

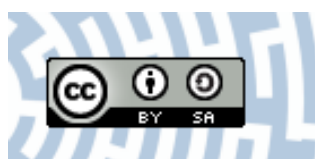


You have downloaded a document from
RE-BUŚ
repository of the University of Silesia in Katowice

Title: Potencjał bibliometryczny bibliograficznej bazy danych CYTBIN w świetle prostych i złożonych wskaźników analitycznych

Author: Anna Małgorzata Kamińska

Citation style: Kamińska Anna Małgorzata. (2017). Potencjał bibliometryczny bibliograficznej bazy danych CYTBIN w świetle prostych i złożonych wskaźników analitycznych. "Bibliotheca Nostra. Śląski Kwartalnik Naukowy" (2017, nr 2, s. 83-94).



Uznanie autorstwa - Na tych samych warunkach - Licencja ta pozwala na kopiowanie, zmienianie, rozprowadzanie, przedstawianie i wykonywanie utworu tak długo, jak tylko na utwory zależne będzie udzielana taka sama licencja.



POTENCJAŁ BIBLIOMETRYCZNY BIBLIOGRAFICZNEJ BAZY DANYCH CYTBIN W ŚWIETLE PROSTYCH I ZŁOŻONYCH WSKAŹNIKÓW ANALITYCZNYCH

Wstęp

Wśród bibliograficznych baz danych rejestrujących informacje o bibliografiach załącznikowych za najbardziej rozwinięte i dojrzałe uważa się te, które informacje o cytowanych jednostkach gromadzą i przechowują w postaci „maksymalnie ustrukturalizowanej”. W odróżnieniu od baz bibliograficznych, udostępniających informacje o bibliografii załącznikowej, np. w postaci skanowanych obrazów, mogą one realizować o wiele więcej zadań, wśród których wymienić należy:

- bogatsze możliwości wyszukiwawcze (realizowane również w kontekście jednostek cytowanych),
- usprawnienie mechanizmów wyszukiwania dla systemów klasy OPAC (poprzez możliwość podpowiadania jednostek podobnych),
- możliwość realizacji prostych i bardziej złożonych analiz bibliometrycznych.

Gromadzenie danych w postaci strukturalnej jest przy tym warunkiem koniecznym, ale niewystarczającym dla pełnego realizowania wymienionych zadań, co jest przedmiotem dalszych rozważań (Kamińska, 2017c).

Bibliograficzna baza CYTBIN, z założenia projektowana jako gromadząca szeroki zakres informacji bibliograficznych ustrukturalizowanych i specyficznych dla danych form wydawniczych, została utworzona z myślą o realizacji celów dydaktycznych. Miała umożliwić studentom kierunku informacji naukowej i bibliotekoznawstwa Uniwersytetu Śląskiego w Katowicach praktyczne uczestniczenie w tworzeniu opisów bibliograficznych. Obecnie gromadzi ona dane o artykułach publikowanych od 2000 r. w dziesięciu wybranych polskich czasopismach naukowych z zakresu bibliologii, informatologii oraz dyscyplin pokrewnych wraz ze wszystkimi źródłami cytowanymi w tych

¹ Uniwersytet Śląski w Katowicach, Instytutu Bibliotekoznawstwa i Informacji Naukowej

wydawnictwach. Od września 2005 r. została ona udostępniona w sieci Internet na zasadach praw publicznego dostępu, tak aby każda zainteresowana osoba miała możliwość pozyskania zgromadzonych w niej danych. Więcej informacji na temat genezy, stanu obecnego i możliwych kierunków rozwoju bazy CYTBIN można znaleźć w artykule badaczy bezpośrednio zaangażowanych w jej rozwój (Seweryn, Swoboda, 2013). Natomiast nakreślenie szerszego kontekstu również dla innych krajowych inicjatyw w opisywanej dziedzinie przedstawiono w (Seweryn, Swoboda, 2014).

Od rozpoczęcia przedsięwzięcia (w styczniu 2002 r.), wraz z przyrostem gromadzonych danych, zaczęto w zgromadzonych informacjach dostrzegać potencjał bibliometryczny, co dało początek realizacji badań naukowych z zakresu bibliometrii, opartych na stosunkowo prostych analizach bibliometrycznych (Drabek, Tomaszczyk, 2008; Seweryn, Swoboda, 2013). Wtedy też zaczęto sygnalizować pewne problemy natury technologicznej, organizacyjnej czy metodyczno-koncepcyjnej utrudniające pełne wykorzystanie potencjału systemu (najnowsze ich podsumowanie zob. Seweryn, Swoboda, 2013). Również inni badacze, analizujący m.in. bazę CYTBIN, zauważają, że kryzys funkcjonalny odnoszący się do krajowych inicjatyw bibliograficznych baz danych, stoi na przeszkodzie w realizacji bardziej zaawansowanych badań i umożliwia prowadzenie jedynie „prostych analiz” (Kolasa, 2011).

