

Piotr Siłka*

Przemysław Śleszyński

Barbara Jaworska

Instytut Geografii i Przestrzennego Zagospodarowania PAN, Warszawa

Cytowania członków komitetów naukowych Polskiej Akademii Nauk według Google Scholar

Abstrakt. W artykule przedstawiono kompleksową analizę cytowań polskich naukowców będących członkami komitetów naukowych Polskiej Akademii Nauk. W szczególności zajmowano się, w miarę możliwości, porównaniem różnych dziedzin nauki, ale też wykazaniem, że nie zawsze takie porównania są uprawnione. Na podstawie wyszukiwarki Google Scholar (Publish or Perish 3.0) według stanu na połowę 2013 roku zgromadzono informację o 1,1 mln cytowań dla 3033 osób, będących członkami 95 komitetów naukowych. Analizę przedstawiono w trzech aspektach: ogólnej populacji badaczy, poszczególnych komitetów naukowych oraz sześciu podstawowych dziedzin nauki (na podstawie podziału nauk OECD). Wyniki potwierdzają istnienie różnic pomiędzy dziedzinami nauki, wynikających zarówno ze specyfiki dziedzinowej, jak też różnej kultury i tradycji cytowań badaczy.

Słowa kluczowe: rozwój nauki, dziedziny nauki, cytowania autorów, bibliometria, Google Scholar, autorytet, organizacje naukowe

Citation Analysis Of Polish Academy Of Sciences Committee Members in Google Scholar

Abstract. The article presents a comprehensive analysis of citations of Polish scientists who are members of the scientific committees of the Polish Academy of Sciences. In particular, they dealt with comparing different fields of science, but also showing that it is not always such comparisons are legitimate. Based on Google Scholar (Publish or Perish 3.0) in 2013 information about the 1.1 million citations for 3033 people, 95 are members of the scientific committees were gathered. The analysis is presented in three aspects: the general population researchers, various scientific committees and the six main fields of science (on the basis of the OECD division of science). The results confirm the existence of differences between the fields of science, under both the specifics of domain, as well as different cultures and traditions citations researchers.

Keywords: the development of science, science, authors citations, bibliometric, Google Scholar, authority, scientist organizations

1. Wprowadzenie

Spektakularny wzrost zainteresowania metodyką i wynikami badań bibliometrycznych i naukometrycznych w ostatniej dekadzie wynika z trzech głównych przyczyn. Po pierwsze, nastąpił imponujący przyrost liczby osób zajmujących

* Adres do korespondencji: Instytut Geografii i Przestrzennego Zagospodarowania PAN, ul. Twarda 51/55, 00-818 Warszawa, e-mail: psilka@twarda.pan.pl; psleszyn@twarda.pan.pl

się badaniami naukowymi. Współcześnie żyje i tworzy prawdopodobnie więcej naukowców, niż było ich dotychczas kiedykolwiek w dziejach. Liczbę osób na świecie, zatrudnionych w sektorze badań i rozwoju, szacuje się na nieco ponad 10 mln (UNESCO 2014)¹. Większość z nich, samodzielnie lub zespołowo, publikuje wyniki swoich badań, co powoduje konieczność nie tylko uporządkowanego, systematycznego i racjonalnego upowszechniania, ale także znajdowania efektywnych sposobów na identyfikowanie prac najbardziej wartościowych i mających największy oddźwięk.

Drugą przyczyną rozwoju badań bibliometrycznych i naukometrycznych są zmiany w obrębie funkcjonowania i zarządzania nauki, w tym jej zasilania finansowego. Część badaczy zwraca uwagę, że współcześnie mamy do czynienia z coraz częstszym przyjmowaniem kryteriów efektywnościowych w ocenie nauki czyli poglądem, że najbardziej wartościowe są te publikacje, które mają najwięcej cytowań (Paasi 2005; Aalbers i Rossi 2009). Autorzy świadomi, że od liczby cytowań ich publikacji, zależy szansa na otrzymanie środków na kolejne badania, coraz mocniej dbają o ten aspekt działalności naukowej.

Po trzecie, następuje szybki rozwój systemów informatycznych. Umożliwia to z jednej strony gromadzenie dużej liczby danych dotyczących wykorzystania prac naukowych, a także prowadzenie różnego rodzaju analiz wyszukiwawczych, klasyfikacyjnych i porównawczych, dotychczas niemożliwych lub bardzo czasochłonnych. Trzeba tu podkreślić, że kompleksowe analizy tego typu w Polsce do niedawna były utrudnione, gdyż wiele tytułów (artykułów jak i monografii) było wydawanych jedynie w tradycyjnej, papierowej formie. Bardzo widoczna zmiana nastąpiła szczególnie w ostatniej dekadzie w wyniku m.in. zmian kryteriów oceny parametrycznej jednostek naukowych i wymaganej sprawozdawczości w kwestii cytowań.

Przy tak zarysowanych uwarunkowaniach, głównym celem artykułu jest próba ustalenia siły oddziaływania naukowców zgromadzonych w komitetach naukowych Polskiej Akademii Nauk oraz występujących w tym zakresie prawidłowości. Przyjmuje się założenie, że cytowania mogą być, pod pewnymi warunkami i ograniczeniami (Garfield 1965), miernikiem oddziaływania dorobku naukowego (idei, koncepcji, wyników wyrażanych przez publikacje) na ustalenia naukowe innych badaczy i tym samym ewolucję badań.

Prezentowana analiza jest pierwszą tak obszerną i kompleksową oceną cytowań polskich naukowców. Obejmuje 1,1 mln cytowań dla 3033 osób we wszystkich dziedzinach nauki uprawianych w Polsce.

¹ http://data.uis.unesco.org/index.aspx?datasetcode=scn_ds, ostatni dostęp 12 października 2016 r.

2. Rola Komitetów Naukowych PAN w sektorze nauki

Współczesna nauka polska pod względem mierzalnych wyników według różnych ocen plasuje się w drugiej lub trzeciej dziesiątce krajów świata (OECD 2012). Poza pewnymi wyjątkami, lepiej rozwinięte są dyscypliny ścisłe i przyrodnicze, niż humanistyczne i społeczne. Wynika to nie tylko z tradycji naukowych, ale także ze swoistego ‘spadku’ po okresie przed 1989 r., w tym szczególnie świadomej polityki władz komunistycznych w zakresie sterowania naukami społecznymi. Było to charakterystyczne w zasadzie dla wszystkich państw bloku komunistycznego, gdzie podejmowano lokalne tematy, będące na peryferiach globalnej nauki.

Analizując poszczególne dziedziny, nauka polska przed 1989 r. zajmowała lepszą pozycję niż obecnie. Szczególne sukcesy dotyczyły m.in. matematyki, astronomii, fizyki, chemii, archeologii i geografii, a także niektórych specjalności medycyny. W opublikowanym po raz pierwszy w 1981 r. rankingu *Institute for Scientific Information*, Polska zajmowała 15 miejsce w świecie. Od tamtego czasu Polska została wyprzedzona przez Koreę Południową, Chiny, Brazylię, Turcję i Tajwan (Wróblewski 2012).

Po 1989 roku nauka w Polsce przeżywa poważny kryzys, związany z niedoinwestowaniem tego sektora oraz zmianami pokoleniowymi. Źródła tego kryzysu tkwią jeszcze w latach 80. ubiegłego wieku, kiedy Polska przechodziła poważny kryzys gospodarczy związany z upadkiem gospodarki centralnie planowanej oraz kiedy to obserwowano nasilenie wyjazdów zagranicznych najlepszych polskich naukowców. Współcześnie na naukę przeznaczają się około 0,8% PKB, z czego ze środków budżetowych zaledwie około 0,3% PKB, co plasuje Polskę wśród krajów OECD i Unii Europejskiej na bardzo odległej pozycji (OECD 2012). Tymczasem jeszcze w 1980 r. na naukę w Polsce przeznaczano się około 1,6% PKB, głównie ze środków budżetowych (Rejn 2003). Należy jednak zauważyć, że od wejścia Polski do Unii Europejskiej w 2004 r., Polscy naukowcy na niespotykaną dotychczas skalę mogą korzystać z różnych zewnętrznych źródeł finansowania badań naukowych (np. w Programach Ramowych), które przyczyniają się do wzmocnienia naukowej współpracy międzynarodowej. Mimo, że wciąż jest to wciąż słabe zaangażowanie – zwłaszcza pod względem koordynacji projektów, w 7 PR było to 1600 zespołów badawczych z Polski² i nie można lekceważyć tego wpływu na pozostałych pracowników naukowych, jak też na całą polską naukę.

W Polsce, według danych GUS, w końcu 2014 r. w działalności badawczo-rozwojowej było zatrudnionych 153,5 tys. osób (w ekwiwalentach pełnego czasu pracy wyniosło to 104,4 tys. EPC). W 2013 r. w krajach Unii Europejskiej liczba personelu w działalności badawczej i rozwojowej, wyrażonego w EPC, przypadająca na 1000 pracujących ogółem wyniosła 12,6, w tym pracowników

² Dane pochodzą ze strony <http://7pr.kpk.gov.pl>

naukowo-badawczych – 8,0. Najwyższym natężeniem personelu B+R (na 1000 pracujących ogółem) wśród analizowanych krajów charakteryzowała się Dania (21,7), Finlandia (21,6) oraz Luksemburg (20,8), natomiast najniższym – Rumunia (3,8), Cypr (3,4) oraz Bułgaria, Łotwa i Polska (6,0). Odsetek pracowników naukowo-badawczych w ogólnej populacji personelu B+R utrzymuje się w Polsce na poziomie wyższym, niż w całej Unii Europejskiej.

Równocześnie według danych wstępnych, w 2014 r. wśród krajów Unii Europejskiej wysoki udział badaczy (powyżej 75%) w personelu B+R odnotowano poza Polską – w Szwecji (79,8%), na Słowacji (77,1%), na Litwie (76,6%) i w Czechach (76,3%). Znacznie niższy poziom tego odsetka charakteryzuje zatrudnienie w działalności B+R w Luksemburgu i Hiszpanii (po 49,6%), na Cyprze (54,4%), w Rumunii i Słowenii (po 57,7%). W analizach porównawczych według bazy *Scopus*³, w roku 2015 Polska znajdowała się na 19 miejscu pod względem bezwzględnej liczby publikacji. Polscy naukowcy publikują dwukrotnie mniej niż autorzy z Hiszpanii, czy Włoch, a czterokrotnie mniej niż autorzy z Niemiec lub Wielkiej Brytanii. Liczba publikacji autorów amerykańskich zaś jest piętnastokrotnie większa. Choć trzeba zaznaczyć, że liczba publikacji naukowych rośnie w Polsce bardziej dynamicznie, niż np. w Czechach, czy nawet w Niemczech (Pilec 2015).

W okresie powojennym ukształtowały się, funkcjonujące do czasów współczesnych, trzy główne 'piony' naukowe: szkoły wyższe, Polska Akademia Nauk oraz tzw. jednostki badawczo-rozwojowe. Największą instytucją skupiającą naukowców w Polsce jest Polska Akademia Nauk. Instytucja ta powstała w 1951 r. i realizuje działania służące rozwojowi promocji, integracji oraz upowszechniania nauki. Jedną z form realizacji tych zadań są komitety naukowe Polskiej Akademii Nauk, będące samorządną ogólnopolską reprezentacją dyscyplin naukowych lub ich grup, służącą integrowaniu uczonych i badań. Zgodnie z odpowiednimi wytycznymi, w skład komitetów naukowych wchodzi krajowi wybitni specjaliści odpowiednich specjalności. Część osób posiadających tytuł profesora – najwyższy w Polsce, przewidziany w systemie nauki stopień awansu zawodowego, legitymujący się uznanym dorobkiem i osiągnięciami, są członkami Polskiej Akademii Nauk. Pozostałe osoby w komitetach stanowią inni wybitni pracownicy naukowcy reprezentujący szkoły wyższe, placówki Polskiej Akademii Nauk i instytuty zaplecza naukowo-badawczego resortów, a także przedstawiciele instytucji i organizacji gospodarczych i społecznych. Członkowie PAN są wybierani przez środowisko naukowe na podstawie powszechnych, tajnych wyborów. Do zadań komitetów należy przede wszystkim oddziaływanie na rozwój danej dyscypliny naukowej, integrowanie środowisk naukowych oraz wymiana wyników badań. Istotna jest też rola opiniotwórcza oraz doradzanie w zakresie kluczowych problemów naukowych i społeczno-gospodarczych w rozwoju kraju.

³ Dane pochodzą ze strony <http://www.scimagojr.com/countryrank.php?year=2015>

Z powyższych powodów, komitety naukowe PAN stanowią stosunkowo najbardziej reprezentatywne grono najlepszych specjalistów w danej dyscyplinie naukowej. Mogą być tym samym dobrym źródłem informacyjnym dla analiz oraz w pewnym zakresie prawidłowości i stopnia oceny rozwoju dziedzin naukowych w Polsce.

