

# Wykorzystanie Google Scholar do identyfikowania najczęściej cytowanych badaczy i ich prac naukowych. Przypadek publikacji z zakresu nauki o informacji w języku polskim

Zbigniew Osiński

ORCID: 0000-0003-4484-7265

*Institut Nauk o Kulturze, Wydział Humanistyczny  
Uniwersytet Marii Curie-Skłodowskiej w Lublinie*

---

## Abstrakt

**Cel/Teza:** Ustalenie, czy Google Scholar (GS) umożliwia wiarygodną identyfikację najczęściej cytowanych badaczy i ich prac naukowych opublikowanych w języku polskim. Określenie, czy liczba cytowań i stosowanie słów kluczowych są istotnymi kryteriami hierarchizacji wyników wyszukiwania w GS. Stworzenie rankingu najczęściej cytowanych polskich badaczy i prac z zakresu nauki o informacji w okresie 2010–2019. Weryfikacja hipotezy, sugerującej, że zarówno obecność danej pracy w wynikach wyszukiwania, jak i pozycję na liście tych wyników, zaburza problem słów kluczowych tej publikacji, np. ich brak.

**Koncepcja/Metody badań:** Weryfikacja i analiza ilościowa danych pobranych z Google Scholar. Próbę badawczą wyodrębniono za pomocą czterdziestu pytań wyszukiwawczych, którymi były wyrażenia kluczowe dla nauki o informacji oraz dodatkowo imiona i nazwiska badaczy z tej dyscypliny. **Wyniki i wnioski:** W przypadku haseł wyszukiwawczych w języku polskim nie znajduje potwierdzenia teza mówiąca, iż wyszukiwarka Google Scholar nadaje się do niezawodnej identyfikacji najczęściej cytowanych prac i badaczy. Znacząca statystycznie część prac indeksowanych w GS nie pojawia się w wynikach wyszukiwania przy zastosowaniu wyrażen związanych z nauką o informacji. Duża grupa często cytowanych tekstów umieszczona jest na bardzo dalekich miejscach w wynikach wyszukiwania. Nie znajduje także potwierdzenia teza mówiąca, że głównym kryterium szeregowania prac naukowych w wynikach wyszukiwania jest liczba cytowań. Zdecydowana większość pozycji bibliograficznych w wynikach wyszukiwania GS nie jest adekwatna do intencji wyszukującego. Właściwe zastosowanie słów kluczowych w metadanych publikacji nie zwiększa jej szansy na wyszukanie w Google Scholar za pomocą polskich terminów wyszukiwawczych.

**Oryginalność/Wartość poznawcza:** Praca obala coraz bardziej popularne przekonanie o tym, że Google Scholar można profesjonalnie wykorzystywać do wyszukiwania najbardziej wartościowych (a raczej stosunkowo często cytowanych) prac naukowych oraz do formułowania ocen (bazujących na cytowaniach) tychże opublikowanych w języku polskim i badaczy publikujących głównie w tym języku.

## Słowa kluczowe

Cytowania pracy naukowej. Google Scholar. Nauka o informacji. Ranking wyników wyszukiwania.

*Otrzymano: 28 stycznia 2021. Zrecenzowano: 18 kwietnia 2021. Poprawiono: 28 kwietnia 2021.*

*Zaakceptowano: 13 maja 2021.*

## 1. Wprowadzenie

Od ponad dekady w polskiej humanistyce i naukach społecznych coraz większy wpływ zdobywa pogląd mówiący, że liczba cytowań uzyskanych przez badacza lub pracę naukową świadczy o sile wpływu na rozwój nauki, a pośrednio także o naukowej jakości tegoż badacza lub pracy. Pogląd ten początkowo, w drugiej połowie XX w., rozpowszechnił się w naukach ścisłych, przyrodniczych, o zdrowiu i inżynierjno-technicznych. We wspomnianych dziedzinach przyjęło się przekonanie, że dla wyrobienia sobie opinii o badaczu i jego pracach istotne jest zliczanie cytowań w publikacjach zarejestrowanych w bibliograficznej bazie danych Web of Science (a ostatnio także Scopus). Obecność w tej (tych) bazie (bazach) uznawana jest bowiem za gwarancję wysokiej jakości naukowej czasopisma, a pośrednio także pracy w nim opublikowanej. Jednakże ten system nie sprawdza się w przypadku większości dyscyplin z nauk humanistycznych i społecznych uprawianych w Polsce. Powodem są zarówno różnice w zwyczajach publikacyjnych pomiędzy badaczami z poszczególnych dyscyplin naukowych, jak i fakt, że czasopisma z nauk humanistycznych i społecznych, które wydawane są w innych językach niż angielski, przez wydawców innych niż międzynarodowe koncerny wydawnicze, nie posiadają odpowiednio licznej reprezentacji w obu bazach, pozwalającej na racjonalne i obiektywne wyliczenia oraz analizy (Osiński, 2019).

Nie stworzono jeszcze odpowiednio zasobnych baz cytowań bibliograficznych, które zapewniłyby wiarygodne wyniki automatycznego zliczania cytowań prac polskich badaczy z nauk humanistycznych i społecznych, toteż pojawia się istotny problem: w jaki sposób pozyskać dane pozwalające w miarę precyzyjnie określić, jak dużo cytowań zdobyły prace poszczególnych polskich badaczy? Rozwiązanie tego problemu jest ważne dla wszystkich prowadzących badania naukometyczne i bibliometryczne oraz dla tych, którzy, przygotowując projekt badawczy, poszukują najważniejszych (a więc najbardziej wpływowych, czyli najczęściej cytowanych) prac na dany temat. Do tego należy uwzględnić fakt, że od ponad dekady upowszechniają się poglądy, których autorzy oceniają pozycję polskiej nauki, uczelni, instytutów i badaczy na podstawie danych dotyczących cytowań (np. Siłka et al., 2016; Śleszyński, 2013; Wróblewski, 2013). Ponadto aktualny system oceny parametrycznej dyscyplin naukowych w kryterium pierwszym opiera się na rankingu czasopism, który w dużym stopniu uwzględnia cytowania zdobywane przez artykuły opublikowane w danym periodyku. Tak więc liczba cytowań staje się informacją kluczową dla naukowych badań i karier. Czy receptą na braki w zakresie informacji o cytowaniach może być wykorzystanie Google Scholar (GS), wyszukiwarki prac naukowych uznanej kilka lat temu za „najbardziej reprezentatywną aproksymację danych o cytowalności z obszaru nauk społeczno-humanistycznych” (Racki & Drabek, 2013)?

## 2. Stan badań

W literaturze naukowej z ostatnich lat pojawił się pogląd sugerujący, że do niezawodnej identyfikacji najczęściej cytowanych badaczy i ich prac wystarczająco dobrze nadaje się wyszukiwarka Google Scholar, która indeksuje literaturę naukową w szerokim zakresie dyscyplin, typów dokumentów i języków (m.in. Beel et al., 2010; Delgado López-Cózar et al., 2019; Martin-Martin et al., 2017; Stribu et al., 2015). GS specjalizuje się w wyszukiwaniu

i identyfikacji pełnych tekstów prac naukowych, danych bibliograficznych o charakterze naukowym, a także świadczy szereg usług, takich jak np. zliczanie cytatów otrzymanych przez każdy dokument. Stwierdzono, że wyniki wyszukiwania w GS w dużym stopniu pokrywają się z wynikami kwerend w komercyjnych bazach danych bibliograficznych. Ponadto GS zapewnia liczne informacje o pracach naukowych, które nie są zarejestrowane w Web of Science i Scopus, co może być przydatne dla badaczy w uzyskaniu pełniejszego obrazu konkretnej problematyki badawczej. Wyszukiwarka ta została określona mianem lidera pod względem liczby wyników bez względu na zastosowane terminy wyszukiwawcze (Stribu et al., 2015).

Ustalono, że co prawda funkcjonowanie algorytmów, za pomocą których odsiewane są treści naukowe od nienaukowych oraz tworzona jest posortowana lista wyników, stanowi tajemnicę handlową firmy Google, to jednak możliwe jest opracowanie zasad optymalizacji pod kątem uzyskania przez daną pracę miejsca na pierwszych stronach wyników. Na miejsce w wynikach wyszukiwania wpływa bowiem częstotliwość występowania w tekście wyrażenia, które zostało użyte jako wyrażenie wyszukiwawcze oraz liczba cytowań tego tekstu. Istotne są także szczegóły techniczne pliku PDF takie, jak: wpisanie właściwych metadanych do metryczki pliku, zamieszczenie na początku tekstu danych autorów, tytułu pracy i abstraktu, stosowanie stylów akapitowych dla śródtytułów, stosowanie szczegółowych słów kluczowych (raczej terminów wielowyrazowych niż pojedynczych wyrazów) i częste ich używanie (a także ich synonimów) we właściwych miejscach nie tylko w samym tekście: w: metadanych, tytule, śródtytułach, nazwach tabel i grafik oraz zachowanie schematu tekstu właściwego dla pracy naukowej – wstęp, stan badań, wyniki, wnioski, bibliografia (Beel et al., 2010).

Za istotną zaletę tej wyszukiwarki uznano uwzględnianie takich źródeł artykułów, które nie występują w Web of Science i Scopus, co pozwala uzyskać szersze spojrzenie na dorobek naukowy wielu badaczy, zwłaszcza w naukach humanistycznych i społecznych (Delgado López-Cózar et al., 2019). Inne badania sugerują, że co prawda GS systematycznie rozszerza zakres indeksowanych dokumentów naukowych, to jednak korzystanie z wyników wyszukiwania powinna cechować duża ostrożność. Mechanizmy tej wyszukiwarki nie odróżniają bowiem prac recenzowanych od nierecenzowanych i są podatne na manipulowanie cytowaniami (Halevi et al., 2017). Ustalono też, że liczba cytowań jest kluczowym czynnikiem w rankingu wyników: istnieje silny związek pomiędzy cytowaniami dokumentu a jego pozycją w wynikach wyszukiwania, a zatem GS jest w stanie skutecznie identyfikować wysoko cytowane prace (Martin-Martin et al., 2017). Wyrażono przy tym przypuszczenie, że czynnikiem zaburzającym związek pomiędzy liczbą cytowań danej pracy, a miejscem tej pracy w wynikach wyszukiwania jest język publikacji. Okazało się bowiem, że w przypadku gdy wysła się tzw. zerowe zapytanie (w wyszukiwaniu zaawansowanym określa się jedynie rok wydania poszukiwanych publikacji, zaś pole do wpisywania wyrażen wyszukiwawczych pozostaje puste), to publikacje w językach innych niż angielski GS umieszcza na odległych pozycjach, nawet jeżeli są stosunkowo często cytowane. Problem ten może istotnie zaburzać wyniki wyszukiwania w stosunku do prac nieanglojęzycznych, bowiem GS udostępnia jedynie pierwszy tysiąc rekordów na każde zapytanie (Martin-Martin et al., 2017). Stąd też, chcąc odpowiedzieć na pytanie, czy GS nadaje się do identyfikowania najczęściej cytowanych prac naukowych i badaczy także w języku polskim, należy przeprowadzić odpowiednie badania, ale z ograniczeniem do stawiania pytań wyszukiwawczych jedynie w języku polskim.

