

**OTWARTA NAUKA W ŚWIETLE POLSKIEJ  
BIBLIOGRAFII NAUKOWEJ**

**EMANUEL KULCZYCKI**

**PRZEMYSŁAW KORYTKOWSKI**

**MAJ 2020**



Publikacja finansowana jest w ramach projektu „Badanie wzorców doskonałości w nauce i sztuce” w ramach programu Ministra Nauki i Szkolnictwa Wyższego pod nazwą „DIALOG” w latach 2017-2020.

## **STRONA PROJEKTU**

<http://excellence-project.zut.edu.pl>

## **JAK CYTOWAĆ RAPORT**

Kulczycki, E., Korytkowski, P. (2020). Otwarta nauka w świetle Polskiej Bibliografii Naukowej. DOI: 10.6084/m9.figshare.12204287

# SPIS TREŚCI

1. Podsumowanie i rekomendacje .....	4
2. Cel raportu i przedmiot analiz .....	6
3. Podstawy prawne raportowania metadanych o otwartym dostępie .....	7
4. Źródła danych .....	11
5. Metadane o otwartym dostępie w rekordach bibliograficznych .....	14
6. Artykuły w czasopismach ujętych w DOAJ .....	19
7. Odsetek publikacji w otwartym dostępie według typów publikacji .....	21
8. Zgodność metadanych PBN i DOAJ .....	25
9. Zgodność metadanych o otwartym miejscu publikacji .....	31
10. Poziom otwartości instytucji naukowych .....	32
11. Poziom otwartości dyscyplin naukowych .....	34

## PODSUMOWANIE I REKOMENDACJE

---

W raporcie przeanalizowano realizację idei otwartej nauki przez polskich naukowców wykorzystując 1,15 mln rekordów bibliograficznych zgłoszonych przez polskie instytucje naukowe za lata 2013–2018 do Polskiej Bibliografii Naukowej (PBN) będącej częścią Zintegrowanego Systemu Informacji o Nauce i Szkolnictwie Wyższym „POL-on”. Raport dostarcza danych na temat tego, w jakim stopniu polscy naukowcy wykorzystują możliwość publikowania wyników swoich prac naukowych w otwartym dostępie oraz o jakości zbieranych danych w ramach PBN. Raport został oparty na danych według stanu na 21 października 2019 roku.

Na podstawie rozporządzenia Ministra Nauki i Szkolnictwa Wyższego w sprawie Systemu Informacji o Nauce z 2015 roku (Dz.U. 2015 poz. 944) instytucje naukowe mają obowiązek raportowania w przypadku publikacji dostępnych w otwartym dostępie informacji o typie otwartej licencji, wersji tekstu, dacie udostępnienia, embargu oraz sposobie publikacji.

Analizy przedstawione w niniejszym raporcie pokazują, że zakres zbieranych danych w PBN odnoszący się do otwartego dostępu, rodzi problemy w trakcie zgłaszania metadanych do rekordów bibliograficznych jak i ich późniejszej analizy. Poniżej przedstawiamy najważniejsze wyzwania:

1. Porównanie danych zebranych w PBN oraz bazie Directory of Open Access Journals (DOAJ) wskazuje, że spośród rekordów bibliograficznych PBN dotyczących artykułów naukowych opublikowanych w czasopismach ujętych w DOAJ tylko w około 35% przypadków są takie same dane dotyczące licencji oraz w około 43% przypadków są takie same dane o miejscu publikacji. Wskazuje to, że instytucje naukowe mają ogromny problem z poprawnym podaniem danych w tym zakresie.
2. Rozporządzenie MNiSW wymaga podania informacji zarówno o sposobie otwartego dostępu: otwarte czasopismo, otwarte repozytorium, inne rozwiązanie oraz o otwartej wersji tekstu: oryginalna wersja autorska, ostateczna wersja autorska, ostateczna wersja opublikowana. Jednakże Rozporządzenie nie wskazuje, do której wersji tekstu ma odnosić się informacja o sposobie otwartego dostępu. Każda z wersji tekstu może mieć inne miejsce publikacji.
3. Nieprecyzyjne zapisy rozporządzenia sprawiły, że w ponad pięciu tysiącach rekordów publikacji opublikowanych w czasopismach ujętych w DOAJ podano błędnie otwarte repozytorium lub inne rozwiązanie w miejsce jedynej poprawnej opcji – otwarte czasopismo.
4. Zestawiając rekordy, w których podano, że publikacja została udostępniona w otwartym dostępie w momencie opublikowania z informacją o liczbie miesięcy, jakie upłynęły od momentu opublikowania do momentu udostępnienia publikacji w sposób otwarty, okazuje się, że jest takich ponad dwa i pół tysiąca rekordów, które jednocześnie mają moment udostępnienia wskazany jako „moment opublikowania” oraz liczbę miesięcy większą od zera, co jest sprzeczne.