W tej sytuacji Autorka mając na uwadze dostępność danych bazy CYTBIN oraz jej potencjał bibliometryczny, sygnalizowany przez badaczy bezpośrednio zaangażowanych w jej rozwój (Seweryn, Swoboda, 2013), podjęła próbę transformacji udostępnionych publicznie danych do modeli umożliwiających realizację bardziej zaawansowanych analiz. Ich podstawą był wypracowany warsztat badawczy oraz wstępna analiza problemów odnoszących się do inicjatyw bibliograficznych baz danych (Kamińska, 2017c). Należy podkreślić, że celem niniejszego opracowania nie jest podjęcie tematu realizacji całościowych badań dotyczących rozwoju nauki w przedmiotowym obszarze, a jedynie weryfikacja hipotezy istnienia potencjału bibliometrycznego bazy CYTBIN. Wobec braku pewności odnośnie do jakości zgromadzonych danych, wnioski wyciągane z przeprowadzonych analiz należy traktować z dużą ostrożnością, podobnie zresztą jak każde inne wynikające z analiz ilościowych, zastosowanych do badań naukometrycznych.

Ponieważ publikowane do tej pory wyniki badań dotyczyły najczęściej analiz w podziale na artykuły czy czasopisma, to większość podjętych badań odnosi się do poziomu autorów i różnych typów relacji między nimi. Stanowi to dodatkowe wyzwanie, gdyż analizowanie tych relacji wymaga agregacji informacji dotyczących poszczególnych artykułów z wcześniejszym uwzględnieniem artykułów mających wielu autorów.

Analizy przeprowadzone zostały dla całego udostępnionego (na 6 grudnia 2016 r.) zakresu czasowego danych. Z jednej strony uniemożliwia to

zobrazowanie trendów wynikających z mniejszych przedziałów czasowych, czyli ogranicza możliwość prognozowania tych trendów. Z drugiej natomiast niektórzy badacze dziedzin humanistycznych zauważają (Kolasa, 2011), że czas „życia publikacji” (tzn. okresu od powstania publikacji do ostatniego jej cytowania) jest na tyle duży, że zawężanie okresów analiz może prowadzić do deformacji obrazu dziedziny.

Proste wskaźniki ilościowe

Analizy danych zamieszczonych w bazie umożliwiają uzyskanie informacji o różnego typu zależnościach, od prostych dotyczących np. liczby cytowań autorów i artykułów, współpracy między autorami czy wydawnictwami, poprzez bardziej zaawansowane, takie jak metoda współcytowań i powiązań bibliograficznych, aż po złożone wskaźniki oraz mapy bibliometryczne. Analiza zawartości danych pozwala na uzyskanie informacji np. o najczęściej cytowanych publikacjach i autorach lub ogólniejszych o cytowaniach pomiędzy czasopismami; można również ustalić, którzy autorzy i jak często cytują innych autorów.

Najczęściej cytowane publikacje to: *Słownik encyklopedyczny informacji, języków i systemów informacyjno-wyszukiwawczych* z 2002 roku (35 razy²), *Encyklopedia wiedzy o książce* z 1971 (21), *Słownik terminologiczny informacji naukowej* z 1979 (17) i „Ustawa z dnia 27 czerwca o bibliotekach” z 1997 (17). Okazało się, że wymieniona jako ostatnia ustawa w bazie występuje również jako „Ustawa z dnia 27 czerwca 1997 r. o bibliotekach” i jest cytowana 12 razy. Po scaleniu tych dwóch zapisów, byłaby drugą najczęściej cytowaną publikacją – 29 razy (o metodach przeciwdziałania i korygowania rekordów w takich sytuacjach zob. Kamińska, 2017a).

Najczęściej cytowanymi autorami³ w bazie są: Jacek Wojciechowski (84 razy), Marek Nahotko (62), Jadwiga Sadowska (62) i Barbara Sosińska-Kalata (53).

Analizując informacje zgromadzone w bazie, można również dowiedzieć się którzy autorzy i jak często cytują innych autorów. Najczęściej, bo 39 razy, Maciej Waltoś cytował Zdzisława Szeligę, Marzena Świgoń cytowała Anthon'ego J. Onwuegbuzie 33 razy, a Quina G. Jiao 29.

Badając zależności między wydawnictwami można również określić liczbę cytowań między czasopismami. W tym przypadku należy podkreślić, że ich obecność w bazie jest różna i wynika ze sposobu gromadzenia danych. Najczęściej, bo 222 razy cytowany był „Przegląd Biblioteczny” w „Przeglądzie Bibliotecznym”; 160 razy „Bibliotekarz” w „Bibliotekarzu”;

² Wszystkie zamieszczone w tej części artykułu informacje wynikają z danych obliczonych i zgromadzonych w tradycyjnych układach tabelarycznych za pomocą aplikacji arkusza kalkulacyjnego.

³ Najczęściej cytowani byli autorzy anonimowi – Anon 5 513 razy.

92 razy „Zagadnienia Informacji Naukowej” przez „Zagadnienia Informacji Naukowej”; a 50 razy „Przegląd Biblioteczny” cytował „Bibliotekarza”.