W przypadku wykorzystywania GS do badania dorobku badaczy z zakresu nauki o informacji (ang. *information science*) dotychczas ustalono, że możliwe jest stosowanie informacji zawartych w profilach badaczy z tej dyscypliny w Google Scholar Citations do wieloaspektowej analizy tej dyscypliny i do dokładnego zobrazowania społeczności naukowców. Jednakże konieczne jest ręczne przetworzenie pobranych informacji – uzupełnienie brakujących danych i usunięcie duplikatów (Martin-Martin et al., 2018).

### 3. Cele i metodologia badań

W związku z powyższym autor przeprowadził badania, które miały na celu ustalenie, czy Google Scholar umożliwia na tyle wiarygodną i kompletną identyfikację najczęściej cytowanych badaczy i prac naukowych opublikowanych w języku polskim, by jakość uzyskanych danych pozwalała na ich naukowe zastosowanie. Istotne było też ustalenie, czy liczba cytowań i stosowanie określonych słów kluczowych w języku polskim, przez autorów i wydawców, wpływa na sortowanie wyników wyszukiwania w GS (co sugerują badania opisane powyżej). Ponadto, poddano weryfikacji hipotezę sugerującą, że zarówno obecność danej pracy w wynikach wyszukiwania, jak i pozycję na liście tych wyników, zaburza problem słów kluczowych publikacji, np. ich brak.

Badania ograniczono do prac i autorów związanych z nauką o informacji (informatologią), która to dyscyplina według formalnego podziału nauk i dyscyplin naukowych w Polsce od 2019 r. należy do zakresu nauk o komunikacji społecznej i mediach. Przyjęte ograniczenie wiąże się z drugim celem badań, którym był zamiar stworzenia rankingu polskich badaczy i prac z tej dyscypliny (tym razem bez ograniczenia do języka polskiego), którego jedynym kryterium byłaby cytawalność. Zakres czasowy pobieranych danych ograniczono więc do okresu 2010–2019. Badanie zrealizowano w trzech etapach opisanych poniżej.

#### 3.1. Etap 1. Gromadzenie materiału badawczego

W pole wyszukiwawcze wpisano czterdzieści wyrażań, którymi były terminy uznane za kluczowe dla nauki o informacji (wzięte w cudzysłów). Autor zastosował następujące kryteria wyboru tych terminów: dwie wersje nazwy dyscypliny: nauka o informacji, informatologia; dziesięć terminów najczęściej występujących jako słowa kluczowe w artykułach opublikowanych w latach 2010–2018 w czasopiśmie związanych z badaną dyscypliną, obecnych na dawnej (istniejącej do 2018 r.) ministerialnej liście B (Osiński, 2019): zarządzanie informacją, biblioteki akademickie, kompetencje informacyjne, komunikacja naukowa, język haseł przedmiotowych, rozwój dyscyplin naukowych, wyszukiwanie informacji, języki informacyjno-wyszukiwawcze, metody ilościowe w informatologii, czasopisma naukowe. Kolejne dwadzieścia osiem terminów autor wybrał na podstawie analizy artykułu Barbary Sosińskiej-Kalaty *Obszary badań współczesnej informatologii (nauki o informacji)* (2013) oraz książki *Nauka o informacji*, wydanej pod redakcją Wiesława Babika (2016) jako nazwy podstawowych obszarów badań nauki o informacji: teoria informacji, architektura informacji, procesy informacyjne, systemy informacyjne, zasoby informacyjne, ekologia informacji, bibliometria, bariery informacyjne, użytkownicy informacji, etyka informacyjna, źródła informacji naukowej, środowisko informacyjne, organizacja informacji, zachowania

informacyjne, kultura informacyjna, bibliografia załącznikowa, wizualizacja informacji, bibliograficzna baza danych, działalność informacyjna, naukometria, infobrokerstwo, czasopisma open access, biblioteki cyfrowe, bezpieczeństwo informacyjne, reprezentacja informacji, potrzeby informacyjne, obiekty informacyjne, przestrzeń informacyjna.

Z uzyskanych rezultatów wykluczono patenty oraz prace, które, ze względu na tematykę oraz miejsce publikacji (czasopismo/wydawnictwo) lub autora, związane były z takimi dyscyplinami jak: informatyka, nauki o zarządzaniu, ekonomia, medioznawstwo, pedagogika, nauki o administracji. Przyjęto założenie, że w przypadku niemożności w miarę jednoznacznego skojarzenia tematyki pracy z nauką o informacji, o zakwalifikowaniu do badań zadecyduje współwystępowanie autora lub czasopisma/wydawnictwa jednoznacznie związanego, według informacji dostępnych na stronach WWW poszczególnych uczelni oraz wiedzy autora niniejszego artykułu, z tą dyscypliną naukową. Do dalszych analiz nie zakwalifikowano też materiałów, które nie miały charakteru publikacji naukowej (brak elementów typowych dla takiej publikacji, np. bibliografii lub właściwej struktury). Ostatecznie materiał badawczy stanowiło 259 prac spełniających opisane powyżej kryteria (plik z danymi i rankingiem cytowań tych prac dostępny w RepOD – <https://repod.icm.edu.pl>).

Konieczna okazała się weryfikacja cytowań, albowiem GS do cytowań danej pracy naukowej zalicza także: cytowania pochodzące z dostępnych w Internecie prezentacji, prospektów reklamowych, nieopublikowanych wystąpień konferencyjnych, programów kształcenia oraz sylabusów przedmiotów, duplikaty cytowań, a także autocytowania. Wśród cytowań trafiają się również takie, które nie prowadzą do jakiegokolwiek pracy cytującej. Liczba cytowań zaburzających obiektywny obraz cytowalności prac związanych z danym wyrażeniem wyszukiwawczym przekracza w niektórych przypadkach nawet 30% wszystkich cytowań wykazanych przez GS (dla wyrażeń wyszukiwawczych: bibliometria, użytkownicy informacji, bibliografia załącznikowa, naukometria, obiekty informacyjne). W przypadku połowy wyrażeń odsetek takich cytowań przekracza 20%.

Tak więc liczbę cytowań podawaną przez GS zweryfikowano poprzez analizę każdego rekordu. Pomniejszono ją o autocytowania, powtórzenia cytowań tej samej pracy, cytowania pochodzące z prezentacji, z programów kształcenia i sylabusów, z wystąpień konferencyjnych, które nie zostały opublikowane oraz takie cytowania, których źródła GS nie podaje w sposób niebudzący wątpliwości skąd pochodzą. W celu ograniczenia ilości informacji do zakresu niezbędnego z punktu widzenia realizacji celów badań, analizy przeprowadzono z uwzględnieniem tekstów, które uzyskały przynajmniej trzy zweryfikowane cytowania.

### **3.2. Etap 2. Weryfikacja kompletności materiału badawczego**

Kompletność danych, w sensie uwzględnienia w zestawie wyników wyszukiwania możliwie wszystkich publikacji spełniających opisane powyżej kryteria i indeksowanych przez Google Scholar, sprawdzono w drugim etapie badania. Polegał on na przeprowadzeniu wyszukiwania za pomocą nazwisk i imion (wziętych w cudzysłów) 41 badaczy, których nazwiska pojawiły się w wynikach wyszukiwania w pierwszym etapie badań i którzy w okresie 2010–2019 zaliczani byli do reprezentantów nauki o informacji (według informacji dostępnych na stronach WWW poszczególnych uczelni oraz wiedzy autora). Początkowo wyniki ograniczono do języka polskiego. Szukano w nich prac, które nie pojawiły się na poprzednich czterdziestu listach z wynikami wyszukiwania, a które uzyskały przynajmniej trzy zweryfikowane

cytowania. Chcąc stworzyć pełny ranking prac stosunkowo często cytowanych, powtórzono wyszukiwanie według autorów, ale z uwzględnieniem wszystkich języków.

### **3.3. *Etap 3. Weryfikacja hipotezy o związku miejsca publikacji na liście wyników z występowaniem wyrażen wyszukiwawczych wśród metadanych tej publikacji***

Trzeci etap badań zrealizowano w związku z tym, że w poprzednich etapach okazało się, iż porządek pozycji bibliograficznych w wynikach wyszukiwania w GS w dużej części przypadków nie był oparty na liczbie cytowań poszczególnych publikacji (wbrew sugestiom innych badań, omówionych w przeglądzie piśmiennictwa). Nie stwierdzono też istotnego związku pomiędzy tym, w ilu wynikach wyszukiwania pojawiała się dana praca, a tym, ile uzyskała cytowań. Fakt ten sprzeczny jest z racjonalnym przekonaniem sugerującym, że im częściej (w odpowiedzi na więcej wyrażen wyszukiwawczych) dana praca pojawia się w wynikach Google Scholar (wyszukiwarki powszechnie stosowanej przez środowisko naukowe), tym więcej ma szans na cytowanie.

W związku z tym postanowiono poddać weryfikacji wstępną hipotezę, sugerującą, że zarówno obecność danej pracy w wynikach wyszukiwania, jak i jej pozycję na liście tych wyników, zaburza problem słów kluczowych. Szukano więc słów kluczowych w tytule, tekście (w pozycji „słowa kluczowe”), metryce pliku, na stronie WWW wydawcy/czasopisma lub w bibliotece cyfrowej/repozytorium, gdzie plik z tekstem został udostępniony. Ponadto przyjrano się drugiemu z czynników mogących wpływać na sortowanie wyników wyszukiwania w GS – zgodności użytych w pracy (w miejscach wymienionych powyżej) słów kluczowych z wyrażeniami wyszukiwawczymi, dzięki którym dana praca została wyszukana. Jak można wstępnie domniemywać, skoro dana praca została wyszukana przy użyciu np. terminu „biblioteka cyfrowa”, znaczy, że tekst lub jego opis na stronie WWW powinien zawierać ten termin przynajmniej kilkukrotnie. Na tym etapie, ze względu na konieczność minimalizowania ilości danych, które pomieściłyby się w ramach jednego artykułu, badanie ograniczono do prac najczęściej cytowanych, co najmniej dwadzieścia razy, lub wyszukanych za pomocą co najmniej siedmiu wyrażen wyszukiwawczych (bez nazwisk i imion autorów). Uwzględniono dwadzieścia prac, których pełne teksty w formacie PDF lub HTML dostępne były w Internecie bez opłat i konieczności logowania się.

Dane zbierano w okresie od maja do października 2020 r., a więc wyniki badań nie uwzględniają późniejszych zmian w zasobach Google Scholar.