3. Analizy bibliometryczne dziedzin i dyscyplin⁴ nauki

Dotychczasowe badania bibliometryczne i naukometryczne wskazują na różnice w trendach dotyczących publikacji, a co za tym idzie, także cytowań pomiędzy poszczególnymi dyscyplinami. Najczęściej zagadnienie te jest omawiane w kontekście terminu ‘kultury cytowań’, stosowane są także określenia: wzorców cytowań (ang. *citation patterns*) i wzorców publikacji (ang. *publication patterns*). Kultura cytowań opisuje cechy, strukturę i reguły cytowań charakterystyczne dla danego obszaru nauki, kraju, czy czasopisma (Wouters 1999). W jej ramach najczęściej analizowaną są takie aspekty jak:

- specyfika liczby publikacji w danej dziedzinie, a przez co dalej idzie, także samej liczby cytowań (Moed i in. 1985). Na przykład prace z zakresu matematyki są długo przygotowywane, stąd mniejsza liczba publikacji, a także mniejsza liczba cytowań;
- liczba autorów prac, czyli zagadnienie, czy częściej w danej dziedzinie są publikowane prace jednoautorskie, jak np. ma to miejsce w humanistce (Hellqvist 2009), czy też wieloautorskie (nauki eksperymentalne i wymagające dużych nakładów, np. fizyka, astronomia, teledetekcja satelitarna);
- kanały komunikacji, które są najczęściej wykorzystywane przez autorów do publikacji badań, tj. czy są to raczej artykuły w czasopismach, czy monografie (Hammarfelt 2012; Konieczna 2002); przykładowo w humanistyce dominującym kanałem komunikacji są monografie i rozdziały w monografiach (Cullars 1998; Drabek i in. 2015);
- dynamiki i wieku cytowań, tj. czy struktura cytowań składa się w większości z nowszych publikacji, czy też starszych (Barnett i in. 1989). Jako pierwszy zauważył to D.J. de Solla Price (1970), który badał udział cytowań, ze względu na wiek, m.in. w poszczególnych dyscyplinach naukowych. Zgodnie z jego badaniem nauki „twarde” charakteryzują się większym udziałem cytowań publikacji z 5 minionych lat (np. fizyka – około 60%), niż nauki „miękkie” (np. humanistyka – zaledwie około 10%). Dlatego tak popularny współczynnik oddziaływania (*impact factor*; Garfield 1979, 1994) liczony

⁴ Podział na dyscypliny i dziedziny wykorzystany w tym artykule został ustanowiony w 2004 r. przez Organizację Współpracy Gospodarczej i Rozwoju (OECD). Zgodnie z nim podzielono naukę na sześć dziedzin nauki oraz jedną dziedzinę sztuki, natomiast w ramach dziedzin wyszczególniono w sumie ponad 40 różnych dyscyplin naukowych.

dla okresu 2 lat jest odpowiedni dla nauk biologicznych, ale już nie dla np. matematyki (Laloë i in. 2009);

- język publikacji oraz kraj pochodzenia autorów (Winclawska 1996).

Z powyższych względów należy mieć świadomość, że rozbieżności wskaźników bibliometrycznych w poszczególnych dziedzinach nauki wynikają z istnienia różnej kultury cytowań, co chociażby przekłada się na osiągnięcia w popularnym wskaźniku Hirscha (Batista i in. 2008). Z tego względu należy bardzo ostrożnie porównywać aktywność i siłę oddziaływaniu naukowców z różnych dyscyplin i dziedzin nauki (Althouse i in. 2009; Petersen i in. 2010; Chang 2013).

Zagadnienia dotyczące bibliometrii w Polsce cieszą się popularnością od lat 1990. Popularne są analizy w zakresie publikacji z konkretnej dziedziny dotyczące poszczególnych autorów, jak też zbiorowej oceny czasopism naukowych oraz ośrodków naukowych. Jedną z pierwszych takich prac był artykuł A. Królikowskiej (1971) dotyczący prac z zakresu mechaniki. W kolejnych latach poszczególne dziedziny nauki były analizowane często przez A. Drabek, B. Stefaniak, M. Kolasa, H. Szarskiego, A. Królikowską, A. Pilca, G. Rackiego, P. Nowaka oraz A. Olechnicką i A. Płoszaja. Coraz częściej także poszczególne ośrodki naukowe zaczynają wydawać specjalne raporty dotyczące cytawalności swoich pracowników, np. Politechnika Wroclawska (m.in. Komperda i Urbańczyk 2011). Podejmowane także analizy porównawcze pomiędzy dziedzinami nauki (Błocki i Życzkowski 2012, Sadowski 1995; Życzkowski 2011; Kozłowski 1994, Wróblewski 2001, 2013; Kierzek 2009, 2010) oraz analizy powiązań między dyscyplinami, np. relacje geografii i z innymi dyscyplinami (Stachowiak i Bajerski 2012).

W powyższym kontekście interesujące badania w tym zakresie przeprowadził m.in. K. Wróblewski (2005), który analizował udział polskich autorów w poszczególnych dyscyplinach w bazie edycji bazy *Science Citation Index* wydawanej jako *National Science Indicators*⁵. Z kolei A. Olechnicka i A. Płoszaj (2008) badali aktywność naukową w podregionach Polski, m.in. poprzez publikacje naukowe z bazy *Web of Science* z lat 2001–2006. K. Klincewicz (2008) w ramach badań nad innowacyjnością analizuje publikacje z obszaru nauk ścisłych i technicznych w bacie INSPEC (obejmującej artykuły w specjalistycznych czasopismach i wystąpienia na międzynarodowych konferencjach). Pozostali autorzy w zakresie aktywności publikacyjnej naukowców i cytowań wskazują m.in. na niedoskonałość istniejących baz danych, brak uwzględniania specyfiki cytawalności wśród dyscyplin, jak też zbyt niską rangę czasopism nieanglojęzycznych (Pindłowa 1989; Marszakowa-Szajkiewicz 1996; Nowak 2006; Stefaniak 2008).

Na tle powyższej analizy literatury, celem artykułu jest sprawdzenie, na ile zaobserwowane w piśmiennictwie różnice pomiędzy wartościami wskaźników bibliometrycznych w różnych dziedzinach nauki, wynikające z odmiennych kultur

⁵ Jest to poprzednia nazwa bazy Thomson Reuters.

cytowań, występują również w Polsce w grupie członków Komitetów Naukowych Polskiej Akademii Nauk. Argumentem za przeprowadzeniem tak ukierunkowanej analizy jest przygotowanie obszernej bazy cytowań, dotyczącej wszystkich komitetów PAN. Dodatkowym argumentem za badaniem był fakt, że do tej pory gremia te nie były przedmiotem dokładniejszych analiz w tak szerokiej skali (temat ten jedynie pokrótce poruszył jeden ze współautorów w przeglądowym artykule w „Forum Akademickim”; Śleszyński 2014). Badania poświęcone naukowcom z konkretnego kraju przedstawiały m.in. sytuacje we Włoszech (Abromo i in. 2015), Słowenii (Perc 2010), Chorwacji (Klaic 1995) oraz Brazylii (del Almeida i in. 2013).

4. Źródła danych i metody badań

Zasadniczo istnieją cztery przedmiotowe sposoby przeprowadzenia badań cytowalności, opierające się na założeniu, że daną pracę naukową publikuje (1) autor reprezentujący (2) instytucję, w której jest afiliowany, poprzez określony (3) kanał komunikacji naukowej (np. książka, rozdział, czasopismo, strona internetowa) wydany przez konkretne (4) wydawnictwo. Obecnie istnieje wiele baz bibliograficznych, jednak najczęściej do analiz bibliometrycznych stosuje się następujące źródła:

- *Web of Science* (m.in. Batista i in. 2006; Petersen i in. 2010; Franceschet 2010; Ho 2014; Mikki 2010),
- *Scopus* (m.in. Bosman i in. 2006; Etxebarría i Gomez-Uranga 2010; Fingerman 2005; Moya-Anegón i in. 2008),
- *Google Scholar* (m.in. Aguillo 2012; Orduna-Malea i Lopez-Cozar 2014; Kousha i Thelwall 2008; Harzing 2013; Smith 2008).

Badania wskazują, że w zależności od dziedzin, a także dyscyplin naukowych stopień pokrycia w bazach jest różny (Bartneck i Kokkelmans 2011; Mingers i Lipitakis 2010; Etxebarría i Gomez-Uranga 2010). A.W. Harzing (2013) oraz J. Mingers i E. Lipitakis (2010) wskazują na Goole Scholar jako najlepsze źródło dla badań porównawczych między dziedzinami. Główną przyczyną jest w tym przypadku większa liczba publikacji zwartych, przez co także zakres źródeł jest bardziej reprezentatywny, a liczba cytowań jest większa (m.in. Bar-Ilan 2010; de Winter i in. 2014; Franceschet 2010; Mikki 2010; Must 2006). Istotnym powodem jest też indeksowanie publikacji w językach innych, niż angielski (Amara i Landry 2012). Dzięki temu Google Scholar prezentuje większą reprezentatywność dla obszaru Polski, niż dwie wyżej wspomniane bazy. Te czynniki, a także stosunkowo duża łatwość w pozyskaniu materiału źródłowego (korzystano ze wspomagającego oprogramowania Publish or Perish; por. Harzing 2011) przyczyniły się do wykorzystania Google Scholar jako źródła danych w niniejszym artykule.

Google Scholar nie jest typową bazą bibliograficzną. Jest to wyszukiwarka tekstów naukowych występujących w szeroko rozumianym Internecie. Serwis ten powstał w 2004 r., a już dwa lata później doczekał się swojej pierwszej analizy naukowej (Neuhaus i in. 2006), w której na podstawie wybranych 50 pozycji porównano ją z innymi popularnymi bazami. Google Scholar indeksuje dane z istniejących baz bibliograficznych, w tym z rozproszonych repozytoriów instytucjonalnych. W roku 2011 udostępnione także usługę *Google Scholar Citations*, umożliwiającą tworzenie profili dla poszczególnych autorów, gdzie m.in. są gromadzone cytowania w ujęciu rocznym. Rok później zaś uruchomiono usługę *Google Scholar Metrics*, która z kolei pozwala na przedstawianie podstawowych danych bibliometrycznych dla czasopism, których artykuły są indeksowane przez wyszukiwarkę Google Scholar⁶. Dotyczy to zarówno stron regularnych czasopism i serii wydawniczych, jak też tzw. „szarej literatury”, w postaci różnego rodzaju zamieszczanych raportów, dokumentów w formacie pdf itp. Jedną z głównych krytyk dotyczy właśnie braku informacji, jakie źródła podczas indeksowania są brane pod uwagę i jaka jest ich selekcja (Jacsó 2005; Kousha i Thelwall 2008). Zatem Google Scholar nie wyszukuje części istniejących cytowań, m.in. ze względu na błędy i niekonsekwencje zapisu metatagów⁷ a z drugiej strony indeksuje prace, które nie zawsze przeszły odpowiednią weryfikację jakościową, np. w postaci recenzji. Należy też mieć świadomość, że Google Scholar jest podatny na przekłamania, gdyż nie weryfikuje wszystkich źródeł, z których korzysta (López-Cózar i in. 2012).

W przypadku analiz ilościowych należy też zauważyć, że Google Scholar nie pomija autocytowań i posiada pewien odsetek błędów obliczeniowych, wynikających z niedokładnego zapisu autorów i wydawnictw, jak też dostępności w wielu źródłach różnych wersji tej samej pracy, co powoduje czasami pojawianie się różnych rekordów dotyczących tej samej pracy. Choć należy pamiętać, że każdy autor ma możliwość korygowania takich błędów w momencie, gdy założy i będzie administrował swoje konto w narzędziu *Google Scholar Citations*, które wpływa na zasoby Google Scholar. Powoduje to utrudnienia w przypadku niektórych analiz, ale generalnie baza ta jest dobrą podstawą do analizy statystycznej cytowań.

Należy mieć także na uwadze fakt, iż zasięg Google Scholar cały czas się powiększa. Jest to związane z tym, że nie tylko poszczególne repozytoria, z których korzysta, są na bieżąco aktualizowane, ale także w wyniku włączania starszych materiałów (np. wskutek zmiany polityki wydawnictwa, które postanawia publikować książki w Internecie, np. poprzez Google Books). Powoduje to niekiedy skokowe wzrosty cytowań dla poszczególnych autorów w czasie. W Polsce zalety

⁶ Uwzględnione zostały artykuły z lat 2007–2011. Indeksowane są tylko czasopisma, które w tym czasie opublikowały co najmniej 100 artykułów i zanotowały minimum jedno cytowanie.

⁷ Metatag (znacznik meta) oznacza zbiór znaczników w sekcji nagłówkowej danego dokumentu, zamieszczonego na stronie internetowej, używanego do opisu jego zawartości merytorycznej, pochodzenia, formatu, itd.

Google Scholar na tle innych źródeł informacyjnych o cytowaniach przedstawili ostatnio G. Racki i A. Drabek (2013), stwierdzając m.in., że pomimo różnych słabości, wyszukiwarka ta jest najbardziej reprezentatywną dla nauk społeczno-humanistycznych.

