdf.
9. KOLASA, W. M. Historia mediów polskich w świetle wskaźników bibliometrycznych. *Zeszyty Prasoznawcze* 2011, s. 3–4, s. 8–27.
10. KOLASA, W. M. Retrospektywny indeks cytowań w humanistyce. Koncepcja, metoda, zastosowanie. *Przegląd Biblioteczny* 2011, z. 4, s. 466–486.
11. KOMPERDA, A., URBAŃCZYK, B. Analiza naukometryczna dorobku publikacyjnego uczelni — kryteria, statystyki, raporty. *PTINT Praktyka i Teoria Informacji Naukowej i Technicznej* 2010, nr 3, s. 46–54.
12. MAMORSKA, K., SIEK, B. O zależnościach między bibliometrią a bibliografią na przykładzie bibliografii GUMed. *Forum Bibliotek Medycznych* [on-line]. 2011, nr 1(7) [Dostęp 28.04.2013], s. 557–563. Dostępny w World Wide Web: <http://www.bg.umed.lodz.pl/images/stories/wydawnictwa/forum/ForumNr7.pdf>.
13. MARSHAKOVA-SHAIKEVICH, I. *Badania ilościowe nauki. Podejście bibliometryczne i webometryczne*. Poznań: Uniwersytet im. A. Mickiewicza. Wydział Pedagogiczno-Artystyczny, 2009. ISBN 978-83-927082-4-7.
14. NOWAK, P. *Bibliometria, webometria. Podstawy. Wybrane zastosowania*. Poznań: Wydaw. Nauk. UAM, 2006. ISBN 83-232-1701-7.
15. OLECHNICKA, A., PŁOSZAJ, G. *Polska nauka w sieci. Przestrzeń nauki i innowacyjności. Raport z Badań* [on-line]. Warszawa: 2008 [Dostęp 28.04.2013]. Dostępny w World Wide Web: http://www.espon.pl/files/10_2/2/Polska%20nauka%20w%20sieci.pdf. ISBN 978-83-928352-1-9.



16. RACKI, G. Geologia Polska w świetle Science Citation Index. *Przegląd Geologiczny* 1996, T. 44, nr 9, s. 928–930.
17. RACKI, G., DRABEK, A. Cytowania i wskaźnik Hirscha. Gdzie szukać, jak obliczać? *Forum Akademickie* [on-line]. 2013, nr 2 [Dostęp 28.04.2013]. Dostępny w World Wide Web: <http://forumakademickie.pl/fa/2013/02/cytowania-i-wskaznik-hirscha-gdzie-szukac-jak-obliczac/>.
18. STEFANIAK, B. i in. Bibliograficzna baza danych — CYTBIN. *PTINT Praktyka i Teoria Informacji Naukowej i Technicznej* 2005, nr 4, s. 3–9.
19. STEFANIAK, B. Niekonwencjonalne zastosowania bibliograficznych baz danych, cz. 1. *PTINT Praktyka i Teoria Informacji Naukowej i Technicznej* 1995, nr 1, s. 3–10.
20. STEFANIAK, B. Niekonwencjonalne zastosowania bibliograficznych baz danych, cz. 2. *PTINT Praktyka i Teoria Informacji Naukowej i Technicznej* 1995, nr 2, s. 3–8.
21. WAGA, M., DRABEK, A. Arton — baza cytowań polskiej literatury humanistycznej (stan prac nad bazą). *Zagadnienia Naukoznawstwa* 2002, T. 38, z. 151–152, s. 83–89.
22. WINCŁAWSKA, B. M. Indeks cytowań socjologii polskiej (założenia ideowe i omówienie pierwszych wyników). *Zagadnienia Naukoznawstwa* 1995, nr 3–4, s. 243–246.

Drabek, A. Do użytku bibliometrycznego... Niebibliograficzne wykorzystanie baz bibliograficznych. W: Bibliograficzne bazy danych i ich rola w rozwoju nauki. II Konferencja naukowa Konsorcjum BazTech, Poznań, 17-19 kwietnia 2013 [on-line]. Stowarzyszenie EBIB, 2013 [Dostęp: 30.08.2013]. Materiały konferencyjne EBIB, nr 24, Dostępny w World Wide Web: http://open.ebib.pl/ojs/index.php/Mat_konf/article/view/26/15. ISBN 978-83-63458-06-5.