Metody współcytowań i powiązań bibliograficznych

Znane od początków powstawania bibliometrii metryki podobieństw dokumentów, nazwane metodami powiązań bibliograficznych (*bibliographic coupling*) oraz współcytowań (*co-citation*), są stosunkowo rzadko używane przez badaczy do realizacji badań bibliometrycznych. Prawdopodobnie wynika to z dość kłopotliwego sposobu ich wyliczania, zarówno jeśli chodzi o samą metodę (algorytm), jak i o nieproporcjonalnie duży narzut obliczeniowy w stosunku do liczby przetwarzanych rekordów. Trudności wpływają na brak praktycznych zastosowań tych metod w krajowych badaniach bibliograficznych i jak dotąd najpewniej sprowadzają się jedynie w rozważaniach teoretycznych (na temat metod i możliwości ich zastosowań zob. Marszakowa-Szajkiewicz, 2009).

Analizując dane z wykorzystaniem metody powiązań bibliograficznych, można wyróżnić pary dokumentów, w których ten sam dobór pozycji w bibliografii wskazuje na powiązania między samymi dokumentami. Najwięcej wspólnych pozycji w bibliografii – 37 – ma artykuł E. Mostowicz, A. Grzecznowskiej i M. Balińskiej *Informacja biznesowa w wybranych krajach Europy Wschodniej* opublikowany w „PTINT. Praktyka i Teoria Informacji Naukowej i Technicznej” w 2001 r., dwukrotnie wprowadzony do bazy pod różnymi identyfikatorami. Opisy różnią się dodatkowo wpisanym tytułem w języku angielskim, symbolem klasyfikacji tematycznej, słowami kluczowymi, abstraktami i co ciekawe liczbą pozycji w bibliografii (w pierwszej wersji jest ich 39, w drugiej 38). Podobna sytuacja wystąpiła w kolejnej parze dokumentów o 16 wspólnych pozycjach w bibliografii – artykuły Barbary Niedźwiedzkiej *Poszukujący informacji, czy też unikający tego zadania użytkownik* opublikowanym w „PTINT. Praktyka i Teoria Informacji Naukowej i Technicznej” w 2001 roku (liczba cytowań wymieniona w pierwszym dokumencie to 23 w drugim 26). Trzecia para dokumentów, o 15 wspólnych pozycjach w bibliografii, to dwa różne artykuły autorstwa Stanisława Skórki. Pierwszy *Strategie poszukiwania informacji w edukacyjnym systemie hipertekstowym* został opublikowany w „PTINT. Praktyka i Teoria Informacji Naukowej i Technicznej” w 2005 r. (liczba cytowań w bazie – 23), a drugi *Strategie poszukiwania informacji w edukacyjnym serwisie internetowym* również w 2005 r. w „Przeglądzie Bibliotecznym” (liczba cytowań wpisanych w bazie 31).

Wykorzystanie metody współcytowań pozwala na wyróżnienie pary dokumentów, których wspólne występowanie w bibliografii innych publikacji wskazuje na możliwość powiązania między samymi dokumentami. Najczęściej wspólnie występującymi w bibliografiach wydawnictwami (dzie-

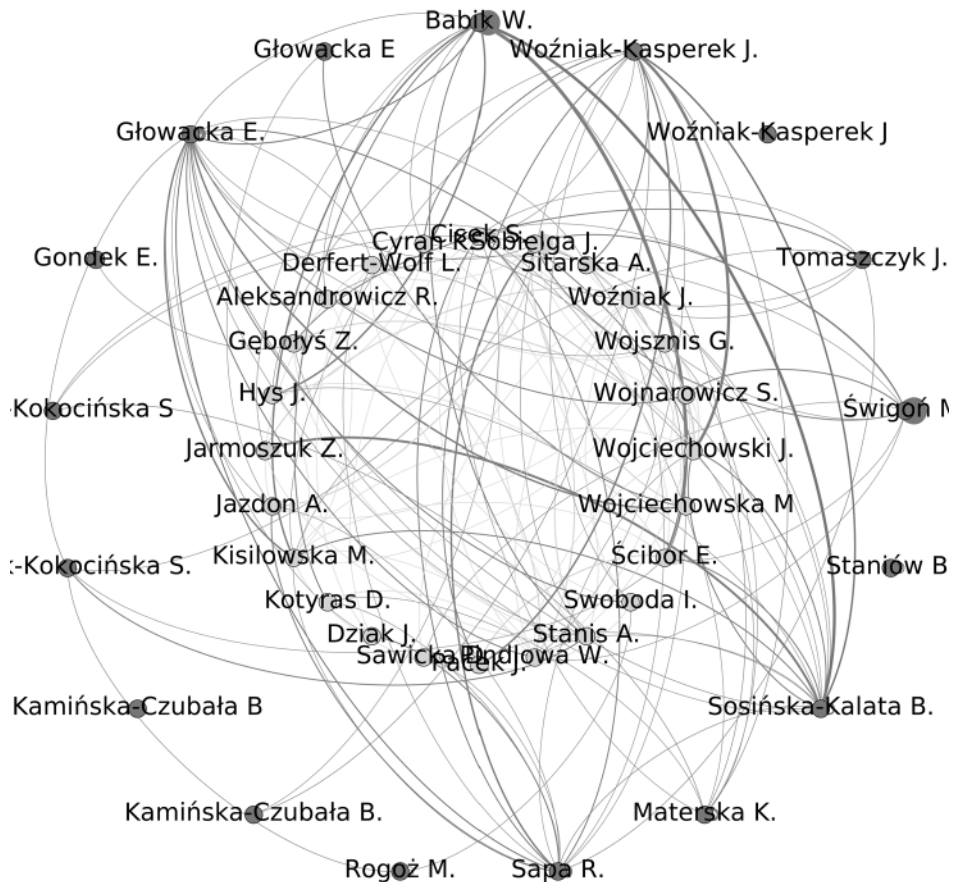
więć razy) są *Słownik terminologiczny informacji naukowej* z 1979 r. i *Słownik encyklopedyczny informacji, języków i systemów informacyjno-wyszukiwawczych* z 2002 r. Kolejne pary dokumentów były cytowane wspólnie po pięć razy, przykładowo wybrane pary to: *Słownik encyklopedyczny terminologii języków i systemów informacyjno-wyszukiwawczych* z 1993 r. i *Słownik terminologiczny informacji naukowej* z 1979 r.; *Słownik terminologiczny informacji naukowej* z 1979 r. i *Tezaurus informacji naukowej* z 1992 r.; *Encyklopedia wiedzy o książce* z 1971 r. i *Encyklopedia współczesnego bibliotekarstwa polskiego* z 1976 r.