W analizach wykorzystano kwerendę wykonaną w drugiej połowie 2013 r., w ramach której zgromadzono dane dla 3033 osób⁸. Baza ta objęła 1,1 mln rekordów. Długi proces weryfikacyjny poszczególnych rekordów spowodował znacznie opóźnienie w procesie przygotowania niniejszego artykułu.

Analizą objęto składy komitetów naukowych PAN, ustalonych na kadencję 2011–2014. Lista stałych komitetów naukowych obejmuje 95 jednostek. Oprócz tego jest także 6 komitetów problemowych i 3 rady o statusie komitetu problemowego działające przy Prezydium PAN oraz 5 komitetów problemowych przy Wydziale I (Nauk Ekonomicznych i Społecznych PAN) i 3 przy Wydziale IV (Nauk Technicznych PAN). Ze względu na interdyscyplinarność komitetów problemowych (trudność w rozdzieleniu pomiędzy dziedziny nauki) kwerenda dotyczyła tylko i wyłącznie komitetów naukowych.

Dla każdej osoby pozyskano informację z Google Scholar (za pomocą aplikacji Publish or Perish 3.0; Harzing 2007) o liczbie publikacji wraz z datą jej opublikowania oraz liczbą cytowań. Należy także zaznaczyć, iż zastosowano metodę zliczania całkowitego dla podstawowej jednostki analizy, którą były publikacje. Metoda zliczania całkowitego polega na tym, że poszczególne dane artykułu traktowane są jako 1. W tym przypadku, jeśli dana praca była napisana przez więcej niż jednego autora, ta publikacja i jej liczba cytowań nie były dzielone przez liczbę jej autorów (np. dla każdego z trzech współautorów dana publikacja liczyła się jako jedna, a także każdy otrzymywał pełną liczbę cytowań tej publikacji)⁹. Taki sposób tworzenia kwerendy został narzucony przez techniczną niemożność użycia metody zliczania udziałowego (format danych nie pozwalał na takie obliczenia). Autorzy dołożyli wszelkich starań by wyeliminować błędy związane z pisownią nazwiska, a także sumowania prac różnych autorów mających to samo imię i nazwisko¹⁰.

W analizach oprócz standardowych wskaźników względnych wykorzystano indeks Hirscha (2005). Indeks ten informuje, że naukowiec o wskaźniku H posiada H publikacji, z których każda ma H lub więcej cytowań. W ciągu ostatniej dekady stał się jednym z najpopularniejszych wskaźników oceniających dorobek naukowy, który obecnie istnieje jako jedna z podstawowych funkcjonalności m.in. w Google

⁸ Liczba rekordów w bazie wyniosła 3328, gdyż jedna osoba mogła być członkiem 2 lub trzech komitetów naukowych. W analizach dotyczących całej grupy autorów, zdublowane rekordy nie były brane pod uwagę.

⁹ Metoda zliczania udziałowego nie zastosowana w niniejszym artykule polega na tym, że poszczególne dane artykułu dla wszystkich autorów sumują się do 1. W tym przypadku polegałoby na dzieleniu pomiędzy autorów liczby samej publikacji jak i jej cytowań (np. każdy z trzech współautorów w bazie otrzymał by 1/3 publikacji i także 1/3 cytowań dla tej publikacji).

¹⁰ W analizowanej grupie znalazły się trzy takie przypadki, gdzie dwie różne osoby miały takie samo imię i nazwisko. Byli to: Janusz Borkowski, Jerzy Brzeziński, Zbigniew Dąbrowski. W każdym z 6 przypadków analizowane wszystkie publikacje tychże autorów, by rozdzielić je pomiędzy dwie osoby.

Scholar, Scopus oraz Web of Science. Warto jednak dodać, że indeks Hirscha został zaprojektowany do wstępnego porównania dorobku naukowców pracujących w jednej dziedzinie i starającego się o to samo stanowisko, o czym należy pamiętać przy porównaniach badaczy z różnych dziedzin czy dyscyplin. Wielu autorów wskazuje na zbyt dużą zależność tego wskaźnika od czasu trwania naukowej działalności badacza (Egghe 2006; Egghe i Rousseau 2008; Kelly i Jennions 2006; Liang 2006). Z tego względu sam Jorge Hirsch zaproponował podział wskaźnika H przez liczbę lat od pierwszej publikacji i taki wskaźnik określił jako m -quotient. Skutkiem kolejnych dyskusji są propozycje nowych wskaźników (Liu i in. 2013). W niniejszej publikacji wykorzystano także wskaźniki a -index (Jin 2006), m -index (Bornmann i in. 2008), które odnoszą się do tzw. rdzenia Hirscha, czyli publikacji wliczanych do wskaźnika H . M -index oznacza medianę cytowań wśród tych publikacji, natomiast a -index – wartość średniej arytmetycznej. Pozwalają one także obliczyć odpowiednio wskaźniki R -index (Jin i in. 2007) i q^2 -index, które łączą informacją o charakterze ilościowym (indeks H) z informacją o charakterze bardziej jakościowym, czyli jak dużo cytowań uzyskały najlepsze prace danego autora. Dzięki temu autor mający taki sam indeks H , jak jego kolega, ale jednocześnie więcej cytowań najpopularniejszych prac, będzie miał wyższy wskaźnik R -index i q^2 -index. R -index jest kwadratem sumy cytowań w rdzeniu Hirscha, a q^2 -index (Cabrerizo i in. 2010) jest średnią arytmetyczną iloczynu wskaźników h i m .

5. Analiza cytowań autorów

Liczba publikacji autorów będących członkami komitetów naukowych PAN do momentu przeprowadzonej kwerendy wyniosła blisko 119 tysięcy zaindeksowanych prac, co daje ponad 39 prac na jedną osobę. Dominantą, czyli najczęstszą liczbą prac (Rycina 1) pojawiającą się w bazie jest 20 (6,3% wszystkich autorów). Ponad 75% autorów posiada od 1 do 50 prac. Pozostałych 681 autorów ma zaindeksowanych więcej niż 50 publikacji, na czele z Zygmuntem Baumanem z ponad 1500 opracowaniami, który ma ich prawie 3 razy więcej niż kolejny autor (Tabela 1). Tak wysoki wynik jest też spowodowany licznymi tłumaczeniami prac autora, które Google Scholar traktuje jako osobne prace.

Analizując daty publikowania prac, badana grupa największą indeksację prac wykazuje w latach 2006–2011, kiedy to ukazuje się corocznie ponad 6 tysięcy publikacji. W ostatnich dwóch latach przed przeprowadzeniem kwerendy liczba artykułów w bazie drastycznie spada – do 600 w 2013 r., co wskazuje, że może występować nawet kilkunastomiesięczny okres opóźnienia (szczególnie w przypadku monografii), zanim publikacja zostanie udostępniona w Internecie przez podmioty mające prawa do nich, a następnie zindeksowane przez Google Scholar.

Rycina 1. Udział autorów mających zaindeksowaną liczbę prac

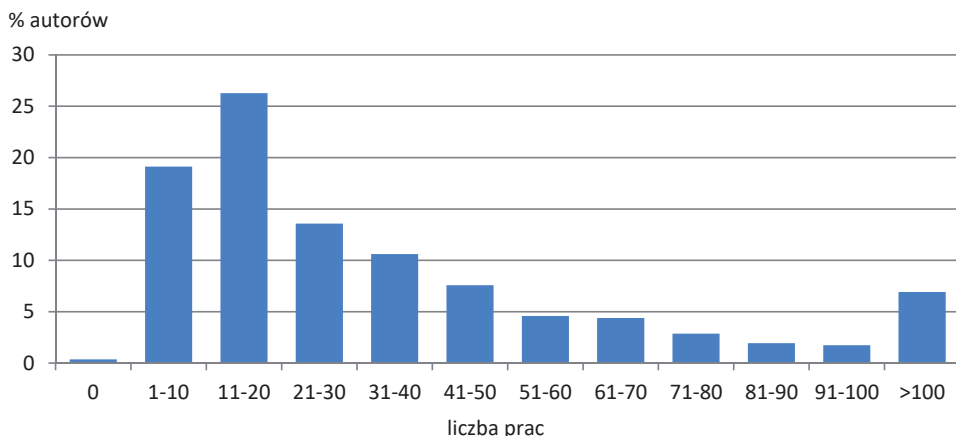


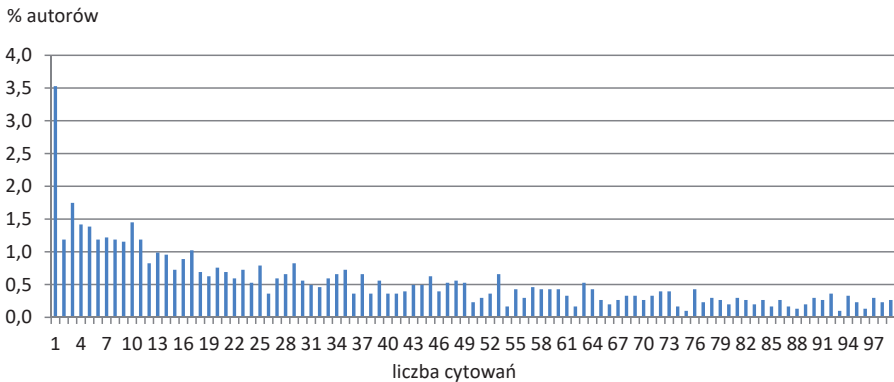
Tabela 1. Wybrane dane dotyczące autorów

Osoby o największej liczbie publikacji		Osoby o największej liczbie cytowań		Osoby o największym indeksie H	
Zygmunt Bauman	1 552	Zygmunt Bauman	78 378	Zygmunt Bauman	100
Janusz Kacprzyk	535	Andrzej Skowron	12 822	Mariusz Ratajczak	55
Ulrich Werner	517	Krzysztof Narkiewicz	10 038	Roman Słowiński	52
Bożena Kamińska	454	Mariusz Ratajczak	9 915	Marek Michałak	52
Andrzej Skowron	437	Roman Słowiński	9 742	Jan Potempa	51
Jan Hertrich Woleński	425	Marek Michałak	9 649	Wacław Szybalski	47
Mieczysława Czerwionka-Szaflarska	395	Wacław Szybalski	9 392	Andrzej Wróblewski	46
Stanisław Konturek	389	Janusz Kacprzyk	9 161	Bogumił Jeziorski	45
Jacek Namiesnik	376	Jan Potempa	8 527	Andrzej Skowron	44
Roman Słowiński	368	Józef Pacyna	7 301	Stanisław Konturek	44

Zindeksowane prace autorstwa członków komitetów były łącznie cytowane ponad 1 139 tys. razy, co daje przeciętnie 375,8 cytowania na 1 autora, a 9,6 na 1 pracę. 107 autorów nie zanotowało cytowania swoich prac (3,4%). Co najmniej 10 cytowań uzyskało ponad 84% autorów, ponad 50 cytowań – nieco poniżej 60%, zaś powyżej 100 – niecałe 50% (Rycina 2). Sam Z. Bauman z prawie 80 tysiącami cytowaniami swoich prac (najwięcej, gdyż 4185 cytowań zanotowało hiszpańskojęzyczne wydanie *Liquid Modernity*) odpowiada za 6,8% wszystkich zgromadzonych cytowań w bazie. Kolejna osoba (Andrzej Skowron) ma ich już 6 razy mniej.

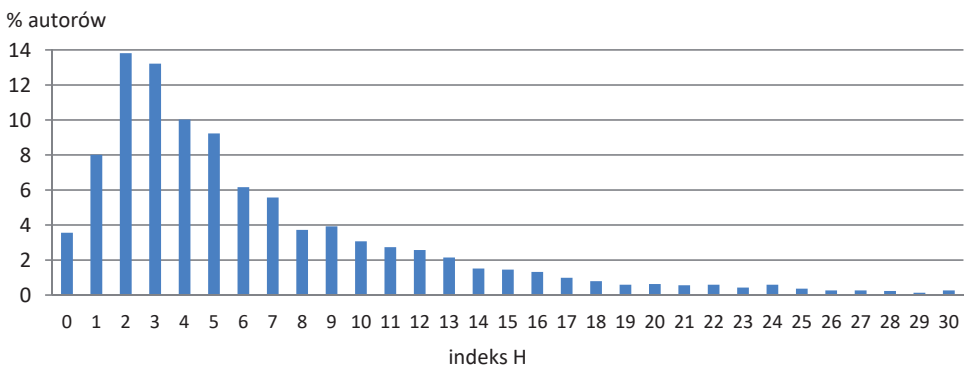
Rozpatrując globalną liczbę cytowań wykorzystanego zbioru znajdujemy potwierdzenie prawa Lotki (Lotka 1926). Rozkład danych zbliża się wręcz do modelowego rozkładu według Vilfreda Pareto, ponieważ 20% członków komitetów odpowiada dokładnie za 81,1% wszystkich cytowań. W grupie osób (67), których wszystkie prace były cytowane chociaż jeden raz, liczba opublikowanych prac waha się od 1 do 23.