## **4. Wyniki badań**

### **4.1. *Identyfikacja stosunkowo często cytowanych badaczy i ich publikacji***

Wyszukiwanie przeprowadzone w zasobie danych Google Scholar wykazało obecność 259 prac spełniających opisane powyżej kryteria (plik z danymi i rankingiem cytowań tych prac dostępny w RepOD – <https://repod.icm.edu.pl>). W tym zestawie znalazło się 59 publikacji, które w realiach polskiej nauki o informacji można uznać za stosunkowo często cytowane: każda z nich otrzymała co najmniej 10 cytowań (Zał. 1).

Autorami lub współautorami prac wyszukanych w GS jest 157 osób w tym: Marzena Świgoń – 17 prac, Wiesław Babik – 10, Aneta Drabek – 10, Ewa Głowacka – 8, Emanuel Kulczycki – 8, Ewa A. Rozkosz – 8, Sabina Cisek – 7, Władysław M. Kolasa – 7, Jacek Wojciechowski – 7, Lidia Derfert-Wolf – 6, Małgorzata Kisilowska – 6, Veslava Osińska – 6, Hanna Batorowska – 5, Justyna Jasiewicz – 5, Małgorzata Kowalska – 5, Małgorzata Fedorowicz-Kruszewska – 4, Grzegorz Gmiterek – 4, Mirosław Górny – 4, Marek Nahotko – 4, Stanisław Skórka – 4, Marcin Werla – 4, Jadwiga Woźniak-Kasperek – 4, Małgorzata Janiak – 3, Barbara Kamińska-Czubala – 3, Katarzyna Materska – 3, Marcin Roszkowski – 3, Barbara Sosińska-Kalata – 3, Piotr Taflowski – 3, Magdalena Wójcik – 3.

Dominują artykuły w czasopismach – 158 (61%), z czego 29 opublikowano w periodykach zagranicznych. Wyszukano też 66 artykułów w pracach zbiorowych (25,5%) oraz 35 monografii (13,5%). W przypadku artykułów w czasopismach polskich dominują opublikowane w: *Biuletynie EBIB* – 22, *Przeglądzie Bibliotecznym* – 20 oraz w *Zagadnieniach Informacji Naukowej* – 18.

Spośród 259 prac, 203 wyszukano przy użyciu jednego z czterdziestu wyrażen wyszukiwawczych. Kolejne 31 wyszukano stosując imię i nazwisko badacza z ograniczeniem do języka polskiego, a następne 25 – likwidując to ograniczenie. Tak więc prawie 22% prac, spośród cytowanych co najmniej trzy razy, nie pojawiło się w wynikach wyszukiwania GS, jeżeli zastosowano związane z nauką o informacji wyrażenia wyszukiwawcze w języku polskim.

Autorami lub współautorami 59 prac cytowanych co najmniej dziesięć razy (Zał. 1) jest 40 autorów, wśród których zdecydowanie dominuje M. Świgoń – autorka lub współautorka 13 publikacji zaindeksowanych przez GS. Liczebnością prac w tym zestawieniu wyróżnił się także W.M. Kolasa – 5 publikacji oraz E. Kulczycki – 4 publikacje.

Do najczęściej cytowanych prac należą monografie – 21 spośród 35 wyszukanych oraz artykuły w czasopismach zagranicznych – 16 spośród 29 wyszukanych.

W zestawie prac cytowanych co najmniej dziesięć razy, 38 wyszukano stosując słowa kluczowe, 8 stosując imię i nazwisko badacza z ograniczeniem do języka polskiego, a następne 13 – likwidując to ograniczenie. Tak więc, jeżeli zastosowano polskie terminy wyszukiwawcze związane z nauką o informacji, spośród prac cytowanych stosunkowo najczęściej, prawie 36% nie pojawiło się w wynikach wyszukiwania w GS.

#### **4.2. Liczba wyrażen wyszukiwawczych związanych z publikacją i jej miejsce na liście wyników a liczba zdobytych cytowań**

Okazało się, że poszczególne prace pojawiały się na listach wyników przy bardzo zróżnicowanej liczbie zapytań wyszukiwawczych. W skrajnych przypadkach zidentyfikowano publikacje, które na listach wyników pojawiły się w odpowiedzi na zastosowanie od jednego do jednego wyrażenia wyszukiwawczego. Stwierdzono, że skojarzenie konkretnego tekstu ze stosunkowo dużą liczbą wyrażen wyszukiwawczych, co teoretycznie powinno dawać większą szansę na znalezienie tej publikacji w zasobach Internetu, nie koresponduje w sposób wyraźny z większą liczbą cytowań (Tab. 1). Wśród 24 prac występujących przynajmniej w pięciu wynikach wyszukiwania znajdujemy 6 publikacji spośród 38 cytowanych co najmniej dziesięć razy (16%). Na pozostałe 165 prac, cytowanych mniej niż dziesięć razy, przypada 18 publikacji występujących przynajmniej w pięciu wynikach wyszukiwania (11%). W prawie połowie przypadków wykazana przez GS praca pojawiła się jedynie

w jednym wyniku wyszukiwania – 99 na 203 (49%). Taka sytuacja dotyczyła 21 spośród 38 prac cytowanych dziesięć i więcej razy (55%), co potwierdza brak wyraźnego związku pomiędzy potencjalną szansą na znalezienie tekstu pracy w GS (liczba skojarzonych z nią wyrażeń wyszukiwawczych), a jej cytawalnością, w sytuacji, gdy stosujemy wyrażenia wyszukiwawcze w języku polskim.

Jednego z wyjaśnień takiego stanu rzeczy dostarcza analiza pozycji danej pracy w wynikach wyszukiwania. Okazuje się, że najczęściej jest to odległa pozycja, poza kilkoma pierwszymi stronami wyników, które zdecydowana większość użytkowników Internetu sprawdza. Na 24 prace z grupy wyszukanych za pomocą pięciu i więcej terminów przypada w sumie 158 rekordów. Z tego jedynie 27 rekordów (17%) pojawiło się na pierwszych trzech stronach wyników. Na trzeci i czwarty kwartył listy wyników wyszukiwania, czyli tam, gdzie niewielu poszukujących dociera, przypada 56 (35.5%) wyszukanych rekordów. Jeżeli podobnej analizie poddamy zbiór 38 prac cytowanych co najmniej dziesięć razy, a wyszukanych z zastosowaniem wyrażeń wyszukiwawczych (Zał. 1), to okazuje się, że uzyskanych rekordów było 91. Z tego 25 (27.5%) znalazło się na pierwszych trzech stronach wyników. Zaś w trzecim i czwartym kwartylu listy wyników wyszukiwania było ich 44 (48%). Tak więc sugestia Beel et al. (2010), iż jednym z czynników wpływających na pozycję pracy w wynikach wyszukiwania GS jest liczba cytowań, w przypadku stosowania haseł wyszukiwawczych w języku polskim nie znajduje potwierdzenia. W zależności od zastosowanego wyrażenia wyszukiwawczego, prawie połowa prac stosunkowo często cytowanych pojawia się dopiero w trzecim i czwartym kwartylu listy wyników.

Tabela 1. Prace wyszukane za pomocą pięciu i więcej wyrażeń wyszukiwawczych

L.p.	Autor	Tytuł	Liczba zweryfikowanych cytowań	Wyrażenia i strona w wynikach / liczba stron wyników
1	2	3	4	5
1	Babik W.	Ekologia informacji katalizatorem równoważenia rozwoju społeczeństwa informacji i wiedzy	9	1. Zarządzanie informacją 50/100 2. Kompetencje informacyjne 28/57 3. Wyszukiwanie informacji 16/100 4. Procesy informacyjne 28/53 5. Ekologia informacji 1/21 6. Bariery informacyjne 6/29 7. Użytkownicy informacji 1/25 8. Środowisko informacyjne 9/21 9. Zachowania informacyjne 15/29 10. Przestrzeń informacyjna 16/44 11. Potrzeby informacyjne 44/100

1	2	3	4	5
2	Sosińska-Kalata B.	Obszary badań współczesnej informatologii (nauki o informacji)	25	<ol style="list-style-type: none"> <li>1. Nauka o informacji 1/51</li> <li>2. Informatologia 19/34</li> <li>3. Kompetencje informacyjne 33/57</li> <li>4. Komunikacja naukowa 16/40</li> <li>5. Wyszukiwanie informacji 3/100</li> <li>6. Procesy informacyjne 37/53</li> <li>7. Bibliometria 18/42</li> <li>8. Etyka informacyjna 2/2</li> <li>9. Zachowania informacyjne 15/29</li> <li>10. Działalność informacyjna 12/82</li> </ol>
3	Batorowska H., Klepka R., Wasiuta O.	Media jako instrument wpływu informacyjnego i manipulacji społeczeństwem	7	<ol style="list-style-type: none"> <li>1. Nauka o informacji 20/51</li> <li>2. Kompetencje informacyjne 19/57</li> <li>3. Teoria informacji 68/100</li> <li>4. Procesy informacyjne 45/53</li> <li>5. Ekologia informacji 21/21</li> <li>6. Wyszukiwanie informacji 93/100</li> <li>7. Architektura informacji 21/28</li> <li>8. Bariery informacyjne 21/29</li> <li>9. Użytkownicy informacji 23/25</li> <li>10. Bezpieczeństwo informacyjne 50/56</li> </ol>
4	Jasiewicz J.	Obszary, problemy i metody badawcze library and information science. Część I	3	<ol style="list-style-type: none"> <li>1. Nauka o informacji 7/51</li> <li>2. Informatologia 3/34</li> <li>3. Biblioteki akademickie 60/71</li> <li>4. Kompetencje informacyjne 5/57</li> <li>5. Wyszukiwanie informacji 17/100</li> <li>6. Teoria informacji 44/100</li> <li>7. Systemy informacyjne 8/100</li> <li>8. Bibliometria 14/42</li> <li>9. Zachowania informacyjne 10/29</li> </ol>
5	Babik W.	Biblioteka akademicka na rozdrożu: o współczesnych przemianach w środowisku informacyjnym bibliotek	6	<ol style="list-style-type: none"> <li>1. Biblioteki akademickie 4/71</li> <li>2. Kompetencje informacyjne 28/57</li> <li>3. Wyszukiwanie informacji 44/100</li> <li>4. Teoria informacji 58/100</li> <li>5. Systemy informacyjne 8/100</li> <li>6. Środowisko informacyjne 4/21</li> <li>7. Kultura informacyjna 15/39</li> <li>8. Przestrzeń informacyjna 12/44</li> <li>9. Potrzeby informacyjne 38/100</li> </ol>