Złożone wskaźniki oraz mapy bibliometryczne

Ciekawym podejściem do badań bibliometrycznych jest postrzeganie danych bibliograficznych jako sieciowej struktury, przedstawiającej ścieżki cytowań między poszczególnymi jednostkami bibliograficznymi czy poszczególnymi autorami oraz relacje autorstwa pomiędzy naukowcami a ich artykułami. Obserwacja ta umożliwia analizowanie zależności w takich strukturach danych za pomocą metod analiz społecznościowych (SNA, *social network analysis*), (na temat sieciowych struktur danych oraz propozycji interpretacji poszczególnych miar SNA w bibliometrii zob. Kamińska, 2018a; Kamińska, 2018b).

Na ilustracji 1 przedstawiono informację pochodzącą z bazy CYTBIN o osobach będących członkami Rady programowej I Ogólnopolskiej Konferencji Naukowej *Organizacja i recepcja treści w środowisku cyfrowym*, która odbyła się 23 maja 2017 r. w Katowicach. Chodziło o ukazanie relacji między nimi a osobami je cytującymi. Na zewnętrznej orbicie zostały rozmieszczone wierzchołki z nazwiskami osób, wraz z inicjałami imienia, będących członkami Rady programowej konferencji. W wewnętrznym okręgu wymieniono nazwiska osób będących w relacji z członkami Rady programowej. Łuki pomiędzy poszczególnymi osobami świadczą o istniejącej relacji. Warto zwrócić uwagę na powtarzające się osoby w zewnętrznym okręgu. Sytuacja taka powstała w wyniku różnego zapisu nazwiska i inicjału imienia (raz z kropką po inicjale, innym razem bez kropki) i doprowadziła do powielenia autorów i rozdzielenia powiązań pomiędzy wykreowanymi bytami. Aby poprawnie zaznaczyć łuki relacji, należałoby początkowo scalić podwójne byty do właściwej, ujednocnionej formy.

Następnie (ilustracja 2) przedstawiono mapę cytowań między autorami bez autocytowań z nieważonym stopniem wchodzącym, wyliczonym na podstawie bazy, gdzie wierzchołki oznaczają autorów, a łuki między nimi wskazują na cytowanie. Cytowanie nie jest relacją zwrotną, dlatego kierunek łuku, zgodnie z ruchem wskazówek zegara, prowadzi od autora cytującego do autora cytowanego. Nieważony stopień wchodzący oznacza, że relacja cytowania zaistniała (nie uwzględniana jest liczba cytowań między konkretną