Rycina 2. Procent autorów posiadających daną liczbę cytowań (do 100 cytowań)



Ponad 100 autorów (3,2% członków komitetów) posiada zerowy indeks H (brak publikacji z choć jednym cytowaniem). Dotyczy to autorów mających niezbyt wiele zaindeksowanych publikacji (maksymalnie 22). Największy odsetek autorów w analizowanej bazie ma indeks H równy 2 lub 3 (w sumie jest to 26,6% autorów; Rycina 3). Liczby autorów z kolejnymi indeksami Hirscha od 3 maleją zgodnie z funkcją wykładniczą ($y = 14,49e^{-0,14x}$ $R^2 = 0,940$). Tylko 38 autorów ma indeks Hirscha większy niż 30, a tylko cztery osoby powyżej 50.

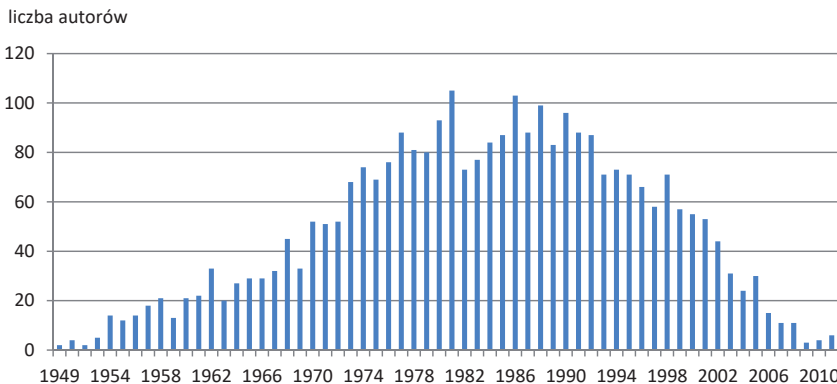
Rycina 3. Procent autorów posiadających dany indeks H (bez autorów z indeksem $h > 30$)



Przyglądając się samym publikacjom w bazie, ponad 40% z nich nie zanotowała ani jednego cytowania w Google Scholar. Co najmniej 10 razy było cytowanych 20,4% prac, ponad 50 cytowań zanotowało 3,6% publikacji, zaś powyżej 100 – 1,3%. Wyniki te są nieco wyższe od danych zaobserwowanych przez E. Garfielda (1989), który podawał, że ponad połowa prac nie jest w ogóle cytowana, więcej niż 50 cytowań uzyska około 1% publikacji, a ponad 100 – zaledwie 0,3%.

Jako data pierwszej publikacji naukowej indeksowanej przez Google Scholar najczęściej pojawiają się lata 1981 i 1986 (Rycina 4). Prawie 30% autorów swój pierwszy artykuł zindeksowany w Google Scholar ma w latach 1981–1990, nieco ponad 24% w latach 1971–1980, a ponad 23% – w latach 1991–2000. Data ta może posłużyć między innymi do obliczenia wskaźnika m („ m -quotient”; Hirsch 2005), który modyfikuje wskaźnik H w zależności od czasu trwania kariery naukowej. Na tej podstawie można określić tzw. „młode gwiazdy”. Wysoki wskaźnik (> 2), osiągnęły 3 osoby: Janusz Bujnicki (Komitet Biologii Ewolucyjnej i Teoretycznej PAN), który w ciągu 13 lat kariery naukowej osiągnął indeks $H = 41$ oraz Tomasz Guzik (Komitet Nauk Fizjologicznych PAN) i Mariusz Ratajczak (Komitet Nauk Fizjologicznych PAN). Wskaźnik pomiędzy 1 a 2 osiągnęło 67 osób, między 0,5–1 kolejne 351 osób, a pozostałych prawie 2,9 tys. autorów uzyskało wskaźnik niższy niż 0,5.

Rycina 4. Liczba autorów a rok rozpoczęcia ich kariery naukowej (daty pierwszej zaindeksowanej publikacji)



Rozkład cytowań nawiązuje do normalnego, z wyjątkiem lat 1982–1985. Ta charakterystyczna „wyrwa” może być dowodem na kryzys w nauce polskiej po stanie wojennym lat 1980., związanym zarówno z mniejszą aktywnością publikacyjną, jak też osłabieniem kadrowym jednostek badawczych.

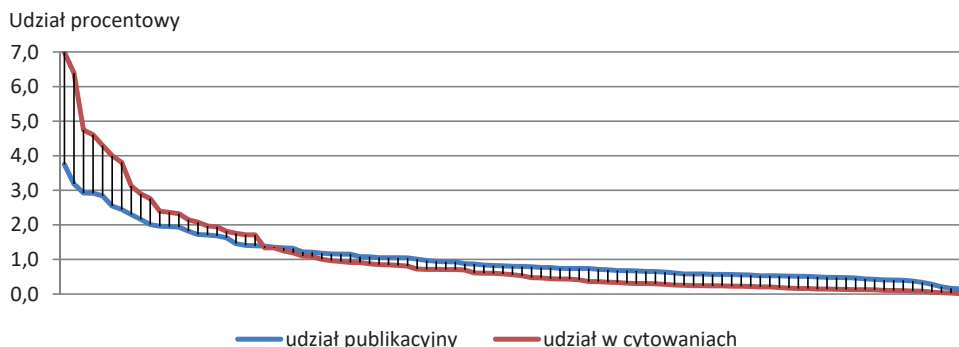
6. Analiza cytowań komitetów naukowych

Zdecydowana większość autorów była członkiem jednego komitetu naukowego. Natomiast 282 osoby było członkami dwóch komitetów, a 17 osób udzielało się w trzech podmiotach. Średnio w jednym Komitecie skład liczył 35 osób, z czego ich liczba waha się od 16 (Komitet Zdrowia Publicznego PAN) do 57 (Komitet Biochemii i Biofizyki PAN). Przeliczając ogólną liczbę publikacji, średnia liczba prac przypadająca na jeden komitet wynosi 1467.

W czołówce komitetów pod względem liczby publikacji są nauki ścisłe i medyczne. Na drugim krańcu znajdują się głównie komitety inżynieryjno-techniczne oraz nauki humanistyczne. Wśród poszczególnych komitetów udział w liczbie publikacji osoby, która najwięcej publikowała jest różny i waha się aż od 44,9% w Komitecie Nauk Filozoficznych (J. Woleński) do 6,7% w Komitecie Nauk Zootechnicznych, dając średnio 13,3% we wszystkich Komitetach.

Średnio jeden komitet „wypracował” 14,4 tysiąca cytowań. Ponad połowę wszystkich cytowań zostało wypracowanych przez 13 z nich, z czego udział cytowań członków Komitetu Socjologii PAN wyniósł prawie 7% w całej analizowanej bazie. Kolejne miejsca w rankingu należą do nauk medycznych (Komitet Biochemii i Biofizyki PAN – 6,4%, Komitet Neurobiologii PAN – 4,7%) i nauk ścisłych (Komitet Chemii PAN 4,6%, Komitet Informatyki PAN – 4,3%). Także te komitety przodują w liczbie cytowań w przeliczeniu na 1 autora, jak też na 1 publikację (Tabela 2). Większe zróżnicowanie występuje wśród komitetów w zakresie udziału w cytowaniach, z czego 30 komitetów wypracowuje większy udział w liczbie cytowań, niż publikacji (Rycina 5).

Rycina 5. Udział w komitetów w liczbie publikacji i cytowaniach



W przypadku 6 komitetów, ponad połowę cytowań zawdzięczają one jednej osobie (Tabela 2). Najwyższy wynik można zauważyć w Komitecie Socjologii PAN, gdyż Z. Bauman odpowiada za 81% wszystkich cytowań tej jednostki. Ośmiu

autorów odpowiada za największą liczbę cytowań w obu komitetach, których są członkami (spośród 282). Zaś geograf L. Starkel w każdym z trzech komitetów, w którym się udziela, jest osobą mającą najwięcej cytowań.

Tabela 2. Wybrane komitety PAN w świetle statystyk dotyczących cytowań

Komitety z początku rankingu		Komitety z końca rankingu	
Liczba publikacji na 1 autora			
Socjologii	123,4	Transportu	14,1
Chemii	112,9	Technologii Drewna	13,7
Nauk Fizjologicznych	106,7	Górnictwa	13,5
Informatyki	94,3	Architektury i Urbanistyki	8,0
Biochemii i Biofizyki	91,4	Nauk Teologicznych	7,5
Udział jednej osoby z największą liczbą publikacji w ogóle publikacji			
Nauk Filozoficznych	44,9	Geodezji	7,2
Socjologii	38,1	Nauk o Żywności	7,1
Zagospodarowania Ziemi Górskich	35,7	Techniki Rolniczej	7,1
Ekologii	24,0	Górnictwa	6,9
Melioracji i Inżynierii Środowiska	23,2	Nauk Zootechnicznych	6,7
Liczba cytowań na 1 autora			
Socjologii	2912,0*	Nauk o Sztuce	28,5
Nauk Fizjologicznych	1832,7	Transportu	27,2
Chemii	1620,6	Technologii Drewna	23,2
Biochemii i Biofizyki	1534,6	Architektury i Urbanistyki	12,4
Neurobiologii	1441,5	Nauk Teologicznych	5,3
Udział jednej osoby z największą liczbą cytowań w ogólnej liczbie cytowań			
Socjologii	81,6	Cytobiologii	10,7
Zagospodarowania Ziemi Górskich	73,1	Immunologii i Etiologii Zakażeń	9,7
Melioracji i Inżynierii Środowiska	57,1	Neurobiologii	8,9
Inżynierii Środowiska	52,7	Geofizyki	8,3
Rozwoju Człowieka	52,6	Biochemii i Biofizyki	6,6
Liczba cytowań na publikacje			
Socjologii	23,6**	Nauk o Kulturze Antycznej	1,9
Biologii Ewolucyjnej i Teoretycznej	18,8	Technologii Drewna	1,7
Mikrobiologii	17,3	Architektury i Urbanistyki	1,6
Nauk Fizjologicznych	17,2	Nauk o Sztuce	1,5
Biochemii i Biofizyki	16,8	Nauk Teologicznych	0,7
Udział publikacji niecytowanych			
Biologii Ewolucyjnej i Teoretycznej	22,7	Górnictwa	59,8
Chemii	24,2	Architektury i Urbanistyki	60,3
Chemii Analitycznej	24,5	Transportu	61,6
Neurobiologii	25,4	Nauk Filozoficznych	63,0
Biochemii i Biofizyki	25,9	Nauk Teologicznych	69,6

cd. Tabeli 2

Komitety z początku rankingu		Komitety z końca rankingu	
Średnia wskaźnika h			
Nauk Fizjologicznych	20,6	Nauk o Kulturze Antycznej	2,3
Biochemii i Biofizyki	19,1	Transportu	2,3
Neurobiologii	18,4	Technologii Drewna	2,0
Chemii	18,3	Architektury i Urbanistyki	1,5
Cytobiologii	15,7	Nauk Teologicznych	1,0

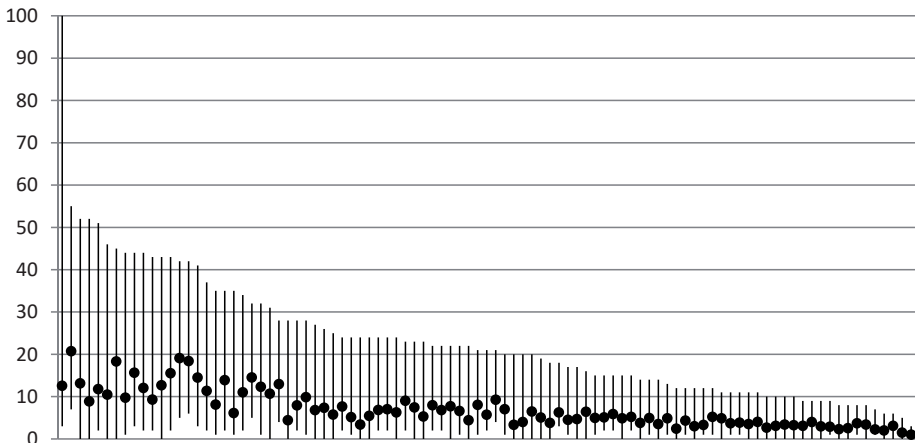
* Bez uwzględnienia prof. Z. Baumana wskaźnik ten spada do 548,8 cytowań na autora.

** Bez uwzględnienia prof. Z. Baumana wskaźnik ten spada do 7 cytowań na publikację.

Jak wspomniano w poprzednim rozdziale, prawie 40% opublikowanych prac członków komitetów nie było cytowanych. Miara ta waha się od 23% dla Komitetu Biologii Ewolucyjnej PAN, do 69% dla Komitetu Nauk Teologicznych PAN. Ponad 21% wszystkich prac w bazie było cytowanych co najmniej 10 razy. Dla dziesięciu komitetów udział ten był większy niż 30%. Ponad 50 cytowań zanotowało jedynie 3,8% publikacji z analizowanej bazy. Średnia ta była większa w 27 komitetach. Publikacji mających powyżej 100 cytowań jest zaledwie 1,3% (z czego 15 komitetów nie miało ani jednej takiej pracy).