5. W przeważającej większości wskazano, że wersja tekstu będąca w otwartym dostępie jest ostateczną opublikowaną wersją. W przypadku ponad pięciu tysięcy rekordów wskazano, że wersją tekstu będącą w otwartym dostępie jest „ostateczna wersja opublikowana”, lecz jednocześnie wskazano, że została ona udostępniona przed publikacją, co nie jest możliwe.
6. Brak w PBN fazy łączenia danych o tych samych publikacjach zgłaszanych przez różne instytucje naukowe powoduje, że nie ma możliwości przeprowadzenia analiz na poziomie rzeczywistych publikacji. Z konieczności w niniejszym raporcie wykorzystywane jest pojęcie rekordu bibliograficznego. Około połowa rekordów bibliograficznych w PBN dotyczy publikacji wielośrodkowych, tj. takich których autorzy pochodzą z więcej niż jednej instytucji naukowej. Jeśli autorzy tych publikacji pochodzą z kilku krajowych instytucji, to PBN zawiera dla tej samej publikacji kilka praktycznie niemożliwych do połączenia rekordów, jeśli każdy z nich nie zawiera unikatowego identyfikatora publikacji (np. DOI lub ISBN). Z tego powodu liczba rzeczywistych publikacji w otwartym dostępie opublikowana przez polskich naukowców jest o nieznany procent niższa od liczby rekordów bibliograficznych w PBN.

### **Rekomendacje dotyczące Polskiej Bibliografii Naukowej:**

1. Należy przebudować PBN w taki sposób, aby była możliwa analiza danych na poziomie rzeczywistych publikacji naukowych. Wymaga to stworzenia mechanizmu uzgadniania rekordów dotyczących publikacji wielośrodkowych. Dzięki wdrożeniu tego mechanizmu będzie można po raz pierwszy dowiedzieć się, ile i gdzie tak naprawdę publikują polscy naukowcy oraz określić poziom naukowej współpracy krajowej.
2. Należy wymagać dodatkowej informacji jednoznacznie wskazującej, czy dana publikacja jest dostępna w trybie otwartego czy zamkniętego dostępu. Obecnie brak informacji o otwartości jest jedynie domyślnie traktowany jako publikacja w trybie zamkniętego dostępu.
3. Należy powiązać PBN z DOAJ w celu zapewnienia spójności danych.
4. Struktura danych w PBN powinna być kontekstowa. Oznacza to, że nie wszystkie kombinacje wprowadzanych danych powinny być dozwolone, np. rekord odnoszący się do artykułu opublikowanego w otwartym czasopiśmie nie powinien dawać możliwości wskazania otwartego repozytorium jako miejsca publikacji.
5. Ośrodek Przetwarzania Informacji powinien współpracować z grupą akademickich bibliotekarzy, którzy podejmowaliby merytoryczne decyzje dotyczące nadsyłanych danych oraz wprowadzali niezbędne korekty. Dla przykładu: w przypadku czasopism naukowych PBN powinien korzystać z zamkniętego słownika, który byłby uzupełniany tylko w razie konieczności przez właśnie takich bibliotekarzy.

Chociaż polskie instytucje naukowe podejmują trud gromadzenia i raportowania danych, to obraz otwartości tych instytucji, jaki się wyłania, analizując tylko dane zapisane w PBN, nie prezentuje ich w adekwatnym świetle. Dlatego też wytyczne i regulacje prawne w zakresie gromadzenia danych warto byłoby ulepszyć, aby praca instytucji i naukowców została w odpowiedni sposób zauważona i doceniona.

## CEL RAPORTU I PRZEDMIOT ANALIZ

---

Celem raportu jest prezentacja wzorów publikacyjnych polskich naukowców z zakresie publikowania w otwartym dostępie. Adresatami raportu są naukowcy oraz specjaliści zajmujący się badaniami nad nauką i szkolnictwem wyższym, którym raport może służyć jako źródło informacji o praktykach publikacyjnych.

**Wzory publikacyjne** to regularności w praktykach publikacyjnych obserwowane w wybieranych przez naukowców typach publikacji (np. artykuły, monografie, prace pod redakcją), czy też właśnie modelu publikacji (otwarty, zamknięty). Wzory publikacyjne mogą być porównywane pomiędzy dyscyplinami (np. jak publikują psychologowie, a jak historycy) czy modelami publikacji (jaki odsetek artykułów w otwartym dostępie publikują fizycy a jaki informatycy).

W niniejszym raporcie uwzględnione zostały publikacje polskich naukowców zgłoszone do Modułu Sprawozdawczego Polskiej Bibliografii Naukowej (PBN) przez ich pracodawców, tj. instytucje naukowe (uczelnie, instytuty Polskiej Akademii Nauk, instytuty badawcze). Zgłaszanie wszystkich publikacji z analizowanego w raporcie okresu, tj. lat 2013–2018, było ustawowym obowiązkiem instytucji naukowych.

W raporcie dane są zagregowane na poziomie instytucji naukowych oraz dziedzin i dyscyplin na podstawie przypisania naukowców do dyscyplin naukowych wykonanego jesienią 2018 roku. Jednostką analizy jest rekord bibliograficzny odpowiadający jednej publikacji zgłoszonej do PBN przez instytucję naukową.