1	2	3	4	5
6	Woźniak-Kasperek J.	Języki informacyjne: między tradycją a nadzieją na cyfrową przyszłość	5	<ol style="list-style-type: none"> <li>1. Nauka o informacji 22/51</li> <li>2. Biblioteki akademickie 51/71</li> <li>3. Kompetencje informacyjne 7/57</li> <li>4. Wyszukiwanie informacji 68/100</li> <li>5. Architektura informacji 18/28</li> <li>6. Użytkownicy informacji 9/25</li> <li>7. Środowisko informacyjne 1/21</li> <li>8. Obiekty informacyjne 1/5</li> </ol>
7	Cisek S.	Metoda analizy i krytyki piśmiennictwa w nauce o informacji i bibliotekoznawstwie w XXI wieku	20	<ol style="list-style-type: none"> <li>1. Nauka o informacji 1/51</li> <li>2. Informatologia 9/34</li> <li>3. Zarządzanie informacją 93/100</li> <li>4. Biblioteki akademickie 60/71</li> <li>5. Komunikacja naukowa 23/40</li> <li>6. Bibliometria 40/42</li> <li>7. Zachowania informacyjne 19/29</li> </ol>
8	Kisilowska M.	Przestrzeń informacyjna jako termin informatologiczny	13	<ol style="list-style-type: none"> <li>1. Kompetencje informacyjne 22/57</li> <li>2. Teoria informacji 67/100</li> <li>3. Ekologia informacji 13/21</li> <li>4. Środowisko informacyjne 1/21</li> <li>5. Zachowania informacyjne 19/29</li> <li>6. Działalność informacyjna 7/82</li> <li>7. Przestrzeń informacyjna 1/44</li> </ol>
9	Pamuła-Cieślak N.	Ukryty Internet jako przedmiot edukacji informacyjnej	8	<ol style="list-style-type: none"> <li>1. Biblioteki akademickie 46/71</li> <li>2. Komunikacja naukowa 26/40</li> <li>3. Wyszukiwanie informacji 9/100</li> <li>4. Teoria informacji 53/100</li> <li>5. Bariery informacyjne 16/29</li> <li>6. Kultura informacyjna 18/39</li> <li>7. Przestrzeń informacyjna 9/44</li> </ol>
10	Babik W.	Słowa kluczowe	35	<ol style="list-style-type: none"> <li>1. Język haseł przedmiotowych 9/14</li> <li>2. Języki informacyjno-wyszukiwawcze 9/12</li> <li>3. Teoria informacji 60/100</li> <li>4. Organizacja informacji 7/17</li> <li>5. Działalność informacyjna 68/82</li> <li>6. Przestrzeń informacyjna 38/44</li> </ol>
11	Kotuła S.D.	Komunikacja bibliologiczna wobec World Wide Web	13	<ol style="list-style-type: none"> <li>1. Nauka o informacji 29/51</li> <li>2. Informatologia 29/34</li> <li>3. Komunikacja naukowa 4/40</li> <li>4. Wyszukiwanie informacji 99/100</li> <li>5. Teoria informacji 80/100</li> <li>6. Obiekty informacyjne 4/5</li> </ol>

1	2	3	4	5
12	Cisek S.	Zachowania informacyjne – wybrane aspekty	4	<ol style="list-style-type: none"> <li>1. Nauka o informacji 17/51</li> <li>2. Zarządzanie informacją 85/100</li> <li>3. Systemy informacyjne 9/100</li> <li>4. Bariery informacyjne 3/29</li> <li>5. Zachowania informacyjne 1/29</li> <li>6. Potrzeby informacyjne 6/100</li> </ol>
13	Jasiewicz J.	Obszary, problemy i metody badawcze library and information science. Część II	3	<ol style="list-style-type: none"> <li>1. Nauka o informacji 19/51</li> <li>2. Informatologia 3/34</li> <li>3. Kompetencje informacyjne 5/57</li> <li>4. Czasopisma naukowe 30/100</li> <li>5. Bibliometria 12/42</li> <li>6. Naukometria 7/13</li> </ol>
14	Materska K.	Aktualność koncepcji zarządzania informacją w dobie big data-perspektywa informatologa	3	<ol style="list-style-type: none"> <li>1. Informatologia 6/34</li> <li>2. Zarządzanie informacją 2/100</li> <li>3. Biblioteki akademickie 58/71</li> <li>4. Wyszukiwanie informacji 67/100</li> <li>5. Bibliometria 21/42</li> <li>6. Naukometria 9/13</li> </ol>
15	Cisek S.	„Metodologia mieszana” w badaniach nauki o informacji i bibliotekoznawstwa	12	<ol style="list-style-type: none"> <li>1. Nauka o informacji 1/51</li> <li>2. Informatologia 16/34</li> <li>3. Użytkownicy informacji 1/25</li> <li>4. Zachowania informacyjne 17/29</li> <li>5. Potrzeby informacyjne 46/100</li> </ol>
16	Babik W.	Kultura informacyjna: spojrzenie z punktu widzenia ekologii informacji	8	<ol style="list-style-type: none"> <li>1. Zarządzanie informacją 75/100</li> <li>2. Ekologia informacji 3/21</li> <li>3. Środowisko informacyjne 9/21</li> <li>4. Kultura informacyjna 1/39</li> <li>5. Działalność informacyjna 6/82</li> </ol>
17	Janiak M.	Informacja naukowa w Polsce na przełomie XX i XXI wieku: dynamika zmian w świetle piśmiennictwa	7	<ol style="list-style-type: none"> <li>1. Nauka o informacji 10/51</li> <li>2. Informatologia 15/34</li> <li>3. Czasopisma naukowe 46/100</li> <li>4. Teoria informacji 38/100</li> <li>5. Działalność informacyjna 5/82</li> </ol>
18	Skórka S.	Nowe wyzwania architektury informacji	4	<ol style="list-style-type: none"> <li>1. Architektura informacji 1/28</li> <li>2. Informatologia 22/34</li> <li>3. Procesy informacyjne 35/53</li> <li>4. Organizacja informacji 1/17</li> <li>5. Przestrzeń informacyjna 16/44</li> </ol>
19	Tafiłowski P.	Architektura informacji jako problem badawczy informatologii	4	<ol style="list-style-type: none"> <li>1. Kompetencje informacyjne 23/57</li> <li>2. Architektura informacji 1/28</li> <li>3. Wyszukiwanie informacji 7/100</li> <li>4. Bariery informacyjne 16/29</li> <li>5. Przestrzeń informacyjna 18/44</li> </ol>

1	2	3	4	5
20	Batorowska H.	Konektywizm w kontekście kształcenia kompetencji informacyjnych	4	1. Kompetencje informacyjne 17/57 2. Ekologia informacji 15/21 3. Zachowania informacyjne 13/29 4. Kultura informacyjna 13/39 5. Potrzeby informacyjne 55/100
21	Nahotko M.	Metadane biblioteczne XXI wieku	4	1. Biblioteki akademickie 55/71 2. Komunikacja naukowa 3/40 3. Wyszukiwanie informacji 45/100 4. Działalność informacyjna 12/82 5. Potrzeby informacyjne 67/100
22	Jasiewicz J.	Zachowania informacyjne młodzieży a elementy edukacji informacyjnej w Polsce	3	1. Kompetencje informacyjne 2/57 2. Czasopisma naukowe 19/100 3. Zachowania informacyjne 1/29 4. Kultura informacyjna 16/39 5. Potrzeby informacyjne 13/100
23	Kamińska-Czubała B.	Kultura informacyjna użytkownika informacji – problemy i metody badań	3	1. Ekologia informacji 6/21 2. Środowisko informacyjne 9/21 3. Kultura informacyjna 1/39 4. Przestrzeń informacyjna 3/44 5. Potrzeby informacyjne 25/100
24	Walczak-Niewiadomska A., Czapnik G., Gruszka Z.	Brokerstwo informacyjne w Polsce – przegląd publikacji	3	1. Wyszukiwanie informacji 19/100 2. Teoria informacji 43/100 3. Użytkownicy informacji 9/25 4. Infobrokerstwo 2/11 5. Przestrzeń informacyjna 32/44

#### 4.3. Wyniki wyszukiwania a zastosowane wyrażenia wyszukiwawcze

Okazało się, że prace cytowane przynajmniej trzy razy (liczba cytowań po weryfikacji), mające związek z nauką o informacji, stanowią jedynie nieznaczny odsetek wyników wyszukiwania w GS mimo że zastosowano wyrażenia wyszukiwawcze ściśle związane z tą dyscypliną (Tab. 2). Odsetek takich prac waha się (pomijając jeden przypadek skrajny, który zawierał jedynie dwie pozycje bibliograficzne) od 0.5% do 9.3% w zależności od zastosowanego wyrażenia wyszukiwawczego. Najmniejszy odsetek prac związanych z nauką o informacji uzyskano stosując wyrażenia wyszukiwawcze, które występują także w takich dyscyplinach, jak nauki o zarządzaniu i ekonomia. Prace głównie z tych dyscyplin zdecydowanie dominowały w wynikach wyszukiwania przy zastosowaniu takich terminów jak: zarządzanie informacją, systemy informacyjne, zasoby informacyjne, działalność informacyjna, bezpieczeństwo informacyjne. Największy odsetek prac przyjętych do analizy wyszukano przy zastosowaniu wyrażen będących albo dwoma nazwami badanej dyscypliny (nauka o informacji, informatologia), albo wyrażeniami wąsko specjalistycznymi (obiekty informacyjne, naukomedia, bibliometria, czasopisma open access, zachowania informacyjne, etyka informacyjna).

Znaczący odsetek wyszukanych prac (wyraźna większość dla każdego terminu, jednakże dokładny procent nie jest istotny dla dalszego wyводу) stanowiły także te, które związane są z nauką o informacji, ale nie uzyskały nawet trzech zweryfikowanych cytowań. Warto przy tym podkreślić, że w przypadku siedmiu spośród czterdziestu wyrażeń wyszukiwawczych nie ma możliwości zapoznania się z kompletem wyników, albowiem GS udostępnia jedynie pierwszy tysiąc pozycji.

W zależności od zastosowanego wyrażenia wyszukiwawczego konieczne okazało się odrzucenie (poza dwoma przypadkami, w których cytowań nie było lub ich suma wynosiła 3 lub mniej) od 7% do 42% cytowań z powodów opisanych już powyżej. W ponad połowie przypadków odrzucono więcej niż 20% cytowań podanych przez Google Scholar.