Średni indeks H (liczony jako średnia arytmetyczna) dla komitetów waha się od 20,7 (Komitet Nauk Fizjologicznych) do 1 (Komitet Nauk Teologicznych), co daje średnią wartość 7 dla wszystkich komitetów naukowych. Analizując komitety pod kątem poszczególnych autorów należy zwrócić uwagę, że dla 46 z nich wartość maksymalna nie przekracza 20 (Rycina 6). Wartość minimalna powyżej 3 (czyli każdy autor ma indeks H co najmniej 3) występuje w 10 komitetach.

Rycina 6. Indeks H dla poszczególnych komitetów (minimum, maksimum i średnia)



7. Analiza cytowań według dziedzin nauki

Dane zebrane w bazie umożliwiły przeprowadzenie analiz porównawczych pomiędzy poszczególnymi dziedzinami nauki w Polsce. Autorzy zostali zagregowani na podstawie przynależności do komitetu, którego byli członkiem, do jednej z 6 podstawowych dziedzin nauki na podstawie systematyki nauk OECD (2006). Ponieważ część osób zasiadała w dwóch lub trzech komitetach, należało wykluczyć powtarzające się osoby w ramach jednej dziedziny. Po agregacji komitetów, 83 osoby występują w dwóch dziedzinach. Przyporządkowanie danego autora do danej dziedziny następowało każdorazowo, po sprawdzeniu jego afiliacji i przejrzaniu całego dorobku naukowego. Najczęstsze połączenie wynikające z tego rozwiązania występuje między naukami rolniczymi a przyrodniczymi, medycznymi a przyrodniczymi oraz medycznymi a inżynierjno-technicznymi.

Pod względem liczby autorów przyporządkowanych według powyższej zasady zdecydowanie przodują nauki przyrodnicze i techniczne (Tabela 3), które łącznie stanowią połowę badanej grupy autorów. Także autorzy z tych dziedzin stworzyli ponad połowę wszystkich analizowanych publikacji. Jednakże to nauki medyczne mogą poszczycić się największą średnią liczbą publikacji na autora – 68 (prawie 60% autorów opublikowało więcej niż 40 prac, czyli tyle, ile wynosi średnia dla całej bazy), wyprzedzając bardzo wyraźnie pozostałe dziedziny. Drugie miejsce pod tym względem zajmują nauki przyrodnicze, gdzie prawie 40% autorów opublikowało więcej prac, niż wynosi średnia.

Tabela 3. Podstawowe dane dotyczące charakterystyki publikacji w ramach dziedzin

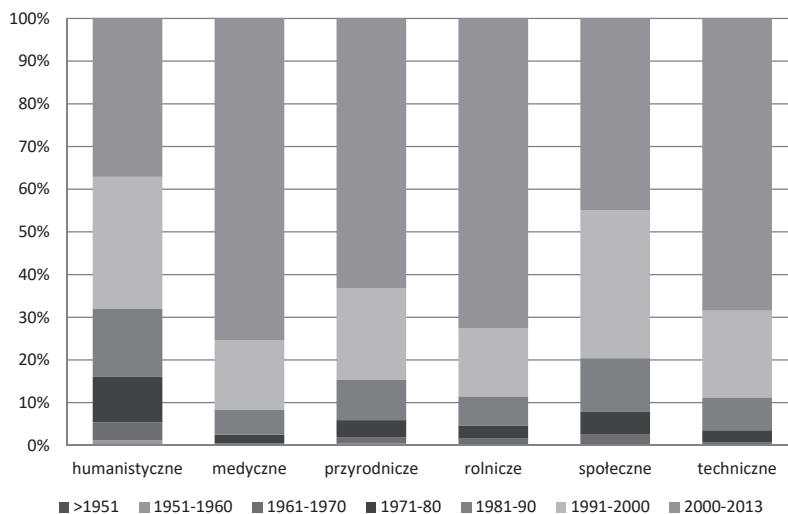
Dziedziny nauki	Liczba autorów	% udział w ogólnej liczbie autorów	Liczba publikacji	% udział w ogólnej liczbie publikacji	Liczba publikacji na 1 autora	Odczynl. standardowe w liczbie publikacji	% udział autorów mających więcej niż 40 prac (średnia)
Humanistyczne	472	15,1	13 131	10,5	27,8	37,9	18,6
Medyczne	328	10,5	22 312	17,9	68,0	65,5	59,8
Przyrodnicze	741	23,7	35 380	28,4	47,7	55,8	38,2
Rolnicze	486	15,5	12 072	9,7	24,8	20,5	17,1
Spółeczne	331	10,6	12 987	10,4	39,2	80,6	28,7
Techniczne	774	24,7	28 685	23,1	37,1	50,2	28,0

Największe zróżnicowania w liczbie publikacji (odchylenie standardowe) występują w naukach społecznych i medycznych. W przypadku nauk społecznych

jest to wynik prac Z. Baumana. Bez publikacji tego autora, odchylenie standardowe wynosi tam 25,6, czyli jest stosunkowo niskie w porównaniu do innych grup.

Analiza publikowanych prac w poszczególnych dziesięcioleciach wskazuje na zmiany w zakresie udziału poszczególnych dziedzin (Rycina 7). Z biegiem czasu zwiększa się udział publikacji z dziedzin technicznych i medycznych, kosztem przede wszystkim prac z zakresu nauk humanistycznych i społecznych. Wynika to z coraz bardziej widocznych różnic między kanałami komunikacji, które są między naukami „twardymi” a „miękkimi”. W ostatniej dekadzie szczególnie prym wiodą nauki przyrodnicze, techniczne i medyczne. Nauki rolnicze zaczęły być bardziej widoczne po roku 2000. Autorzy nauk medycznych wykazują się największą aktywnością publikacyjną po 1990 r. wśród analizowanych dziedzin (ponad 90% prac opublikowanych zostało po tej dacie). Wśród pozostałych dziedzin udział ten waha się od 68 (humanistyka) do 88 (nauki techniczne). Nauki humanistyczne charakteryzują się najbardziej zrównoważoną strukturą pod względem daty publikacji.

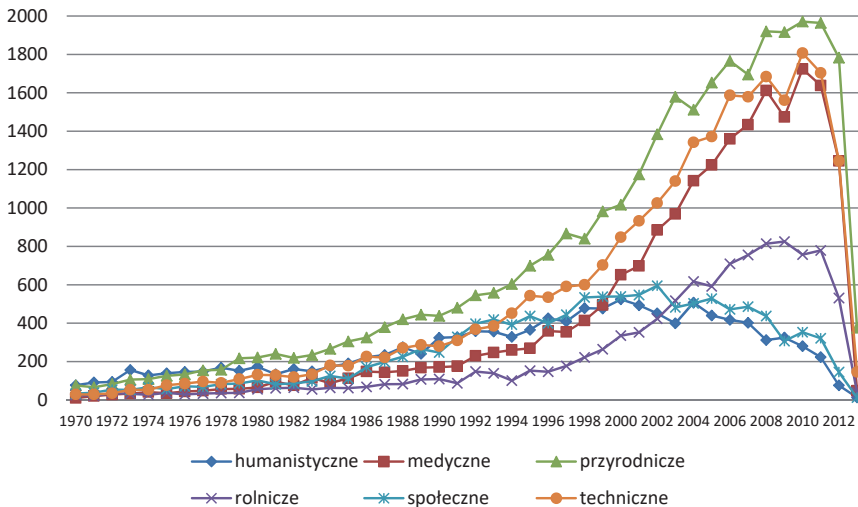
Rycina 7. Procent liczby publikacji w kolejnych dekadach w ramach dziedzin nauki



Analizując liczbę publikacji w kolejnych latach zauważyć można, że bardzo podobną tendencję wzrostową mają nauki medyczne, przyrodnicze i techniczne, osiągając szczególnie po 2000 r. bardzo wysoką dynamikę, która maleje dopiero po 2010 r. Nauki rolnicze mają podobną tendencję, jednakże z mniejszą dynamiką. Najwyższe wartości osiągają w latach 2007–2008, po czym liczba publikacji

systematycznie maleje (Rycina 8). Drugą grupę stanowią nauki społeczne i humanistyczne, notujące mniej dynamiczny wzrost, który już po 2002 r. zamienia się w systematyczny spadek. Prawidłowości te są związane z zasilaniem cyfrowych baz danych publikacjami i pokazują różnice w dostępie dla różnych dziedzin.

Rycina 8. Liczba publikacji w latach 1970–2013 według grup dziedzin



Duża liczba publikacji w naukach przyrodniczych powoduje, iż ta dziedzina odpowiada za największą liczbę cytowań w bazie (ponad 37%; tabela 4). Najmniej cytowań odnosi się do nauk humanistycznych i rolniczych. W przypadku wyłączenia z analizy prof. Baumana, to również nauki społeczne byłyby obecne w tym gronie. Analiza liczby cytowań w przeliczeniu na 1 autora wskazuje na nauki medyczne, jako dziedzinę najbardziej cytowaną. W przypadku liczby cytowań w przeliczeniu na 1 publikację, najlepsze wyniki osiągają zarówno nauki medyczne, przyrodnicze i społeczne. Także te trzy dziedziny nie przekraczają średniej (40%) w przypadku udziału prac bez cytowań. Takich prac w naukach humanistycznych i technicznych jest prawie połowa. W przypadku większości dziedzin, ponad 80% cytowań zostało wypracowanych przez publikacje, które ukazały się po 1990 r. (Rycina 9). Jedynie nauki humanistyczne (52%) wyłamują się z tego schematu, co potwierdza obserwacje poczynione już przez J. Cullarsa (1985), który wskazywał, że najczęściej cytowane prace w tej dziedzinie mają co najmniej dekadę „życia”. Prace z ostatniej dekady są najważniejsze dla pięciu dziedzin (od 43% do 60% cytowań dotyczy prac tego okresu). Jedynie w przypadku nauk humanistycznych to prace z lat 1991–2000 uzyskały największą liczbę cytowań.

Rycina 9. Procent cytowań z poszczególnych dekad według dziedzin

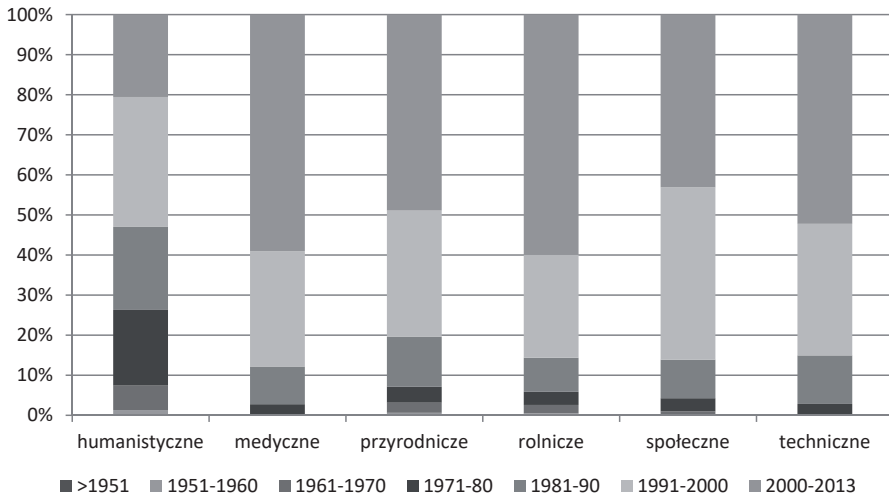
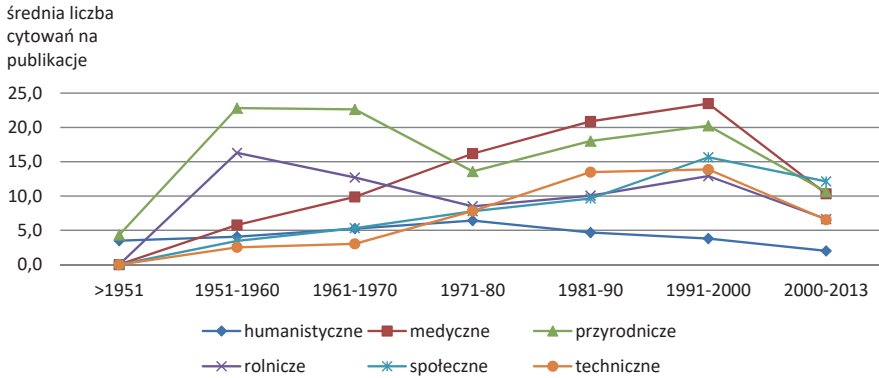


Tabela 4. Wybrane statystyki dziedzin dotyczące cytowań

Dziedzina	Liczba cytowań	Udział cytowań (%)	Cytowania w przeliczeniu		Udział publikacji z liczbą cytowań (%)				
			na 1 autora	na 1 publikację	0	> 10	> 30	> 50	> 100
Humanistyczne	46 116	3,8	97,7	3,5	48,1	8,3	1,9	0,9	0,2
Medyczne	268 966	22,1	820,0	11,9	39,7	26,0	10,0	5,4	1,9
Przyrodnicze	453 272	37,3	611,7	12,7	32,62	29,5	10,4	5,3	1,8
Rolnicze	84 805	7,0	174,5	6,9	40,4	17,0	5,0	2,3	0,8
Społeczne	145 533	12,0	439,7	11,2	38,8	14,9	5,2	2,9	1,3
Techniczne	216 359	17,8	279,5	7,5	46,7	15,7	5,1	2,7	1,0

Obliczając średnią liczbę cytowań na liczbę publikacji można zauważyć, że w zależności od dziedziny nauki, prace z różnych okresów cieszyły się różną popularnością (Rycina 10). W przypadku nauk przyrodniczych, największa średnia cytowań prac należy do okresu 1951–70. Dla nauk rolniczych i społecznych najlepszy okres to lata 1971–80. W naukach medycznych i technicznych widać, że im nowsze publikacje, tym średnia cytowania jest większa, co wiąże się zapewne z dużą zależnością od rozwoju technologicznego, który odpowiada za stosowane narzędzia, a co za tym idzie metody badawcze.