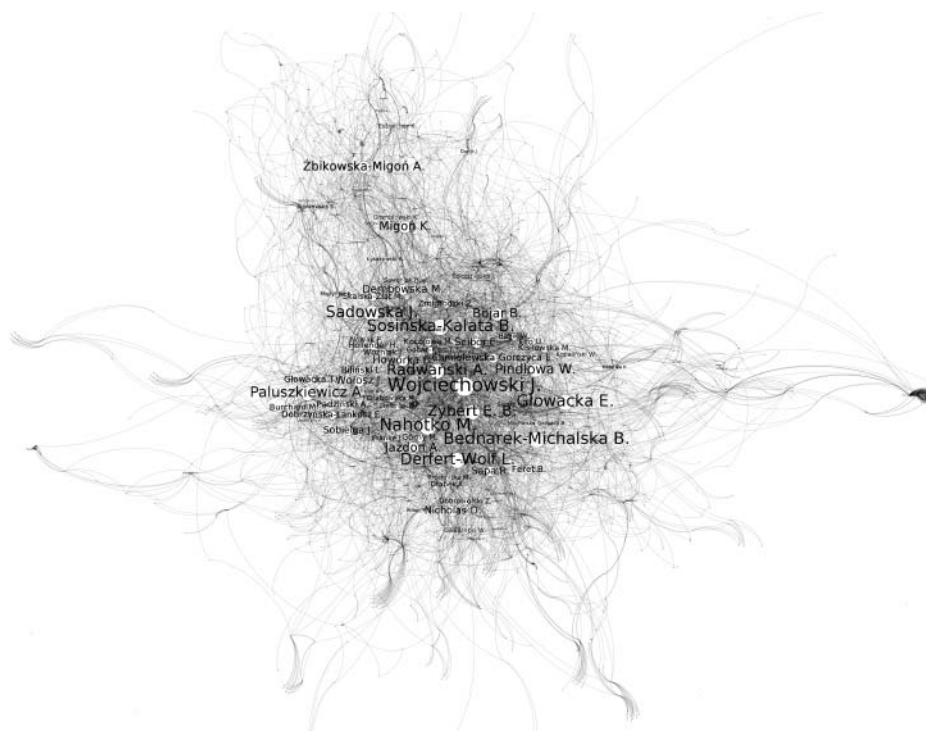


II. 1. Mapa cytowań członków Rady programowej konferencji. Źródło: opracowanie własne, dostępne w: https://zenodo.org/record/1069777/files/Rys_1.png

parą autorów, a jedynie jej fakt) i w tym przypadku wskazuje na autorów najczęściej cytowanych. Im autor jest częściej cytowany przez innych autorów, tym więcej łuków wskazuje na wierzchołek go reprezentujący.

Na ilustracji 3 przedstawiono mapę cytowań pomiędzy autorami bez autocytowań z nieważonym stopniem wychodzącym, wyliczonym na podstawie danych z bazy, gdzie wierzchołki oznaczają autorów, a łuki między

⁴ Wszystkie ilustracje zamieszczone w artykule udostępniono jako dane badawcze pod wspólnym identyfikatorem DOI:10.5281/zenodo.1069777. Dzięki temu można przeanalizować grafy w wysokiej rozdzielczości i w oryginalnych barwach.

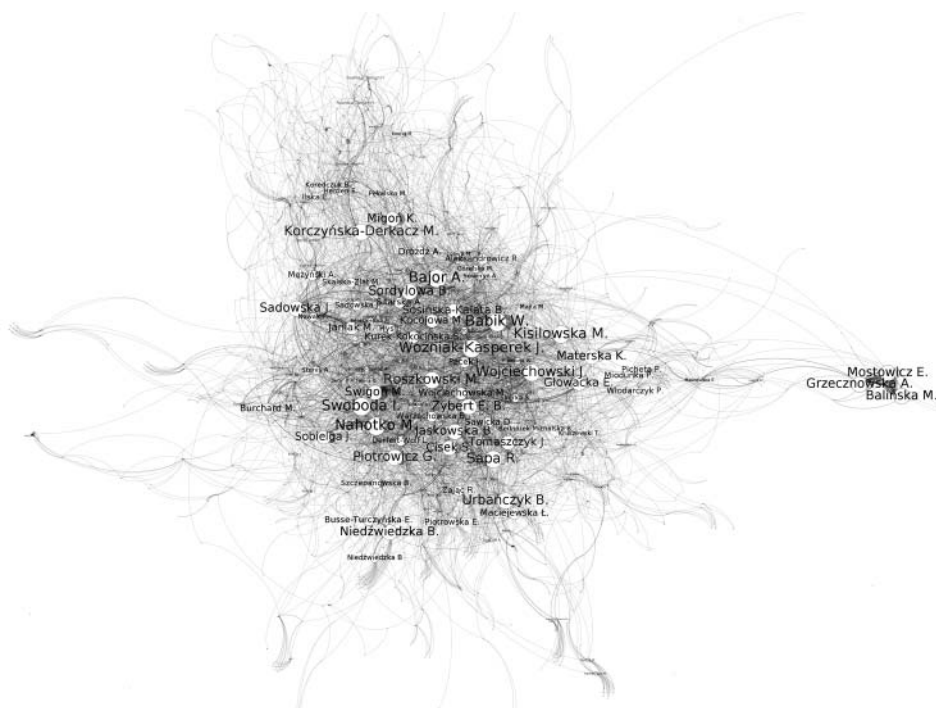


II 2. Mapa cytowań – nieważony stopień wchodzący bez autocytowań. Źródło: opracowanie własne, dostępne w: https://zenodo.org/record/1069777/files/Rys_2.png

nimi wskazują na cytowania. Kierunek łuku, zgodnie z ruchem wskazówek zegara, prowadzi od autora cytującego do autora cytowanego. Nieważony stopień wychodzący oznacza, że relacja cytowania zaistniała i w tym przypadku wskazuje na autorów, którzy najczęściej cytują innych. Im autor częściej cytuje różnych innych, tym więcej łuków wychodzi z wierzchołka go reprezentującego i tym większy jest punkt reprezentujący danych wierzchołek.