Tabela 2. Wyniki wyszukiwania za pomocą słów kluczowych dla dyscypliny nauka o informacji

L.p.	Wyrażenie wyszukiwawcze	Liczba wyników w GS (max. 1 tys. widocznych)	Liczba prac zakwalifikowanych do analizy spośród max. 1 tys. pozycji w wynikach wyszukiwania: liczba/odsetek	Suma cytowań prac zakwalifikowanych do analizy	
				Wg GS	Po weryfikacji: liczba/odsetek
1	2	3	4	5	6
1	Nauka o informacji	510	27 / 5.29%	208	160 / 76.9%
2	Informatologia	340	20 / 5.88%	210	154 / 73.33%
3	Zarządzanie informacją	2890	13 / 1.3%	135	104 / 77.04%
4	Biblioteki akademickie	709	34 / 4.79%	277	237 / 85.56%
5	Kompetencje informacyjne	573	21 / 3.66%	166	133 / 80.12%
6	Komunikacja naukowa	400	23 / 5.75%	280	221 / 78.93%
7	Język haseł przedmiotowych	135	1 / 0.74%	48	35 / 72.92%
8	Rozwój dyscyplin naukowych	71	1 / 1.41%	6	5 / 83.33%
9	Wyszukiwanie informacji	1580	27 / 2.7%	235	185 / 78.72%
10	Języki informacyjno-wyszukiwawcze	121	4 / 3.31%	66	51 / 77.27%
11	Metody ilościowe w informatologii	2	0	0	0
12	Czasopisma naukowe	2090	21 / 2.1%	165	117 / 70.91%
13	Teoria informacji	1120	25 / 2.5%	275	204 / 74.18%
14	Architektura informacji	279	8 / 2.87%	70	54 / 77.14%
15	Procesy informacyjne	528	9 / 1.7%	87	72 / 82.76%

1	2	3	4	5	6
16	Systemy informacyjne	2840	7 / 0.7%	34	27 / 79.41%
17	Zasoby informacyjne	1340	5 / 0.5%	57	47 / 82.46%
18	Ekologia informacji	205	10 / 4.88%	80	60 / 75%
19	Bibliometria	418	25 / 5.98%	345	239 / 69.27%
20	Bariery informacyjne	288	11 / 3.82%	97	82 / 84.54%
21	Użytkownicy informacji	243	11 / 4.53%	141	98 / 69.5%
22	Etyka informacyjna	20	1 / 5%	30	25 / 83.33%
23	Źródła informacji naukowej	67	1 / 1.49%	3	3 / 100%
24	Środowisko informacyjne	202	10 / 4.95%	73	56 / 76.71%
25	Organizacja informacji	167	6 / 3.59%	119	98 / 82.35%
26	Zachowania informacyjne	287	16 / 5.57%	197	149 / 75.63%
27	Kultura informacyjna	388	13 / 3.35%	123	106 / 86.18%
28	Bibliografia załącznikowa	188	6 / 3.19%	62	43 / 69.35%
29	Wizualizacja informacji	230	4 / 1.74%	54	35 / 64.81%
30	Bibliograficzna baza danych	128	3 / 2.34%	14	11 / 78.57%
31	Czasopisma open access	53	3 / 5.66%	14	13 / 92.86%
32	Infobrokerstwo	102	5 / 4.9%	33	29 / 87.88%
33	Biblioteki cyfrowe	981	19 / 1.94%	108	96 / 88.89%
34	Naukometria	129	12 / 9.3%	123	84 / 68.29%
35	Działalność informacyjna	819	10 / 1.22%	143	107 / 74.83%
36	Bezpieczeństwo informacyjne	552	2 / 0.36%	12	10 / 83.33%
37	Reprezentacja i informacji	74	2 / 2.7%	9	8 / 88.89%
38	Potrzeby informacyjne	1760	15 / 1.5%	122	95 / 77.87%
39	Obiekty informacyjne	41	3 / 7.32%	38	22 / 57.89%
40	Przestrzeń informacyjna	435	20 / 4.6%	172	133 / 77.33%

#### 4.4. Zastosowane wyrażenia wyszukiwawcze a słowa kluczowe towarzyszące publikacji

Jedynie słaby związek został potwierdzony przez badanie zgodności wyrażeń wyszukiwawczych wpisywanych do wyszukiwarki Google Scholar ze słowami kluczowymi zamieszczonymi w tytule odszukanej pracy, w metadanych towarzyszących jej tekstowi (w pozycji „słowa kluczowe”), metryce pliku, na stronie WWW wydawcy/czasopisma lub w bibliotece cyfrowej/repozytorium, gdzie plik z tekstem został udostępniony (Tab. 3). Jakakolwiek zgodność wystąpiła w przypadku 12 na 20 prac cytowanych co najmniej dwadzieścia razy lub wyszukanych za pomocą co najmniej siedmiu wyrażeń (bez nazwisk i imion autorów). Sytuacja, w której wszystkie wyrażenia wyszukiwawcze miały swoje odpowiedniki w słowach kluczowych, wystąpiła jedynie w przypadku czterech tekstów odszukanych za pomocą jednego lub dwóch wyrażeń. Jeżeli praca była wyszukana za pomocą większej liczby wyrażeń, to jedynie część z tych wyrażeń miała swoje odpowiedniki w słowach kluczowych: połowa w jednym przypadku, co trzecie (lub ponad 28%) w trzech przypadkach, co czwarte (lub ponad 22%) w dwóch przypadkach, a co piąte lub mniej także w dwóch przypadkach. Taki wynik oznacza, że najczęściej wyrażenia, za pomocą których wyszukano daną pracę, nie miały swoich odpowiedników w słowach kluczowych związanych z tym tekstem (74 na 98 rekordów). Na cytowalność pracy oraz na szansę jej wyszukania za pomocą Google Scholar (mierzoną liczbą wyrażeń wyszukiwawczych) nie ma więc istotnego wpływu obudowanie tej pracy słowami kluczowymi (w sposób zgodny z wymaganiami tej wyszukiwarki). W takiej sytuacji nie można uznać, że na ranking wyników wyszukiwania w GS, ustalony cytowaniami, ma wpływ kwestia używania lub nie używania tychże słów kluczowych w języku polskim.

Tabela 3. Związek pomiędzy wyrażeniami wyszukiwawczymi a słowami kluczowymi związanymi z opublikowaną pracą

(Oznaczenia w kolumnie piątej: 1: na stronie WWW wydawcy/czasopisma w opisie publikacji; 2: w metryce pliku PDF; 3: w bibliotece lub repozytorium cyfrowym; 4: w pliku z tekstem pracy).

L.p.	Autor	Tytuł	Wyrażenia wyszukiwawcze	Słowa kluczowe w tekście lub Internecie
1	2	3	4	5
1	Babik W.	<u>Biblioteka akademicka</u> na rozdrożu: o współczesnych przemianach w <u>środowisku informacyjnym</u> bibliotek	<u>Biblioteki akademickie</u> Kompetencje informacyjne Wyszukiwanie informacji Teoria informacji Systemy informacyjne <u>Środowisko informacyjne</u> <u>Kultura informacyjna</u> Przestrzeń informacyjna Potrzeby informacyjne	1. brak 2. brak 3. <u>biblioteki akademickie</u> , <u>kultura informacyjna</u> , społeczeństwo informacyjne, społeczeństwo oparte na wiedzy 4. brak

1	2	3	4	5
2	Babik W.	<u>Ekologia informacji</u> katalizatorem równoważenia rozwoju społeczeństwa informacji i wiedzy	Zarządzanie informacją Kompetencje informacyjne Wyszukiwanie informacji Procesy informacyjne <u>Ekologia informacji</u> <u>Bariery informacyjne</u> Użytkownicy informacji Środowisko informacyjne Zachowania informacyjne Przestrzeń informacyjna Potrzeby informacyjne	1. brak 2. ekologia informacji; antropoinfosfera; zrównoważony rozwój; społeczeństwo informacji i wiedzy; nadmiar informacji; <u>bariery informacyjne</u> ; <u>informacja ekologiczna</u> 3. brak 4. ekologia informacji, antropoinfosfera, zrównoważony rozwój, społeczeństwo informacji i wiedzy, nadmiar informacji, <u>bariery informacyjne</u> , <u>informacja ekologiczna</u>
3	Babik W.	Słowa kluczowe	Język haseł przedmiotowych Języki informacyjno-wyszukiwawcze Teoria informacji Organizacja informacji Działalność informacyjna Przestrzeń informacyjna	1. informacja naukowa, termin, systemy wyszukiwania informacji, bibliotekoznawstwo, słowa kluczowe 2. keywords; information retrieval; indexing; Internet; information organization; indexing and retrieval languages; documentary indexing and retrieval system 3. brak 4. brak
4	Batorowska H.	Od alfabetyzacji informacyjnej do <u>kultury informacyjnej</u> : rozważania o dojrzałości informacyjnej	<u>Kultura informacyjna</u>	1. kultura, społeczeństwo informacyjne 2. brak 3. brak 4. brak
5	Batorowska H., Klepka R., Wasiuta O.	Media jako instrument wpływu informacyjnego i manipulacji społeczeństwem	Nauka o informacji Kompetencje informacyjne Teoria informacji Procesy informacyjne Ekologia informacji Wyszukiwanie informacji Architektura informacji Bariery informacyjne Użytkownicy informacji Bezpieczeństwo informacyjne	1. brak 2. media, manipulacja, wpływ polityczny, media, politics, media influence, media manipulations 3. brak 4. brak

1	2	3	4	5
6	Błocki Z., Życzkowski K.	Czy można porównywać jabłka i gruszki? O danych bibliometrycznych w różnych dziedzinach nauki	Bibliometria	1. brak 2. brak 3. brak 4. brak
7	Cisek S.	Metoda analizy i krytyki piśmiennictwa w <u>nauce o informacji</u> i bibliotekoznawstwie w XXI wieku	<u>Nauka o informacji</u> <u>Informatologia</u> Zarządzanie informacją Biblioteki akademickie Komunikacja naukowa Bibliometria Zachowania informacyjne	1. analiza i krytyka piśmiennictwa; badania literaturowe; bibliotekoznawstwo; <u>informatologia</u> ; <u>nauka o informacji</u> ; metaanaliza; metodologia; metody badań; przegląd systematyczny; stan badań; XXI wiek 2. brak 3. brak 4. brak
8	Jasiewicz J.	<u>Kompetencje informacyjne</u> młodzieży: analiza-stan faktyczny-kształcenie na przykładzie Polski, Niemiec i Wielkiej Brytanii	<u>Kompetencje informacyjne</u>	1. brak 2. brak 3. informacja cyfrowa; media masowe – edukacja; nowe technologie; społeczeństwo informacyjne 4. brak
9	Jasiewicz J.	Obszary, problemy i metody badawcze library and information science. Część I	Nauka o informacji <u>Informatologia</u> Biblioteki akademickie Kompetencje informacyjne Wyszukiwanie informacji Teoria informacji Systemy informacyjne <u>Bibliometria</u> Zachowania informacyjne	1. brak 2. brak 3. brak 4. bibliologia i <u>informatologia</u> , library and information science, obszary badawcze, metody badawcze, <u>bibliometria</u>
10	Kisilowska M.	<u>Przestrzeń informacyjna</u> jako termin informatologiczny	Kompetencje informacyjne Teoria informacji Ekologia informacji <u>Środowisko informacyjne</u> Zachowania informacyjne Działalność informacyjna <u>Przestrzeń informacyjna</u>	1. brak 2. brak 3. brak 4. <u>przestrzeń informacyjna</u> , <u>środowisko informacyjne</u> , infosfera