Rycina 10. Średnia liczba cytowań na publikację w dekadach



Sprawdzono również, jak prawo Lotki jest obecne w poszczególnych dziedzinach, szczególnie do proporcji zaproponowanych przez Vilfreda Pareto. Jedyne w przypadku liczby cytowań wśród trzech grup (nauki humanistyczne, społeczne, techniczne) można zauważyć zbliżony procent do wzorcowych 80%. W pozostałych dziedzinach udział ten jest bliższy 70%.

Tabela 5. Wybrane wskaźniki dotyczące koncentracji cytowań

Dziedzina nauki	Udział 20% autorów w liczbie publikacji	Udział 20% autorów w liczbie cytowań	wskaźnik Herfindahla	h-index	m-quotient	m-index	a-index	R-index	q ² -index	Autor największej liczby cytowań (liczba cytowań, %)	Udział cytowań 10 najpopularniejszych autorów w %
Humanistyczne	56,6	77,8	128,5	3,7	0,11	7,7	9,6	6,0	5,2	Andrzej Walicki (1989, 4,2%)	30,0
Medyczne	49,8	70,6	117,2	11,6	0,44	23,7	34	20,4	16,4	Krzysztof Narkiewicz (10038, 3,7%)	25,0
Przyrodnicze	52,9	71,3	54,5	9,9	0,35	21,0	29,7	17,1	14,1	Wacław Szybalski (9392, 2,0%)	14,3

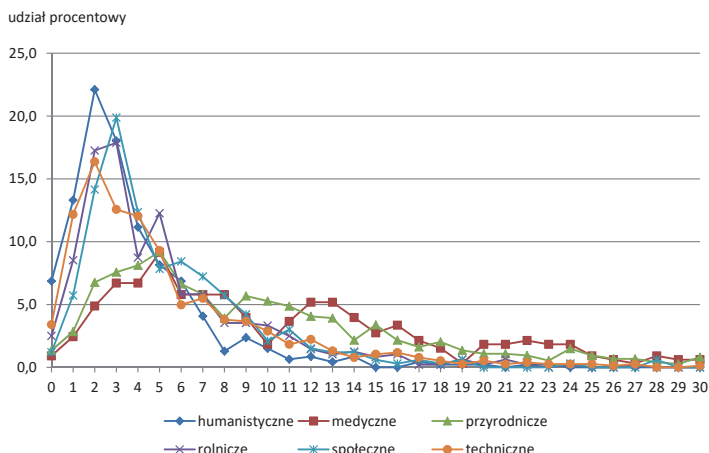
cd. Tabeli 5

Dziedzina nauki	Udział 20% autorów w liczbie publikacji	Udział 20% autorów w liczbie cytowań	wskaznik Herfindahla	h-index	m-quotient	m-index	a-index	R-index	q ² -index	Autor największej liczby cytowań (liczba cytowań, %)	Udział cytowań 10 najpopularniejszych autorów w %
Rolnicze	47,3	74,6	131,8	5,2	0,21	11,7	15,6	8,8	7,5	Jacek Oleksyn (3588, 6,5%)	27,2
Społeczne	57,9	86,8	2933,5	5,8	0,19	13,2	18,7	10,2	8,5	Zygmunt Bauman (78399, 53,7%)	68,2
Techniczne	57,2	82,9	144,9	5,6	0,21	12,5	17,8	9,9	8,1	Andrzej Skowron (12822, 5,9%)	30,5

Obliczono także miarę koncentracji dla poszczególnych dziedzin (tabela 5) z pomocą wskaźnika Herfindahla. Widoczną koncentrację wskazują nauki społeczne, co jest wynikiem dominacji prof. Baumana. Jego cytowania stanowią ponad 54% wszystkich cytowań w tej dziedzinie. Z tego samego powodu nauki te wyróżniają się na tle innych pod względem udziału dziesięciu najpopularniejszych autorów. Zazwyczaj odpowiadają oni za mniej niż 30% cytowań, zaś w przypadku nauk społecznych jest to prawie 70%. Najmniejszą koncentrację zaś wykazują nauki przyrodnicze, gdzie udział 10 najpopularniejszych autorów stanowi 14%.

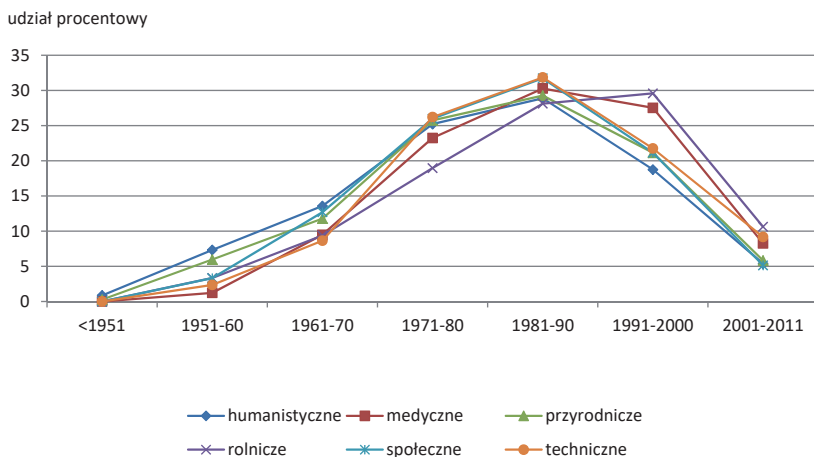
Największy udział procentowy autorów o indeksie $h = 0$ (Rycina 11) dotyczy nauk humanistycznych (blisko 7%), najmniejszy zaś – nauk medycznych i społecznych (około 1%). Dla większości nauk najwyższy udział autorów z danym indeksem występuje dla wartości 2 lub 3. Jedynie nauki medyczne i przyrodnicze mają największy udział dla większych wartości (5). To także te dziedziny dominują nad pozostałymi w przypadku udziału osób z indeksami w przedziale 11–30, a także mają największy udział osób z indeksem $h > 30$. Dla nauk medycznych takie osoby stanowią 4,6%, zaś dla nauk przyrodniczych – 2,3%. W przypadku pozostałych czterech dziedzin występują nieliczne przypadki osób mających tak wysoki indeks, co składa się na udział poniżej 1%, a w przypadku nauk humanistycznych nie ma ani jednej takiej osoby.

Rycina 11. Udział procentowy autorów mających danych wskaźnik h w dziedzinach nauki



Dla większości dziedzin autorzy rozpoczynali swoją karierę naukową (rozumianą jako opublikowanie pierwszej zaindeksowanej publikacji) w okresie 1981–1990 (Rycina 12). Wyjątkiem są nauki rolnicze, gdzie najwięcej autorów debiutowało w latach 1990–2000. Porównując dziedziny pod kątem rozpoczęcia kariery naukowej można stwierdzić, że struktura wieku jest porównywalna. Chociaż patrząc na udział osób z młodszym stażem naukowym widać, że nauki rolnicze i medyczne są najbardziej „otwarte” na młodych naukowców. Oczywiście w przypadku tej charakterystyki jest to wynik przede wszystkim wyboru osób do komitetów naukowych i nie odzwierciedla trendu dla całych dziedzin.

Rycina 12. Okresy pierwszej publikacji wśród autorów wyróżnionych dziedzin

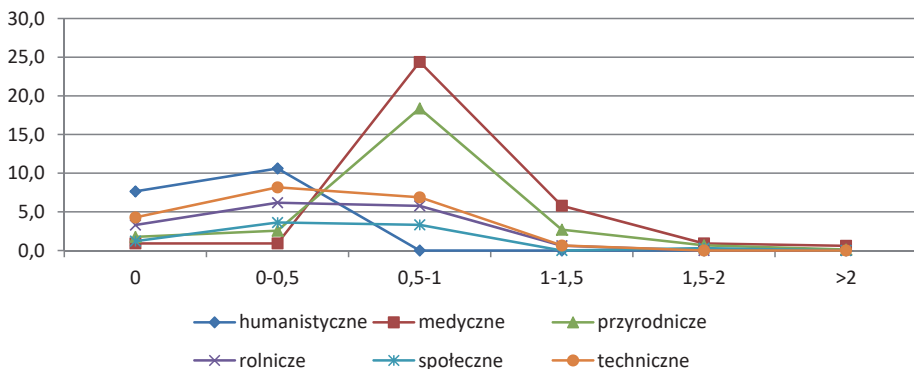


Rok pierwszej publikacji umożliwia obliczenie wyżej wspomnianego wskaźnika m-quotient (Rycina 13) dla poszczególnych dziedzin (rozumianego, jako średnia arytmetyczna autorów). Pod tym kątem przodują nauki medyczne, które wydają się być najlepszym środowiskiem do osiągnięcia wysokiego wskaźnika h, jak też m-quotient. Na drugim miejscu znajdują się nauki przyrodnicze, a po nich mniej więcej na tym samym poziomie znajdują się nauki rolnicze, techniczne i społeczne. Najgorzej pod tym względem prezentują się nauki humanistyczne, gdzie najtrudniej osiągnąć dobry wskaźnik m-quotient, co potwierdza zależność wskaźników od kultury cytowań.

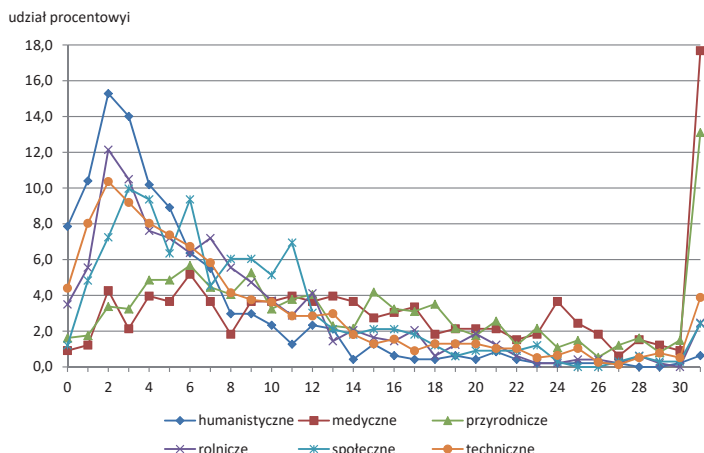
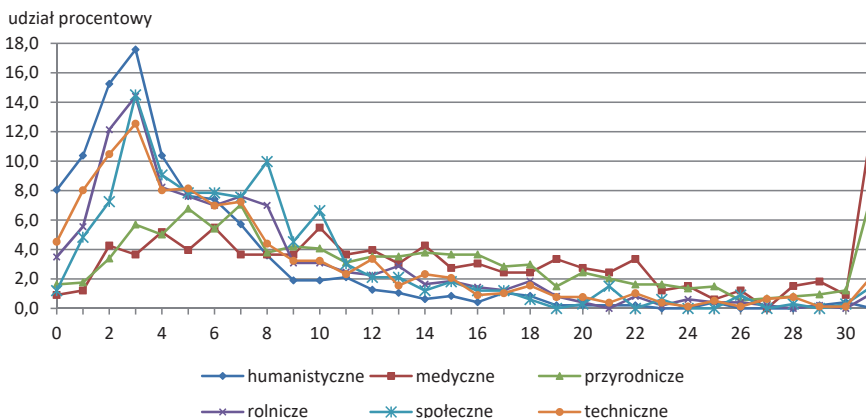
Wskaźniki m-index i a-index odnoszą się do tzw. rdzenia Hirscha, czyli publikacji wliczanych do wskaźnika H. M-index oznacza medianę cytowań wśród tych publikacji, natomiast a-index wartość średniej arytmetycznej. Potwierdzają one prymat nauk medycznych, jako publikacji o największym oddziaływaniu. Pozwalają one także obliczyć odpowiednio wskaźniki R-index (Rycina 14) i q^2 -index (Rycina 15), które łączą informacją o charakterze ilościowym (indeks H) z informacją bardziej o charakterze jakościowym, czyli jak dużo cytowań uzyskały najlepsze prace danego autora. Dzięki temu autor mający taki sam indeks H, jak jego kolega, ale jednocześnie więcej cytowań najpopularniejszych prac, będzie miał wyższy wskaźnik R-index i q^2 -index. Stąd zauważyć można, że nauki medyczne i przyrodnicze mając zbliżony średni indeks H, różnią się pod względem cytowań, co pokazują oba wskaźniki R-index i q^2 -index, gdzie różnica pomiędzy tymi dziedzinami jest wyższa. Podobnie wśród nauk rolniczych, społecznych i technicznych, gdzie zaznacza się wyraźniej prymat nauk społecznych.