Na ilustracji 4 również przedstawiono cytowania. Rozmiar wierzchołka określa indeks Hirscha danego autora, wyliczony na podstawie zawartości bazy. Jest to miara całkowitoliczbowa, a maksymalna jej wartość w tym przypadku wynosi cztery, zatem możemy obserwować wierzchołki o trzech rozmiarach. Wartość Hirscha cztery wystąpiła tylko dla dwóch największych wierzchołków (J. Sadowska i J. Wojciechowski).

Cytowania między czasopismami, przedstawione w części artykułu dotyczącym analiz ilościowych, mogą również zostać zobrazowane za pomocą struktury sieciowej. Sposób pokazania tych relacji (cytowania



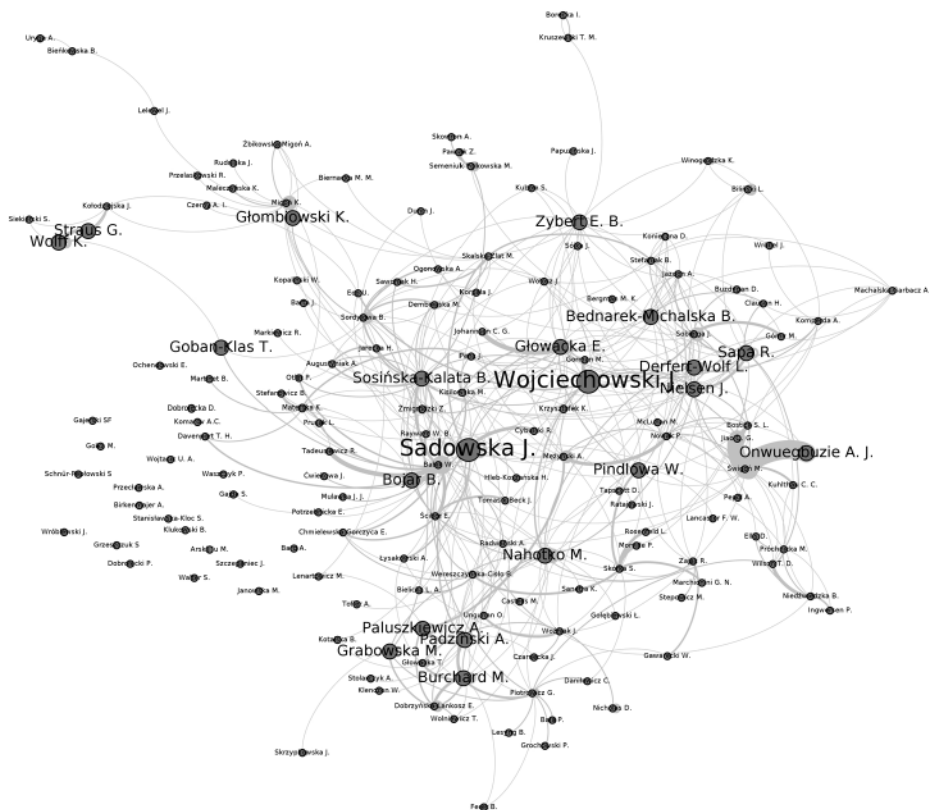
II. 3. Mapa cytowań – nieważony stopień wychodzący bez autocytaowań. Źródło: opracowanie własne, dostępne w: https://zenodo.org/record/1069777/files/Rys_3.png

między czasopismami razem z autocytaowaniami) zaprezentowano na ilustracji 5. Wierzchołki reprezentują poszczególne tytuły czasopism, łuki cytowania. Im grubszy łuk, tym większa liczba cytowań. Pogrubione łuki przy konkretnych tytułach obrazują autocytaowania. Wielkość wierzchołka wskazuje na stopień wchodzący, czyli rozmiar jest proporcjonalny do liczby otrzymanych cytowań.

Więcej wizualizacji wyników analiz bibliometrycznych dla bazy CYTBIN przedstawiono w (Kamińska, 2017e), natomiast dla bibliograficznej bazy danych GRUBA stworzonej przez autorkę w (Kamińska, 2017b).