1	2	3	4	5
11	Klincewicz K., Mijał M., Żemigajła M.	<u>Bibliometria</u> w zarządzaniu technologiami i badaniami naukowymi.	<u>Bibliometria</u>	1. brak 2. brak 3. <u>bibliometria</u> 4. brak
12	Kola- sa W.M.	Historiografia prasy polskiej (do 1918 roku): naukome- tryczna analiza dyscypliny 1945–2009	<u>Bibliometria</u> <u>Naukometria</u>	1. brak 2. brak 3. historiography, history of the press, research, Poland, citation analysis, scientometrics, bibliometrics, historiografia, historia prasy, badania, Polska, analiza cytowań, <u>naukometria</u> , <u>bibliometria</u>
13	Kulczycki E.	Wykorzystanie mediów społeczności- wych przez akademickie uczelnie wyż- sze w Polsce. Badania w for- mule otwartego notatnika.	Komunikacja naukowa	1. brak 2. Social Media, Open Science, Blog, Blogging, Blogsphere, Science, Communication, University 3. brak 4. brak
14	Nahotko M.	<u>Komunika- cja naukowa</u> w środowisku cyfrowym: glo- balna biblioteka cyfrowa w in- formatycznej infrastrukturze nauki	<u>Komunikacja naukowa</u> Wyszukiwanie informacji Użytkownicy informacji Organizacja informacji	1. informacja naukowa – technologia, biblioteki cyfrowe 2. brak 3. brak 4. brak
15	Pamuła- -Cieślak N.	Ukryty Internet jako przedmiot edukacji informacyjnej	Biblioteki akademickie Komunikacja naukowa Wyszukiwanie informacji Teoria informacji Bariery informacyjne Kultura informacyjna Przestrzeń informacyjna	1. brak 2. brak 3. brak 4. brak

1	2	3	4	5
16	Sapa R., Krakowska M., Janiak M.	Information seeking behaviour of mathematicians: scientists and students	Zachowania informacyjne	1. brak 2. brak 3. brak 4. brak
17	Sosińska-Kalata B.	Obszary badań współczesnej <u>informatologii</u> ( <u>nauki o informacji</u> )	<u>Nauka o informacji</u> <u>Informatologia</u> Kompetencje informacyjne Komunikacja naukowa Wyszukiwanie informacji Procesy informacyjne Bibliometria Etyka informacyjna Zachowania informacyjne Działalność informacyjna	1. <u>Informatologia</u> . Definicje. Interdyscyplinarność. <u>Nauka o informacji</u> . Nurty badawcze. Pole badawcze. Problemy badawcze 2. brak 3. brak 4. <u>Informatologia</u> . Definicje. Interdyscyplinarność. <u>Nauka o informacji</u> . Nurty badawcze. Pole badawcze. Problemy badawcze
18	Świgoń M.	Information barriers in libraries: types, typologies and Polish empirical studies	Bariery informacyjne	1. Information management, Libraries, Barriers, Anxiety 2. brak 3. brak 4. Information management, Libraries, Barriers, Anxiety
19	Wojciechowski J.	Biblioteka w komunikacji publicznej	Biblioteki akademickie	1. biblioteki publiczne 2. brak 3. brak 4. brak
20	Woźniak-Kasperek J.	Języki informacyjne: między tradycją a nadzieją na cyfrową przyszłość	Nauka o informacji Biblioteki akademickie Kompetencje informacyjne <u>Wyszukiwanie informacji</u> Architektura informacji <u>Użytkownicy informacji</u> <u>Środowisko informacyjne</u> <u>Obiekty informacyjne</u>	1. brak 2. brak 3. brak 4. Ilość informacji. Integracja zasobów informacyjnych. Integracja wyszukiwania. Język informacyjno-wyszukiwawczy. <u>Obiekt informacyjny</u> . Sieć. <u>Środowisko informacyjne</u> . <u>Użytkownicy informacji</u> . <u>Wyszukiwanie informacji</u> . Zmiana

## 5. Wnioski

Zastosowana metodologia ma pewne ograniczenia. Firma Google nie ujawnia wszystkich mechanizmów i algorytmów odpowiedzialnych za tworzenie listy wyników w odpowiedzi na określone wyrażenie wyszukiwawcze oraz za porządkowanie wyników. W związku z tym trudno o jednoznaczne określenie powodów, dla których czasami stosunkowo często cytowane publikacje w języku polskim umieszczane są na początku listy wyników, a czasami na bardzo dalekich pozycjach. Podobnie trudno o jednoznaczne wyjaśnienie zjawiska polegającego na tym, że dana praca pojawia się na liście wyników w odpowiedzi na zastosowanie wyrażenia wyszukiwawczego, które nie jest tożsame ze słowami kluczowymi zawartymi w metadanych tej pracy. Trzecie zjawisko, które trudno jednoznacznie wyjaśnić, polega na tym, że na liście wyników związanej z wyrażeniem wyszukiwawczym w języku polskim pojawia się praca w języku angielskim (informacje od wydawcy na stronie WWW także są po angielsku). Sytuacja taka dotyczyła jednak jedynie nielicznych publikacji. Z kolei wątpliwość dotycząca odsetka prac polskich autorów opublikowanych w języku angielskim, które nie zostały wykazane w wynikach wyszukiwania przy zastosowaniu wyrażenia wyszukiwawczego w języku polskim, znalazła jednak swoje wyjaśnienie po zastosowaniu wyrażen wyszukiwawczych w postaci imion i nazwisk tych autorów, o czym już wspomniano omawiając wyniki badań. Publikacje w języku angielskim, pominięte ze względu na stosowanie wyrażen wyszukiwawczych w języku polskim, to ok. 10% ogólnej liczby prac uwzględnionych w badaniach.

Przeprowadzone za pomocą GS wyszukiwanie prac z zakresu nauki o informacji wykazało, że w przypadku wyrażen wyszukiwawczych w języku polskim nie znajduje potwierdzenia teza Stribu et al. (2015) mówiąca, iż ta wyszukiwarka nadaje się do niezawodnej identyfikacji najczęściej cytowanych prac i badaczy. Okazało się bowiem, że algorytmy decydujące o doborze prac do danego wyrażenia wyszukiwawczego oraz porządkujące wyniki wyszukiwania w przypadku użycia języka polskiego są wysoce zawodne. Znacząca statystycznie część prac indeksowanych w GS nie pojawia się w wynikach, gdy wyszukiwanie przeprowadza się na podstawie słów kluczowych fundamentalnych dla nauki o informacji. Publikacje te zostają wyszukane dopiero wtedy, gdy wyrażeniem wyszukiwawczym jest imię i nazwisko autora. Jeszcze większa grupa tekstów stosunkowo często cytowanych umieszczona jest na bardzo dalekich miejscach w wynikach wyszukiwania, co powoduje, że niewielu szukających do nich dotrze. Nie można ustalić, jakie prace GS umieszcza powyżej tysięcznej pozycji w wynikach wyszukiwania. Nie znajduje także potwierdzenia teza Martina-Martina et al. (2017) mówiąca, że głównym kryterium porządkowania wyników wyszukiwania jest liczba cytowań wyszukanych prac. Okazało się bowiem, że znacząca statystycznie grupa prac w języku polskim, stosunkowo często cytowanych, znajduje się dopiero w trzecim i czwartym kwartylu wyników wyszukiwania.

Stwierdzono też, że zdecydowana większość publikacji wskazanych w wynikach wyszukiwania nie jest adekwatna do intencji wyszukującego. GS nie odróżnia bowiem prac pochodzących z różnych dyscyplin naukowych, które co prawda stosują niekiedy identyczne słowa kluczowe, ale zakres ich badań i treść artykułów jest znacząco różna. Ponadto, wyszukiwarka GS nie dość dobrze profiluje zainteresowania użytkownika. Uwagę tę należy umieścić w kontekście faktu, że prawie każda wyszukiwarka zapisuje aktywność każdego adresu IP i profiluje użytkownika korzystającego z urządzenia, któremu ten adres jest

przyporządkowany. Autor niniejszej pracy przez kilka miesięcy korzystał z tego samego komputera, był zalogowany na konto Google i w tym czasie w GS wpisywał jedynie hasła związane z nauką o informacji. W wynikach wyszukiwania korzystał jedynie z informacji na temat prac, które zakwalifikował jako przynależne do tej dyscypliny. Najwyraźniej GS nie stworzyło właściwego profilu zainteresowań autora, albowiem relewancja wyników pod koniec badań była na podobnym poziomie jak na początku.

Wyszukiwarka GS nie odróżnia też prac naukowych od innych materiałów zamieszczonych w zasobach indeksowanych przez własne roboty internetowe. Znaczącą statystycznie część cytowań przypisanych każdej prawie pracy trudno uznać za świadczącą o poziomie wpływu tej pracy na rozwój badań. Są to bowiem autocytowania, duplikaty cytowań, cytowania pochodzące ze źródeł innych niż opublikowana praca naukowa oraz różne błędy.

Tak więc pojawienie się tekstu w wynikach wyszukiwania nie oznacza, że ma on związek z intencjami wyszukującego, wyrażonymi za pomocą określonych terminów wyszukiwawczych i dotychczasowymi poszukiwaniami w zasobach indeksowanych przez GS, a towarzysząca mu informacja o liczbie cytowań nie zawsze oznacza, że tyle właśnie cytowań zdobyła ta praca. Doszukiwanie się związku pomiędzy liczbą cytowań pracy podawaną przez GS, a jej popularnością (bo o jakości nie można wyrokować na podstawie cytowań), zbyt często wiedzie na manowce.

W świetle przeprowadzonych badań można stwierdzić, że czynnikiem wpływającym w sposób zauważalny na kolejność prac w wynikach wyszukiwania raczej nie jest dobór słów kluczowych towarzyszących danej pracy. Autor tekstu i jego wydawca, dzięki właściwemu zastosowaniu słów kluczowych, nie zwiększają szansy na wyszukanie tego tekstu w GS za pomocą polskich wyrażen wyszukiwawczych. Teza o tym, że możliwe jest opracowanie zasad optymalizacji pod kątem uzyskania przez daną pracę miejsca na pierwszych stronach wyników, w których to zasadach kluczowe będzie odpowiednie użycie słów kluczowych (Beel et al., 2010), jest prawdziwa w sytuacji używania języka angielskiego, ale mocno wątpliwa w przypadku języka polskiego. Przeprowadzone badanie nie dało odpowiedzi na pytanie, co zwiększa szanse tekstu naukowego na odnalezienie go w Internecie przez potencjalnie zainteresowanego, polskiego czytelnika. Nie wpływa na to ani cytowalność pracy, ani zabiegi autora i wydawców związane z wyborem słów kluczowych. Także stosowanie wyrażen wyszukiwawczych będących terminami kluczowymi dla nauki o informacji oraz przeglądanie więcej niż kilku pierwszych stron wyników wyszukiwania nie daje gwarancji odnalezienia każdego tekstu tematycznie związanego z danym terminem, nawet jeżeli uzyskał on stosunkowo dużą liczbę cytowań.