Rycina 13. Udział autorów w dziedzinach nauki według wartości m-quotient

udział procentowy



Rycina 14. Udział autorów w dziedzinach nauki z poszczególnym zakresem indeksu R

Rycina 15. Udział autorów w dziedzinach nauki z poszczególnym indeksem q^2 

8. Dyskusja i wnioski

Wnioski z analiz są metodyczne i poznawcze. W przypadku korzystania z wyszukiwarki Google Scholar (*Publish or Perish*) przede wszystkim należy podkreślić, że pozwala ona na przeprowadzanie analiz bibliometrycznych na skalę wcześniej w Polsce nieznaną. Natomiast w kwestiach szczegółowych nasuwa się więcej uwag metodycznych, związanych zarówno ze specyfiką językową, jak też narodową w sensie prawidłowości rozwoju nauki.

Analizy wykazały, że problemem jest brak rozróżnialności polskich liter, co sprawia, że dane autorów z Polski są bardziej podatne na błędy wynikające z pisowni nazwisk i imion. Skutkuje to albo brakiem części prac, albo też przypisaniem prac

autorstwa innych osób. W chwili obecnej brak jest innych danych identyfikujących autora, co w przypadku osób o tym samym imieniu i nazwisku powoduje trudności i pochłania dużo czasu w odpowiednim przypisaniu prac autorów.

Wyszukiwarka Google Scholar nie ma umiejętności łączenia tej samej pracy występującej w różnych wersjach językowych, co powoduje pewne zniekształcenia statystyk na poziomie poszczególnych prac i przez to modyfikację niektórych wskaźników dla autorów. Problem ten dotyka jednak nielicznej grupy badaczy, którzy publikują monografię w języku innym, niż angielskim.

Na fakt pojawienia się polskich prac w Google Scholar wpływa wiele czynników. Przede wszystkim zależy to od sposobu przygotowanego przez właściciela praw autorskich materiału źródłowego, czyli głównie wydawnictw. Jeśli taki podmiot jest świadomy, w jaki sposób należy przygotować publikacje, aby była indeksowana w Google Scholar, to znacznie zwiększają się szanse na szybkie pojawienie się pracy w wyszukiwarce. W rzeczywistości może to jednak trwać nawet kilka miesięcy. Jednakże w wielu przypadkach proces dostosowywania wydawnictwa i jego prac do założeń wyszukiwarki wciąż trwa i z tego powodu wiele prac jest skutecznie indeksowanych nawet po kilkunastu miesiącach.

Analiza prac członków komitetów naukowych Polskiej Akademii Nauk w Google Scholar wskazuje na istnienie istotnych różnic pomiędzy członkami, komitetami, a także dziedzinami naukowymi, potwierdzając tym samym zależność wskaźników bibliograficznych od różnej kultury cytowań.

Stwierdzono dosyć niską średnią liczbę publikacji, wynosząca około 40 na jednego autora. Tylko 1/4 autorów posiadała w wyszukiwarce więcej niż 50 prac. To słaby wynik, jeśli uznamy, że członkowie komitetów naukowych są czołowymi przedstawicielami nauki polskiej, a średnia data pierwszej publikacji zindeksowanej przez Google Scholar to rok 1973. Daje to, bowiem mniej niż 1 pracę na rok. Wynik ten jest spowodowany jednak w dużej mierze brakiem wielu starszych prac w bazie Google Scholar, zarówno periodyków jak i monografii. Współcześnie młodszy naukowcy w Polsce mają większą skłonność do dbania, by wersje elektroniczne ich prac były powszechnie dostępne w Internecie, a przez to indeksowane przez wyszukiwarkę, nawet czasami wyprzedzając w tym zakresie wydawnictwa, które są najczęściej właścicielami praw autorskich.

Niewysoka liczba prac w Google Scholar skutkuje także tym, że ponad 100 autorów (3,4% całej populacji członków komitetów PAN) nie zanotowało w bazie ani jednego cytowania. Brak jakiegokolwiek cytowania występuje u ponad 40% wszystkich prac, co jest wynikiem jednak lepszym, niż od klasycznej proporcji zauważonej przez E. Garfielda (1989), który podawał, że ponad połowa prac nie jest cytowanych.

Różnice w zakresie charakterystyk bibliograficznych są obecne także wśród 95 komitetów naukowych. Odnosząc się do średniej około 40 prac na naukowca, wskaźnik ten waha się od ponad 100 (komitety: socjologiczny, chemiczny, nauk

fizjologicznych) do poniżej 10 na osobę (architektura i urbanistyka oraz nauki teologiczne). Średnio, najbardziej twórczy autor odpowiada za 13% prac w komitecie, z czego udział ten dochodzi do nawet ponad 40% (Komitet Nauk Filozoficznych PAN). Jeszcze większą dominację widać w zakresie cytowań, gdyż średnio na jednego najpopularniejszego autora przypada 25% cytowań, jednakże w przypadku 7 komitetów jest to ponad 50%, osiągając skrajną wartość 82% dla Komitetu Socjologii PAN w przypadku Z. Baumana.

Analiza cytawalności poszczególnych komitetów wskazuje, że aż dla 29 z nich, więcej prac nie było cytowanych niż te, które zanotowały chociaż jedną cytację, osiągając najwyższą wartość 69% dla nauk teologicznych. Z kolei publikacji mających powyżej 100 cytowań w całej bazie jest zaledwie 1,3% (z czego 15 komitetów nie miało ani jednej takiej pracy). Analiza relacji pomiędzy liczbą publikacji, a cytowaniami poszczególnych prac w ramach komitetów w dużej mierze potwierdza istnienie takiej zależności (współczynnik korelacji Spearmanna wyniósł 0,76, przy istotności statystycznej 0,05). Jest jednak grupa komitetów, które wymykają się takiemu schematowi. Część z nich pomimo dużej liczby publikacji publikowanych przez swoich członków nie notuje tak wysokiej cytawalności publikowanych prac. Są to komitety PAN m.in. Językoznawstwa, Rozwoju Człowieka, Biocybernetyki i Inżynierii, Fizyki Medycznej Radiobiologii, Mechaniki, Nauk Ekonomicznych. Z drugiej strony jest kilka komitetów, które mimo relatywnie niewielkiej liczby publikacji, osiągają cytawalność prac na wyższym poziomie. Są to m.in. komitety: Astronomii, Botaniki oraz Ochrony Przyrody.

Podział na 6 grup dziedzin nauki dowiódł istotnych różnicowań zarówno w zakresie wzorców publikowania, jak i cytowania. W naukach medycznych odnotowano największą średnią liczbą prac przypadającą na 1 autora (68 prac). Ponadto blisko 60% autorów opublikowało więcej niż 40 prac (czyli tyle, ile wynosi średnia dla całej zgromadzonej bazy), wyprzedzając bardzo wyraźnie pozostałe dziedziny. Drugą dziedziną pod tym względem są nauki przyrodnicze, gdzie prawie 40% autorów opublikowało więcej prac, niż wynosi średnia. Autorzy publikujący zatem w tych dwóch dziedzinach, mają większą szansę na otrzymanie cytowania, niż autorzy z innych dziedzin nauki.

W przypadku średniej liczby cytowań na publikacje, najlepsze wskaźniki odnotowują zarówno nauki medyczne, jak też przyrodnicze i społeczne. Także te trzy dziedziny nie przekraczają średniej (40%) w przypadku udziału prac bez cytowań. Takich prac w naukach humanistycznych i technicznych jest prawie połowa, co jest podobnym zjawiskiem jak i w innych krajach. W przypadku większości dziedzin ponad 80% cytowań zostało wypracowanych przez publikacje, które ukazały się po 1990 roku. Jedynie nauki humanistyczne (52%) wyłamują się tego schematu. Prace z ostatniej dekady są najważniejsze dla pięciu dziedzin (od 43% do 60% cytowań pochodzi z prac tego okresu). Jedynie w przypadku nauk humanistycznych to prace z lat 1991–2000 uzyskały największą liczbę cytowań. Potwierdza to obserwacje

poczynione już przez J. Cullarsa (1985), który wskazywał, że najczęściej cytowane prace w tej dziedzinie mają co najmniej dekadę „życia”.

Liczba cytowań poszczególnych prac wskazuje, że dobra praca może być cytowana przez kolejne pokolenia. Większą szansę na takie „długie życie” mają autorzy z nauk humanistycznych, przyrodniczych oraz społecznych.

Wykazane duże różnice w liczbie cytowań skutkują także różnymi indeksami Hirscha w poszczególnych dziedzinach. Największy udział procentowy autorów z indekse $H = 0$ charakteryzuje nauki humanistyczne (blisko 7% autorów), najmniejszy zaś w naukach medycznych i społecznych (około 1%). Dla większości nauk największy udział autorów z danym indeksem występuje dla wartości 2 lub 3. Jedynie nauki medyczne i przyrodnicze mają największy udział dla większych wartości a ($H = 5$). To także te dziedziny dominują nad pozostałymi w przypadku udziału osób z indeksami w przedziale 11–30, a także mają największy udział osób z indeksem powyżej 30. Dlatego analizy potwierdziły, że nieuprawnione jest ocenienie dorobku publikacyjnego naukowców z różnych dziedzin na podstawie wskaźników bibliometrycznych tego typu, jak indeks Hirscha.

Literatura

- Aalbers M.B., Rossi U., 2009, *Anglo-American/Anglophone hegemony*, w: Kitchin R., Thrift N., red., „International Encyclopedia of Human Geography”, 1, New York: Elsevier, 116–121.
- Abramo G., D’Angelo C.A., 2015, *A methodology to compute the territorial productivity of scientists: The case of Italy*, „Journal of Informetrics”, 9.4: 675–685.
- Aguillo I.F., 2012, *Is Google Scholar useful for bibliometrics? A webometric analysis*, „Scientometrics”, 91.2: 343–351.
- Althouse B.M., West J.D., Bergstrom C.T., Bergstrom T., 2009, *Differences in impact factor across fields and over time*, „Journal of the Association for Information Science and Technology”, 60.1: 27–34.
- Amara N., Landry R., 2012, *Counting citations in the field of business and management: Why use Google Scholar rather than the Web of Science*, „Scientometrics”, 93:553–581.
- Bajerski A., 2008a, *Polskie czasopisma geograficzne w bazie Scopus: próba analizy komunikacji naukowej w polskiej geografii*, „Czasopismo Geograficzne”, 79.3: 367–382.
- Bajerski A., 2008b, *Ranking ośrodków geografii i społeczno-ekonomicznej w Polsce na podstawie cytowań w bazach Web of Science*, „Przegląd Geograficzny”, 80.4: 579–589.
- Bar-Ilan J., 2008, *Which h-index? – A comparison of WoS, Scopus and Google Scholar*, „Scientometrics”, 74.2: 257–271.
- Bar-Ilan J., 2010, *Citations to the Introduction to informetrics indexed by WOS, Scopus and Google Scholar*, „Scientometrics”, 82: 495–506.
- Batista P.D., Campiteli M.G., Kinouchi O., Martinez A.S., 2006, *Is it possible to compare researchers with different scientific interests?*, „Scientometrics”, 68.1: 179–189.
- Błocki Z., Życzkowski K., 2013, *Czy można porównywać jabłka i gruszki? O danych bibliometrycznych w różnych dziedzinach nauki*, „Nauka”, 2: 37–46.
- Bornmann L., Daniel H.D., 2007, *What do we know about the h-index?*, „Journal of the American Society for Information Science and Technology”, 58.9: 1381–1385.