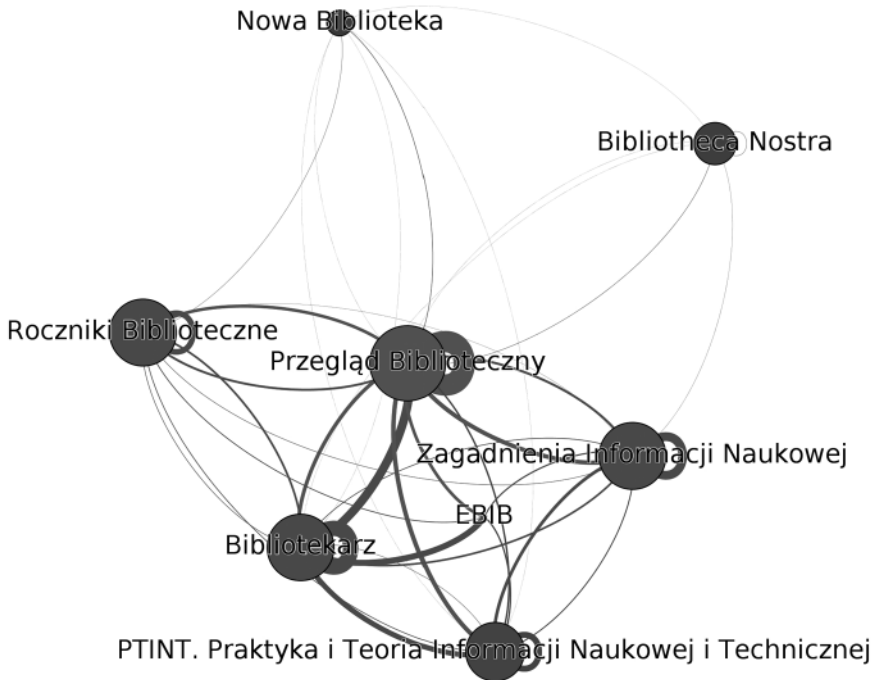
Podsumowanie

Przedstawione badania wyraźnie ukazują duży potencjał bibliometryczny bazy CYTBIN. Po odpowiednim przekształceniu danych bibliograficznych możliwe okazało się zrealizowanie zarówno prostych, jak i bardziej zaawansowanych analiz. Dzięki niektórym miarom podobieństw dokumentów (metoda powiązań bibliograficznych) ujawniono występujące anoma-



Il. 4. Cytowania autorów posiadających wskaźnik Hirsch większy niż 1. Źródło: opracowanie własne, dostępne w: https://zenodo.org/record/1069777/files/Rys_4.png

lie w postaci rekordów zduplikowanych, co mogło mieć oczywisty wpływ na zaburzenie uzyskanych wyników. Jednak nie zmienia to przekonania o możliwości stosowania zaawansowanego aparatu analitycznego do wolumenu zgromadzonych danych. W celu uwiarygodnienia informacji zgromadzonych w bazie oraz realizacji koniecznych deduplikacji najkorzystniej byłoby zastosować metody automatyczne, ograniczające do minimum potrzebę interakcji ze strony operatora (zob. Kamińska, 2017a). Natomiast reorganizując dane (zob. Kamińska 2017c) można uzyskać korzyści bezpośrednio wpływające zarówno na jakość gromadzonych danych, wydajność tego procesu, jak i możliwość realizowania bardziej zaawansowanych analiz bezpośrednio na źródłowych strukturach danych, bez potrzeby ich transformacji do postaci bardziej sprzyjającej tym analizom (jak miało to



II. 5. Cytowania pomiędzy czasopismami z autocytoowaniami. Źródło: opracowanie własne, dostępne w: https://zenodo.org/record/1069777/files/Rys_5.png

miejsce przy realizacji niniejszych badań). Należy dodać, że baza CYTBIN powstawała w czasach, kiedy jeszcze nie było gotowych narzędzi, które można by wykorzystać do jej zasilania (zob. Kamińska, 2017b), a sposób organizowania tych danych nie podlegał jeszcze współczesnym standaryzacji np. w postaci ontologii bibliograficznych. Dane zgromadzone podczas całego okresu funkcjonowania bazy CYTBIN są cenne, a inicjatywa ich gromadzenia i uaktualniania powinna być rozwijana i kontynuowana.

Bibliografia

Drabek, A. (2009). Konferencja Bibliograficzne Bazy Danych: Kierunki rozwoju i możliwości współpracy. Bydgoszcz 27–29 maja 2009 roku. *Bibliotheca Nostra*, 2(18), 116–121.

Drabek, A., Tomaszczyk, J. (2008). Czasopismo „Praktyka i Teoria Informacji Naukowej i Technicznej” w świetle danych bazy CYTBIN. W: Pietruch-Reizes, D. (red.), *Zarządzanie informacją w nauce*