Wobec tego sugestie, o tym że Google Scholar można profesjonalnie (np. w celach badawczych) wykorzystywać do wyszukiwania najbardziej popularnych (czyli stosunkowo często cytowanych) prac naukowych oraz do formułowania ocen (bazujących na cytowaniach) tychże prac opublikowanych w języku polskim i badaczy publikujących głównie w tym języku, w realiach nauki o informacji zbyt często okazują się błędne.

### Załącznik 1. Ranking prac, które uzyskały przynajmniej 10 zweryfikowanych cytowań

W ostatniej kolumnie informacja „Autor” – pozycja znaleziona, gdy wyrażeniem wyszukiwawczym było imię i nazwisko autora plus ograniczenie wyników do języka polskiego; „Autor ang.” – brak ograniczenia wyników do języka polskiego.

L.p.	Autor	Tytuł	Czasopismo/ Praca zbiorowa	Rok pu- bli- kacji	Liczba zwery- fikowa- nych cyto- wań	Wyrażenia wyszukiawcze i strona w wyni- kach/liczba stron
1	2	3	4	5	6	7
1	Nicholas D., Boukacem-Zeghmouri Ch., Rodríguez-Bravo B., Xu J., Watkinson A., Abrizah A., Herman E., <u>Świgoń M.</u>	Where and how early career researchers find scholarly information	<i>Learned Publishing</i>	2017	57	Autor ang.
2	Nicholas D., Rodríguez-Bravo B., Boukacem-Zeghmouri Ch., Rodríguez-Bravo B., Xu J., Watkinson A., Abrizah A., Herman E., <u>Świgoń M.</u>	Early career researchers and their publishing and authorship practices	<i>Learned Publishing</i>	2017	50	Autor ang.
3	Świgoń M.	Library anxiety among Polish students: Development and validation of the Polish Library Anxiety Scale	<i>Library &amp; Information Science Research</i>	2011	44	Autor ang.
4	Batorowska H.	<i>Od alfabetyzacji informacyjnej do kultury informacyjnej: rozważania o dojrzałości informacyjnej</i>		2013	43	Kultura informacyjna 1/39

1	2	3	4	5	6	7
5	Klincewicz K., Mijał M., Żemigala M.	<i>Bibliometria w zarządzaniu technologiami i badaniami naukowymi</i>		2012	40	Bibliometria 1/42
6	Nahotko M.	<i>Komunikacja naukowa w środowisku cyfrowym: globalna biblioteka cyfrowa w informatycznej infrastrukturze nauki</i>		2010	40	Komunikacja naukowa 1/40 Wyszukiwanie informacji 61/100 Użytkownicy informacji 23/25 Organizacja informacji 14/17
7	Wojciechowski J.	<i>Biblioteka w komunikacji publicznej</i>		2010	40	Biblioteki akademickie 61/71
8	Nicholas D., Rodríguez-Bravo B., Boukacem-Zeghmouri Ch., Rodríguez-Bravo B., Xu J., Watkinson A., Abrizah A., Herman E., Świgoń M.	Early career researchers: Scholarly behaviour and the prospect of change	<i>Learned Publishing</i>	2017	36	Autor ang.
9	Babik W.	<i>Słowa kluczowe</i>		2010	35	Język haseł przedmiotowych 9/14 Języki informacyjno-wyszukiwawcze 9/12 Teoria informacji 60/100 Organizacja informacji 7/17 Działalność informacyjna 68/82 Przestrzeń informacyjna 38/44
10	Gmiterek G.	<i>Biblioteka w środowisku społecznościowego Internetu: biblioteka 2.0</i>		2012	33	Autor
11	Woźniak-Kasperek J.	<i>Wiedza i język informacyjny w paradygmacie sieciowym</i>		2011	32	Autor

1	2	3	4	5	6	7
12	Kolasa W.M.	<i>Historiografia prasy polskiej (do 1918 roku): naukowemetryczna analiza dyscypliny 1945–2009</i>		2013	30	Bibliometria 14/42 Naukometria 9/13
13	Świgoń M.	Information barriers in libraries: types, typologies and Polish empirical studies	<i>Library Management</i>	2011	27	Bariery informacyjne 10/29
14	Wojciechowski J.	<i>Biblioteki w nowym otoczeniu</i>		2014	27	Autor
15	Sapa R., Krakowska M., Janiak M.	Information seeking behaviour of mathematicians: scientists and students	<i>Information Research</i>	2014	25	Zachowania informacyjne 26/29
16	Sosińska-Kalata B.	Obszary badań współczesnej informatologii (nauki o informacji)	<i>Zagadnienia Informatyki Naukowej</i>	2013	25	Nauka o informacji 1/51 Informatologia 19/34 Kompetencje informacyjne 33/57 Komunikacja naukowa 16/40 Wyszukiwanie informacji 3/100 Procesy informacyjne 37/53 Bibliometria 18/42 Etyka informacyjna 2/2 Zachowania informacyjne 15/29 Działalność informacyjna 12/82
17	Świgoń M.	Personal knowledge and information management—conception and exemplification	<i>Journal of Information Science</i>	2013	25	Autor ang.

1	2	3	4	5	6	7
18	Jasiewicz J.	<i>Kompetencje informacyjne młodzieży: analiza-stan faktyczny-kształcenie na przykładzie Polski, Niemiec i Wielkiej Brytanii</i>		2012	24	Kompetencje informacyjne 1/57
19	Świgoń M.	<i>Zarządzanie wiedzą i informacją. Podstawy teoretyczne. Badania w wymiarze indywidualnym</i>		2012	24	Autor
20	Błocki Z., Życzkowski K.	Czy można porównywać jabłka i gruszki? O danych bibliometrycznych w różnych dziedzinach nauki	<i>Nauka</i>	2013	23	Bibliometria 26/42
21	Nicholas D., Rodríguez-Bravo B., Boukacem-Zeghmouri Ch., Rodríguez-Bravo B., Xu J., Watkinson A., Abrizah A., Herman E., Świgoń M.	Peer review: The experience and views of early career researchers	<i>Learned Publishing</i>	2017	23	Autor ang.
22	Kisilowska M.	Biblioteka w sieci-sieć w bibliotece: wybrane społeczne i kulturowe aspekty współczesnego bibliotekarstwa		2010	21	Autor
23	Kulczycki E.	Wykorzystanie mediów społecznościowych przez akademickie uczelnie wyższe w Polsce. Badania w formule otwartego notatnika	<i>Komunikologia. Teoria i praktyka komunikacji</i>	2012	21	Komunikacja naukowa 18/40

1	2	3	4	5	6	7
24	Cisek S.	Metoda analizy i krytyki piśmiennictwa w nauce o informacji i bibliotekoznawstwie w XXI wieku	<i>Przegląd Biblioteczny</i>	2010	20	Nauka o informacji 1/51 Informatologia 9/34 Zarządzanie informacją 93/100 Biblioteki akademickie 60/71 Komunikacja naukowa 23/40 Bibliometria 40/42 Zachowania informacyjne 19/29
25	Kowalska M.	<i>Crowdsourcing internetowy-pozytywny wymiar partycypacji społecznej. Konteksty-istota-uwarunkowania</i>		2015	19	Informatologia 15/34 Komunikacja naukowa 24/40 Teoria informacji 83/100 Architektura informacji 22/28
26	Nicholas D., Rodríguez-Bravo B., Boukacem-Zeghmouri Ch., Rodríguez-Bravo B., Xu J., Watkinson A., Abrizah A., Herman E., Świgoń M.	So, are early career researchers the harbingers of change?	<i>Learned Publishing</i>	2019	18	Autor ang.
27	Roman W.K.	<i>Podstawy zarządzania informacją</i>		2012	18	Zarządzanie informacją 2/100
28	Gmerek K.	Archiwa internetowe po obu stronach Atlantyku—Internet Archive, Wayback Machine oraz UK Web Archive	<i>Biuletyn EBIB</i>	2012	16	Biblioteki akademickie 43/71
29	Kamińska-Czubala B.	<i>Zachowania informacyjne w życiu codziennym: informacyjny świat pokolenia Y</i>		2013	16	Zachowania informacyjne 1/29

1	2	3	4	5	6	7
30	Malak P.	<i>Indeksowanie treści: porównanie skuteczności metod tradycyjnych i automatycznych</i>		2012	15	Autor
31	Świgoń M.	Information limits: Definition, typology and types	<i>Aslib Proceedings</i>	2011	15	Bariery informacyjne 27/29
32	Wojciechowska M.	<i>Studium zarządzania niematerialnymi zasobami organizacyjnymi bibliotek</i>		2014	15	Biblioteki akademickie 63/71 Zasoby informacyjne 39/100
33	Górny M., Lewandowski R.	The state of development of digital libraries in Poland	<i>Program</i>	2010	14	Autor ang.
34	Kolasa W.M.	Specific character of citations in historiography (using the example of Polish history)	<i>Scientometrics</i>	2012	14	Autor ang.
35	Kulczycki E.	Zasady oceny czasopism humanistycznych i ich rola w parametryzacji jednostek naukowych	<i>Nauka</i>	2014	14	Czasopisma naukowe 22/100
36	Kulczycki E., Rozkosz E.A., Drabek A.	Ocena ekspercka jako trzeci wymiar ewaluacji krajowych czasopism naukowych	<i>Nauka</i>	2016	14	Bibliografia załącznikowa 14/19
37	Babik W.	O natłoku informacji i związanym z nim przeciążeniu informacyjnym	<i>Człowiek–media–edukacja</i>	2010	13	Zarządzanie informacją 33/100 Teoria informacji 8/100 Ekologia informacji 6/21 Kultura informacyjna 2/39
38	Boruszewski J.	Jakość i wiarygodność informacji w infobrokerstwie	<i>Lingua ac Communitas</i>	2012	13	Infobrokerstwo 2/11

1	2	3	4	5	6	7
39	Derfert-Wolf L.	Archiwizacja Internetu – wprowadzenie i przegląd wybranych inicjatyw	<i>Biuletyn EBIB</i>	2012	13	Biblioteki akademickie 49/71
40	Głowacka E.	Badania wartości ekonomicznej usług biblioteczno-informacyjnych i ich wpływu na otoczenie	<i>Biblioteka</i>	2011	13	Biblioteki akademickie 45/71 Zasoby informacyjne 42/100
41	Kisilowska M.	Przestrzeń informacyjna jako termin informatologiczny	<i>Zagadnienia Informacji Naukowej</i>	2011	13	Kompetencje informacyjne 22/57 Teoria informacji 67/100 Ekologia informacji 13/21 Środowisko informacyjne 1/21 Zachowania informacyjne 19/29 Działalność informacyjna 7/82 Przestrzeń informacyjna 1/44
42	Kotuła S.D.	<i>Komunikacja bibliologiczna wobec World Wide Web</i>		2013	13	Nauka o informacji 29/51 Informatologia 29/34 Komunikacja naukowa 4/40 Wyszukiwanie informacji 99/100 Teoria informacji 80/100 Obiekty informacyjne 4/5
43	Nicholas D., Rodríguez-Bravo B., Boukacem-Zeghmouri Ch., Rodríguez-Bravo B., Xu J., Watkinson A., Abrizah A., Herman E., Świgoń M.	Sci-Hub: The new and ultimate disruptor? View from the front	<i>Learned Publishing</i>	2019	13	Autor ang.