- Bornmann L., Mutz R., Daniel H.D., 2008, *Are there better indices for evaluation purposes than the h-index? A comparison of nine different variants of the h-index using data from biomedicine*, „Journal of the American Society for Information Science and Technology”, 59.5: 830–837.
- Bosman J., Van Mourik I., Rasch M., Sieverts E., Verhoeff H., 2006, Scopus Reviewed and Compared, Dokument internetowy: <http://igiturarchive.library.uu.nl/DARLIN/2006-1220-200432/Scopus%20doorgelicht%20&%20vergeleken%20-%20translated.pdf> [dostęp: 12.03.2016].
- Chang Y.W., 2013, *A comparison of citation contexts between natural sciences and social sciences and humanities*, „Scientometrics”, 96.1: 535–553.
- Costas R., Bordons M., 2007, *Advantages, limitations and its relation with other bibliometric indicators at the micro level*, „Journal of Informetrics”, 1.3: 193–203.
- Cullars J.M., 1985, *Characteristics of the Monographic Literature of British and American Literary Studies*, „College and Research Libraries”, 46.6: 511–522.
- Cullars J.M., 1988, *Characteristics of the Monographic Literature of British and American Literary Studies*, „College and Research Libraries”, 49.2: 157–170.
- De Almeida E.C.E., Guimaraes J.A., 2013, *Brazil's growing production of scientific articles—how are we doing with review articles and other qualitative indicators?*, „Scientometrics”, 97.2: 287–315.
- De Winter J.C.F., Zadpoor A.A., Dodou D., 2014, *The expansion of Google Scholar versus Web of Science: A longitudinal study*, „Scientometrics”, 98: 1547–1565.
- Drabek A., 2001, *Bibliometryczna analiza czasopism naukowych w dziedzinie nauk społecznych*, Praca doktorska, Katowice.
- Drabek A., Rozkosz E.A., Hołowiecki M., Kulczycki E., 2015, *Polski Współczynnik Wpływu a kultura cytowań w humanistyce*, „Nauka i Szkolnictwo Wyższe”, 2.46: 121–138.
- Egghe L., 2006, *Theory and practice of the g-index*, „Scientometrics”, 69.1: 131–152.
- Egghe L., 2006a, *Dynamic h-index: The Hirsch Index in Function of Time*, Dokument internetowy: <http://doclib.luc.ac.be/dspace/bitstream/1942/980/1/dynamic+h.pdf> [dostęp: 12.03.2016].
- Egghe L., Rousseau R., 2008, *An h-index weighted by citation impact*, „Information Processing & Management”, 44.2: 770–780.
- Etxebarria G., Gomez-Uranga M., 2010, *Use of Scopus and Google Scholar to measure social sciences production in four major Spanish universities*, „Scientometrics”, 82: 333–349.
- Fingerman S., 2005, *Scopus: profusion and confusion*, „Online Medford”, 29.2: 36–38.
- Franceschet M., 2010, *A comparison of bibliometric indicators for computer science scholars and journals on Web of Science and Google Scholar*, „Scientometrics”, 83: 243–258.
- Garfield G., 1965, *Can citation indexing be automated?*, w: Stevens M.E., Giuliano V.E., Heilprin L.B., red., *Statistical Association Methods for Mechanized Documentation. Symposium Proceedings*. Washington 1964, National Bureau of Standards Miscellaneous, 269, Washington, 189–192.
- Garfield E., 1989, „Citation Classics and citation behavior revisited”, *Current Contents*, 5, 3–9.
- Hammarfelt B., 2012, *Harvesting footnotes in a rural field: Citation patterns in Swedish literary studies*, „Journal of Documentation”, 68.4: 536–558.
- Harzing A.W., 2013, *A preliminary test of Google Scholar as a source for citation data: A longitudinal study of Nobel prize winners*, „Scientometrics”, 94: 1057–1075.
- Harzing A.W., 2007, Publish or Perish, <http://www.harzing.com/pop.htm>
- Harzing A.W., 2011, *The Publish or Perish Book. Part 1–3*, Melbourne: Tarma Software Research.
- Hellqvist B., 2009, *Referencing in the Humanities and its Implications for Citation Analysis*. *Journal of the American Society for Information Science and Technology*, 61.2: 310–318.
- Hirsch J.E., 2005, *An Index to Quantify an Individual's Scientific Research Output*, „PNAS”, 102: 16569–16572.
- Ho Y.S., 2014, *Classic articles on social work field in Social Science Citation Index: A bibliometric analysis*, „Scientometrics”, 98.500: 137–155.

- Jin B.H., Liang L.M., Rousseau R., Egghe L., 2007, *The R- and AR-indices: Complementing the h-index*, „Chinese Science Bulletin”, 52.6: 855–863.
- Jin B., 2006, *h-Index: An evaluation indicator proposed by scientist*, „Science Focus”, 1.1: 8–9.
- Kelly C.D., Jennions M.D., 2006, *The h-index and career assessment by numbers*, „TRENDS in Ecology and Evolution”, 21.4: 167–170.
- Jacsó P., 2005, *Google Scholar: the pros and the cons*, „Online Information Review”, 29.2: 208–214.
- Kierzek R., 2009, *Jak porównać „apples and oranges”, czyli o różnych metodach analizy publikowalności i dorobku naukowego*, „Sprawy Nauki”, 2: 33–41.
- Kierzek R., 2010, *Publikowalność naukowa w Polsce*, „Forum Akademickie”, 7–8.
- Klaic B., 1995, *Analysis of the scientific productivity of researchers from the Republic of Croatia for the period 1990–1992*, „Scientometrics”, 32.2: 133–152.
- Klinczewicz K., 2007, *Polski system innowacji w obszarze informatyk i analiza bibliometryczna*, w: Okoń Horodyńska E., Zachorowska-Mazurkiewicz A., red., *Innowacje w rozwoju gospodarki i przedsiębiorstw: siły motoryczne i bariery*, Warszawa: Instytut Wiedzy i Innowacji, 17–42.
- Klinczewicz K., 2008, *Polska innowacyjność. Analiza bibliometryczna*, Warszawa: Wydawnictwo Naukowe Wydziału Zarządzania Uniwersytetu Warszawskiego.
- Komperda A., Urbańczyk B., 2010, *Analiza naukometryczna dorobku publikacyjnego PWr.*, „Pryzmat”, 239, Dokument internetowy: http://pryzmat.pwr.wroc.pl/Pryzmat_239/pryzmat239.pdf [dostęp: 12.03.2016].
- Kosmulski M., 2006, *A new Hirsch-type index saves time and works equally well as the original h-index*, „ISSI Newsletter”, 2.3: 4–6.
- Kousha K., Thelwall M., 2008, *Sources of Google Scholar citations outside the Science Citation Index: A comparison between four science disciplines*, „Scientometrics”, 74.2: 273–294.
- Laloë F., Mosseri R., 2009, *Not even right, not even wrong*, „Europhysics News”, 40.5: 27–29.
- Liang L., 2006, *H-index sequence and h-index matrix: Constructions and applications*, „Scientometrics”, 69.1: 163–169.
- Liu Y., Zuo W., Gao Y., Qiao Y., 2013, *Comprehensive geometrical interpretation of h-type indices*, „Scientometrics”, 96: 605–615.
- Lotka A.J., 1926, *The frequency distribution of scientific productivity*, „Journal of the Washington Academy of Sciences”, 16: 317–323.
- López-Cózar E.D., Robinson-García N., Torres Salinas D., 2012, *Manipulating Google Scholar Citations and Google Scholar Metrics: simple, easy and tempting*, „EC3 Working Papers”, 6.
- Marszakowa-Szajkiewicz I., 1996, *Bibliometryczna analiza współczesnej nauki*, Katowice.
- Mikki S., 2010, *Comparing Google Scholar and ISI Web of Science for earth sciences*, „Scientometrics”, 82.123: 321–331.
- Mingers J., Lipitakis E.A.E.C.G., 2010, *Counting the citations: A comparison of Web of Science and Google Scholar in the field of business and management*, „Scientometrics”, 85: 613–625.
- Moed H.F., Burger W.J.M., Frankfort J.G., Van Raan A.F.J., 1985, *The application of bibliometric indicators: Important field- and time-dependent factors to be considered*, „Scientometrics”, 8.3: 177–203.
- Moya-Anegón F., Chinchilla-Rodríguez Z., Vargas-Quesada B., Corera-Álvarez E., Muñoz-Fernández F.J., González-Molina A., Herrero-Solana V., 2007, *Coverage analysis of Scopus: A journal metric approach*, „Scientometrics”, 73.1: 53–78.
- Neuhaus C., Neuhaus E., Asher A., Wrede C., 2006, *The depth and breadth of Google Scholar: An empirical study portal*, „Libraries and the Academy”, 6.2: 127–141.
- Nowak P., 2006, *Bibliometria. Webometria. Podstawy. Wybrane zastosowania*, Poznań.
- OECD, 2006, *Klasyfikacja dziedzin nauki i techniki według OECD*, Paris: OECD.
- OECD, 2012, *Main Science and Technology Indicators*, Paris: OECD.
- Olechnicka A., Płoszaj A., 2008, „Polska nauka w sieci?”, Warszawa.

- Osiński Z., 2012, Bibliometria metodą analizy i oceny dorobku naukowego historyków najnowszych dziejów Polski, w: Matczuk A., Znajomski A., *Kultura, historia, książka*, Lublin: UMCS, 605–616.
- Orduna-Malea E., Lopez-Cozar E.D., 2014, *Google Scholar Metrics evolution: an analysis according to languages*, „Scientometrics”, 98.3: 2353–2367.
- Paasi A., 2005, *Globalisation, academic capitalism and the uneven geographies of international journal publishing spaces*, „Environment and Planning A”, 37.5: 769–789.
- Perc M., 2010, *Growth and structure of Slovenia's scientific collaboration network*, „Journal of Informetrics”, 4.4: 475–482.
- Petersen A.M., Wang F., Stanley H.E., 2010, *Methods for measuring the citations and productivity of scientists across time and discipline*, „Physical Review E – Statistical, Nonlinear, and Soft Matter Physics”, 81.i: 1–25.
- Pilc A., 2004, *Ocena dokonania instytucji naukowej poprzez analizę cytowań na przykładzie Instytutu Farmakologii PAN w Krakowie*, „Nauka”, 2: 149–153.
- Pile A., 2015, *Nauka w kraju – może nie jest tak źle*, „PAUza”, 281.
- Pindlowa W., 1989, *Bibliometria, informetria i scientometria – refleksje terminologiczne i wzajemne relacje*, „Zeszyty Naukowe Uniwersytetu Jagiellońskiego Zeszyty Historycznoliterackie”, 74: 63–73.
- Price D.J. de S., 1970, *Citation Measures of Hard Science, Soft Science, Technology, and Nonscience*, w: C.E. Nelson, D.K. Pollock, red., „Communication among scientists and engineers”, Lexington, MA: Heath Lexington, 3–22.
- Racki G., 1998, *Najbardziej znane polskie publikacje zagraniczne w dziedzinie nauk o Ziemi z lat 1981–1995 na podstawie National Citation Report – Poland*, „Przegląd Geologiczny”, 2: 133–137.
- Racki G., 2001, *Najczęściej cytowane polskie publikacje z dziedziny nauk o Ziemi z lat 90. na podstawie National Citation Reports – Poland 1999 Instytutu Informacji Naukowej (ISI) w Filadelfii*, „Przegląd Geologiczny”, 7: 584–590.
- Racki G., 2003, *Polskie czasopisma geograficzne a międzynarodowy obieg informacji naukowej*, „Przegląd Geograficzny”, 75.1: 101–119.
- Racki G., Drabek A., 2013, *Cytowania i wskaźnik Hirscha: gdzie szukać, jak obliczać?*, „Forum Akademickie”, 2: 40–43.
- Rejn B., 2003, *Działalność badawczo-rozwojowa (B+R) – nakłady, efekty*, Warszawa: Zakład Badań Społeczno-Ekonomicznych.
- Sadowski J., 1995, *Porównawcza ocena wagi publikacji w różnych dziedzinach nauki – czyli jak nie skrzywdzić hodowców frezji*, „Zagadnienia Naukoznawstwa”, 3–4: 59–60.
- Smith A.G., 2008, *Benchmarking Google Scholar with the New Zealand PBRF research assessment exercise*, „Scientometrics”, 74.2: 309–316.
- Stefaniak B., 2008, *Bibliometria w zarządzaniu informacją* w: Pietruch-Reizes D., red., „Zarządzanie informacją w nauce”, Katowice 2008.
- Śleszyński P., 2009, *Pozycja polskich czasopism i serii geograficznych w świetle baz Google Scholar*, „Przegląd Geograficzny”, 81.4: 551–578.
- Śleszyński P., 2013, *Cytowania polskich czasopism naukowych z zakresu geografii i badań regionalnych po 1990 r.*, „Studia Regionalne i Lokalne”, 3.53: 75–88.
- Śleszyński P., 2013, *Cytowania i oddziaływanie polskich ośrodków geograficznych według Google Scholar*, „Przegląd Geograficzny”, 85.4: 599–628.
- Śleszyński P., 2014, *Niedoceniana wartość baz*, „Forum Akademickie”, 2.
- Wróblewski A.K., 2001, *Bibliometryczne nieporozumienia*, „Forum Akademickie”, 9.
- Wróblewski A.K., 2005, *Nauka w Polsce według rankingów bibliometrycznych*, „Nauka”, 2: 13–28.
- Wróblewski A.K., 2013, *Pozycja nauki polskiej w międzynarodowych rankingach*, „Studia BAS”, 3.35: 89–106.

UNESCO, 2014, *UNESCO Data Centre*, Dokument internetowy: http://data.uis.unesco.org/index.aspx?datasetcode=scn_ds [dostęp: 12.03.2016].

Życzkowski K., 2011, *Ile waży jedno cytowanie?*, „Forum Akademickie” 10/2011.