- (s. 365–375). Katowice: Wydawnictwo Uniwersytetu Śląskiego.
- Garczyńska, M. (2013). Bibliograficzne bazy danych w ocenie pracowników i jednostek szkół wyższych w Polsce. W: *Bibliograficzne bazy danych i ich rola w rozwoju nauki, Poznań, 17-19 kwietnia 2013. II Konferencja naukowa Konsorcjum BazTech*. Pobrane 7 września 2017, z: http://open.ebib.pl/ojs/index.php/Mat_konf/article/view/24/164
- Kamińska, A.M. (2017a). Miary podobieństw łańcuchów znakowych a deduplikacja rekordów w bibliograficznych bazach danych. *Przegląd Biblioteczny*, 4 [w druku].
- Kamińska, A.M. (2017b). Od druków źródłowych po mapy nauki. Bibliograficzna baza danych GRUBA. W: Kowalska M., Osińska V. (red.). *Wizualizacja informacji w humanistyce*. (s. 17–36). Toruń: Wydawnictwo Uniwersytetu Mikołaja Kopernika.
- Kamińska, A.M. (2017c). ProBIT – perspektywna metoda tworzenia trawersowalnych indeksów cytowań a współczesne problemy organizacji przestrzeni informacji w tradycyjnych bibliograficznych bazach danych. *Zagadnienia Informatyki Naukowej*, 1(109), 66–82.
- Kamińska, A.M. (2017d). Tam, gdzie zaczyna się bibliometria, czyli jak pozyskać materiał analityczny z autopsji. *Biuletyn EBIB*, 3(173). Pobrane 7 września 2017, z: <http://open.ebib.pl/ojs/index.php/ebib/article/view/534/698>
- Kamińska, A.M. (2017e). Wizualizacje wybranych wskaźników bibliometrycznych na przykładzie bibliograficznej bazy danych CYTBIN. *Toruńskie Studia Bibliologiczne*, 2(19) [w druku].
- Kamińska, A.M. (2018a). Zastosowanie metod analizy sieci społecznościowych w dziedzinach bibliometrii i webometrii. Miary i narzędzia. *Nowa Biblioteka. Usługi, technologie informacyjne i media*, 2(29) [w druku].
- Kamińska, A.M. (2018b). *Zastosowanie struktur grafowych do analiz bibliometrycznych i webometrycznych. Modele i metody* [w druku].
- Kolasa, W.M. (2011). Retrospektywny indeks cytowań w humanistyce. Koncepcja, metoda, zastosowanie. *Przegląd Biblioteczny*, 79(4), 466–486.
- Marszakowa-Szajkiewicz, I. (2009). *Badania ilościowe nauki. Podejście bibliometryczne i webometryczne*. Poznań: Uniwersytet im. Adama Mickiewicza.
- Seweryn, A., Swoboda, I. (2014). Cytowania w polskich bazach bibliograficznych. *PTINT Praktyka i Teoria Informatyki Naukowej i Technicznej*, 1(85), 3–20.
- Seweryn, A., Swoboda, I. (2013). CYTBIN – krajowy indeks cytowań z zakresu bibliologii i informatologii. W: *Bibliograficzne bazy danych i ich rola w rozwoju nauki, Poznań, 17-19 kwietnia 2013 : II Konferencja naukowa Konsorcjum BazTech*. Pobrane 7 września 2017, z: http://open.ebib.pl/ojs/index.php/Mat_konf/article/view/42/150

Anna Małgorzata Kamińska

The bibliometric potential of the bibliographic database CYTBIN in the light of simple and more complex bibliometric indicators

Abstract

In this paper the author tries to verify whether data collected in the bibliographic database CYTBIN enable basic and more complex bibliometric analyses. After transformation of data collected from public website, the author conducts several bibliometric analyses, from quite simple study of citations between authors to co-citation and bibliography coupling methods and visualizations of obtained results. In few cases some anomalies of duplicated records were detected (both for bibliographic units and authors). However, taking into account that these anomalies are easy to detect and correct by automatic methods we may say that possibilities of conducting bibliometric research on the CYTBIN database are great.

Key words: CYTBIN, citations analyses, bibliometrics, indicators, co-citation method, bibliography coupling method

Anna Małgorzata Kamińska

Potencjał bibliometryczny bibliograficznej bazy danych CYTBIN w świetle prostych i złożonych wskaźników analitycznych

Streszczenie

W artykule podjęto próbę weryfikacji potencjału analitycznego danych zgromadzonych w bibliograficznej bazie danych CYTBIN. Po przekształceniu danych udostępnionych publicznie z poziomu aplikacji WWW przeprowadzono wiele analiz bibliometrycznych – od stosunkowo prostych analiz cytowań autorów, poprzez analizę bardziej złożonych miar wyliczanych za pomocą metody współcytowań czy powiązań bibliograficznych po wizualizacje wyliczonych wcześniej miar zarówno dla danych detalicznych, jak i zagregowanych do poziomu czasopism naukowych. W niektórych przypadkach w trakcie realizacji badań zaobserwowano pewne anomalie, polegające w głównej mierze na duplikacji zgromadzonych opisów bibliograficznych dotyczących tak jednostek bibliograficznych, jak i ich autorów. Biorąc jednak pod uwagę, że anomalie te są możliwe do wykrycia i usunięcia metodami automatycznymi, można potwierdzić, że po odpowiedniej reorganizacji danych, możliwości prowadzenia analiz bibliometrycznych na podstawie bazy CYTBIN są wysokie.

Słowa kluczowe: CYTBIN, analiza cytowań, bibliometria, wskaźniki, metoda współcytowań, metoda powiązań bibliograficznych