1	2	3	4	5	6	7
44	Nicholas D., Rodríguez- -Bravo B., Boukacem- -Zeghmo- uri Ch., Rodríguez- -Bravo B., Xu J., Watkinson A., Abrizah A., Herman E., Świgoń M.	Early career researchers' quest for reputation in the digital age	<i>Journal of Scholarly Publishing</i>	2018	13	Autor ang.
45	Wendland M.	Działanie komunikacyjne a przekazywanie informacji	<i>Komuniko- logia. Teoria i praktyka komunikacji</i>	2012	13	Teoria informacji 2/100
46	Cisek S.	„Metodologia mieszana” w badaniach nauki o informacji i bibliotekoznaw- stwa	<i>Biblioteki, informacja, książka: interdyscypli- narne bada- nia i praktyka w XXI wieku</i>	2010	12	Nauka o informacji 1/51 Informatologia 16/34 Użytkownicy informacji 1/25 Zachowania in- formacyjne 17/29
47	Dahlberg I.	Wissensorganisa- tion: Entwicklung, Aufgabe, Anwen- dung, Zukunft	<i>Zagadnienia Informacji Naukowej</i>	2015	12	Organizacja informacji 1/17 Bibliografia za- łącznikowa 3/19
48	Kolasa W.M.	Kierunki badań nad historią prasy polskiej 1918–1939, Cz. 1. Tendencje roz- wojowe, typologia	<i>Rocznik Historii Prasy Polskiej</i>	2011	12	Bibliometria 12/42
49	Osińska V.	<i>Wizualizacja informacji: studium informatologiczne</i>		2016	12	Wizualizacja informacji 1/23
50	Skórka S.	Architekt informacji – kreator przestrzeni informacyjnych	<i>Przegląd Biblioteczny</i>	2011	12	Architektura informacji 1/28 Zasoby informa- cyjne 4/100 Środowisko infor- macyjne 4/21 Przestrzeń infor- macyjna 1/44
51	Górny M.	Key users of Polish digital libraries	<i>The Electronic Library</i>	2012	11	Autor ang.

1	2	3	4	5	6	7
52	Hetmański M.	Świat informacji		2015	11	Teoria informacji 1/100 Procesy informacyjne 46/53
53	Kulczycki E., Drabek A., Rozkosz E.A.	Publikacje a zgłoszenia ewaluacyjne, czyli zniekształcony obraz nauki w Polsce	<i>Nauka</i>	2015	11	Czasopisma naukowe 28/100
54	Osińska V.	<i>Wizualizacja i wyszukiwanie dokumentów</i>		2010	11	Wizualizacja informacji 2/23
55	Świgoń M.	Personal Knowledge Management (PKM) and Personal Employability Management (PEM)–Concepts Based on Competences	<i>Proceedings of the 3rd European conference on Intellectual Capital. University of Nicosia, Cyprus, 18–19 April 2011</i>	2011	11	Autor ang.
56	Drabek A.	<i>Bibliometryczna analiza czasopism naukowych w dziedzinie nauk społecznych</i>		2010	10	Autor
57	Głowacka E.	<i>Kultura oceny w bibliotekach. Obszary, modele i metody badań jakości zasobów oraz usług biblioteczno-informacyjnych</i>		2015	10	Autor
58	Kolasa W.M.	Biblioteki cyfrowe na świecie – powstanie i rozwój	<i>Biblioteki cyfrowe</i>	2012	10	Komunikacja naukowa 35/40 Biblioteki cyfrowe 1/99
59	Kolasa W.M.	Retrospektywny indeks cytowań w humanistyce	<i>Przegląd Biblioteczny</i>	2011	10	Teoria informacji 85/100 Bibliometria 38/42

## Bibliografia

- Beel, J., Gipp, B., Eilde, E. (2010). Academic Search Engine Optimization. *Journal of Scholarly Publishing*, 41(2), 176–190, <https://doi.org/10.3138/jsp.41.2.176>
- Delgado López-Cózar, E., Orduna-Malea, E., Martín-Martín, A. (2019). Google Scholar as a Data Source for Research Assessment. In: W. Glänzel, H.F. Moed, U. Schmoch, M. Thelwall, (eds.), *Springer Handbook of Science and Technology Indicators*, (95–127). Springer International Publishing.
- Halevi, G., Moed, H., Bar-Ilan, J. (2017). Suitability of Google Scholar as a Source of Scientific Information and as a Source of Data for Scientific Evaluation. Review of the Literature. *Journal of Informetrics*, 11(3), 823–834, <https://doi.org/10.1016/j.joi.2017.06.005>
- Martin-Martín, A., Orduna-Malea, E., Harzing, A.-W., Delgado López-Cózar, E. (2017). Can We Use Google Scholar to Identify Highly-cited Documents? *Journal of Informetrics*, 11(1), 152–163, <https://doi.org/10.1016/j.joi.2016.11.008>
- Martin-Martín, A., Orduna-Malea, E., Delgado López-Cózar, E. (2018). A Novel Method for Depicting Academic Disciplines Through Google Scholar Citations: The Case of Bibliometrics. *Scientometrics*, 14, 1251–1273, <https://doi.org/10.1007/s11192-017-2587-4>
- Babik, W., red. (2016). *Nauka o informacji*, Warszawa: Wydaw. SBP.
- Osiński, Z. (2019). Analysis of the Thematic Overlap Between Library and Information Science and Other Sub-disciplines of the Social Communication and Media Sciences in Poland. *Zagadnienia Informacji Naukowej – Studia Informacyjne*, 57(2A), 28–44, <https://doi.org/10.36702/zin.468>
- Osiński, Z. (2019). The Usefulness of Data from Web of Science and Scopus Databases for Analyzing the State of a Scientific Discipline. The Case of Library and Information Science. *Zagadnienia Informacji Naukowej – Studia Informacyjne*, 57(2A), 45–93, <https://doi.org/10.36702/zin.469>
- Racki, G., Drabek, A. (2013). Cytowania i wskaźnik Hirscha: gdzie szukać, jak obliczać? [online]. *Forum Akademickie*, (2) [28.04.2021], <https://prenumeruj.forumakademickie.pl/fa/2013/02/cytowania-i-wskaznik-hirscha-gdzie-szukac-jak-obliczac/>
- Siłka, P., Śleszyński, P., Jaworska, B. (2016). Cytowania członków komitetów naukowych Polskiej Akademii Nauk według Google Scholar, *Zagadnienia Naukoznawstwa*, 4(210), 529–560, <https://doi.org/10.24425/118025>
- Sosińska-Kalata, B. (2013). Obszary badań współczesnej informatologii (nauki o informacji), *Zagadnienia Informacji Naukowej – Studia Informacyjne*, 51(2), 9–41, <https://doi.org/10.36702/zin.600>
- Stirbu, S., Thirion, P., Schmitz, S., Haesbroeck, G., Greco, N. (2015). The Utility of Google Scholar When Searching Geographical Literature: Comparison with Three Commercial Bibliographic Databases. *The Journal of Academic Librarianship*, 41(3), 322–329, <https://doi.org/10.1016/j.acalib.2015.02.013>
- Śleszyński, P. (2013). Cytowania i oddziaływanie polskich ośrodków geograficznych według Google Scholar. *Przegląd Geograficzny*, 85(4), 599–627, <https://doi.org/10.7163/PrzG.2013.4.5>
- Wróblewski, A.K. (2013). Pozycja nauki polskiej w międzynarodowych rankingach [online]. *Studia BAS*, 3(35), 89–106 [28.04.2021], [http://orka.sejm.gov.pl/WydBAS.nsf/0/5FA50C2A6481F-62FC1257BC200435A20/\\$file/Infos\\_152.pdf](http://orka.sejm.gov.pl/WydBAS.nsf/0/5FA50C2A6481F-62FC1257BC200435A20/$file/Infos_152.pdf)

## Using Google Scholar to Identify the Most Cited Researchers and Their Research Papers. The Case of Polish Publications in the Field of Information Science

### Abstract

**Purpose/Thesis:** The study seeks to determine whether Google Scholar allows the user to identify the most frequently cited researchers and their research papers published in Polish, and if the number of citations and the presence of keywords affect the publication's position in the ranking of the results yielded by GS. The author ranks Polish researchers and publications in the field of information science published in the period 2010–2019 to verify the hypothesis that the presence of a given publication in the search results and its position in the ranking is affected by the absence of keywords.

**Approach/Methods:** The author verifies and analyzes the data collected from Google Scholar. The sample comprises the results of forty queries, with key terms relating to information science, as well as names and surnames of researchers in the field.

**Results and conclusions:** The analysis of the sample disproves the hypothesis that Google Scholar search engine allows the user to identify the most frequently cited works and researchers. A significant part of scholarship indexed in GS does not appear in search results when the query is a phrase related to information science. Many frequently cited texts place very low in the ranking. The study also disproves the thesis that the number of citations determines a given publication's ranking. The vast majority of the bibliographic entries in the GS search results does not meet the user's needs. Use of appropriate key terms does not make it more likely that a given publication will be found.

**Originality/Value:** The work refutes the increasingly popular belief that Google Scholar can be professionally used to search for the most important (or rather, the most frequently cited) scientific papers and to rely on the number of citations as a basis for evaluating the papers published in Polish and the researchers publishing mainly in this language.

### Keywords

Google Scholar. Information science. Research citation. Search results ranking.

---

*Dr hab. ZBIGNIEW OSIŃSKI jest profesorem w Instytucie Nauk o Kulturze UMCS w Lublinie; specjalności: nauka o informacji, cyfrowa humanistyka. Najnowsze publikacje: Information Infrastructure of Contemporary Humanities and the Digital Humanities Development as a Cause of Creating New Information Barriers. A Polish case (Digital Scholarship in the Humanities, 2019), Analysis of the Thematic Overlap Between Library and Information Science and Other Sub-disciplines of the Social Communication and Media Sciences in Poland (Zagadnienia Informacji Naukowej – Studia Informacyjne, 2019), The Usefulness of Data from Web of Science and Scopus Databases for Analyzing the State of a Scientific Discipline. The Case of Library and Information Science (Zagadnienia Informacji Naukowej – Studia Informacyjne, 2019), Big data w praktyce badawczej humanistów – problemy metodyczne (Big Data w humanistyce i naukach społecznych, 2020).*

### Kontakt z autorem:

zbigniew.osinski@gmail.com

Instytut Nauk o Kulturze

Wydział Humanistyczny

Uniwersytet Marii Curie-Skłodowskiej

pl. Marii Curie-Skłodowskiej 4

20-031 Lublin