## PODSTAWY PRAWNE RAPORTOWANIA METADANYCH O OTWARTYM DOSTĘPIE

---

Obowiązek przesyłania danych do Polskiej Bibliografii Naukowej (<https://pbn.nauka.gov.pl>) jako elementu sprawozdawczości instytucji naukowych pojawił się wraz z wprowadzeniem następujących aktów prawnych:

1. Nowelizacja Ustawy z dnia 30 kwietnia 2010 roku o zasadach finansowania nauki wprowadzonej Ustawą z dnia 15 stycznia 2015 roku o zmianie ustawy o zasadach finansowania nauki oraz niektórych innych ustaw (Dz.U. 2015 poz. 249).
2. Rozporządzenie Ministra Nauki i Szkolnictwa Wyższego z dnia 29 czerwca 2015 roku w sprawie Systemu Informacji o Nauce (Dz.U. 2015 poz. 944).

W wyniku nowelizacji dodano do ustawy zapis o prowadzeniu Systemu Informacji o Nauce, którego częścią jest Polska Bibliografia Naukowa:

*Art. 4c. 1. Minister prowadzi System Informacji o Nauce w ramach Zintegrowanego Systemu Informacji o Nauce i Szkolnictwie Wyższym „POL-on”, zwany dalej „Systemem”.*

W ustawie określono szczegółowy zakres informacji, które obejmował system. Odnośnie do publikacji ustawa mówiła o:

*2) publikacjach w czasopismach naukowych oraz monografiach naukowych lub w rozdziałach w monografiach naukowych autorstwa pracowników jednostki naukowej oraz publikacjach osób niebędących pracownikami jednostki, które afilowały te publikacje do tej jednostki.*

Ustawa zobowiązywała wszystkie instytucje naukowe (jednostki naukowe według wskazanej ustawy i rozporządzenia) posiadające kategorię naukową i starające się o środki naukowe na działalność statutową do systematycznego wprowadzania danych do Systemu Informacji o Nauce. Odpowiedzialnym za tę czynność czyniła kierownika instytucji naukowej.

*6. Jednostki naukowe, które posiadają kategorię naukową i ubiegają się o środki finansowe na zadania, o których mowa w art. 18 ust. 1, są obowiązane do wprowadzania do Systemu oraz systematycznego aktualizowania danych i informacji, o których mowa w ust. 3 i 10. Za prawidłowość, rzetelność i terminowość danych i informacji wprowadzanych do Systemu odpowiada kierownik jednostki naukowej.*

Ustęp 3 obejmował dane o działalności badawczo-rozwojowej i jej efektach, w tym ww. dane o publikacjach.

Ustawa precyzowała, kiedy dane mają być wprowadzone i za jaki okres:

*Art. 9. Jednostki naukowe (...) wprowadzą do Systemu Informacji o Nauce dane i informacje (...) w terminie 3 miesięcy od dnia wejścia w życie niniejszej ustawy. Dane i informacje wprowadza się za okres od dnia 1 stycznia 2013 r.*

Ustawa weszła w życie po upływie 3 miesięcy od dnia ogłoszenia (tj. 2015-07-03). Po raz pierwszy dane o publikacjach miały być przekazane do 3 października 2015 roku i miały dotyczyć okresu od 1 stycznia 2013 roku.

Rozporządzenie określało:

1. Szczegółowy zakres danych i informacji objętych Systemem Informacji o Nauce (podany w załączniku 1 do rozporządzenia).
2. Terminy aktualizacji danych w Systemie.
3. Sposób i formę udostępniania informacji zgromadzonych w Systemie.

W załączniku określono, jakie informacje powinny być dostarczone do Systemu. Nie wyodrębniono w tej grupie informacji obowiązkowych i fakultatywnych. Oznaczało to, że w przypadku publikacji instytucje miały obowiązek wprowadzić wszystkie dane mające zastosowanie do danej publikacji.

Informacje o otwartym dostępie były określone odrębnie dla publikacji w czasopismach naukowych oraz monografiach naukowych. Oczekiwano od instytucji naukowych podania następujących danych: sposobu udostępnienia, wersji tekstu, rodzaj licencji, daty udostępnienia publikacji w sposób otwarty. Szczegółowe dane podano odrębnie dla publikacji w czasopismach naukowych, monografiach naukowych (monografiach i monografiach wieloautor-skich) oraz rozdziałów.

Publikacje w czasopismach naukowych:

*i) informacje o otwartym dostępie do publikacji (Open Access): (otwarte czasopismo, otwarte repozytorium, inne):*

- *wersja tekstu, która jest dostępna w sposób otwarty (oryginalna wersja autorska, ostateczna wersja autorska, ostateczna wersja opublikowana),*
- *otwarta licencja, na jakiej publikacja jest dostępna (Creative Commons – Uznanie Autorstwa (CC BY); Creative Commons – Uznanie Autorstwa – Na Tych Samych Warunkach (CC BY-SA); Creative Commons – Uznanie Autorstwa – Użycie niekomercyjne (CC BY-NC); Creative Commons – Uznanie Autorstwa – Bez utworów zależnych (CC BY-ND); Creative Commons – Uznanie Autorstwa – Użycie niekomercyjne – Na tych samych warunkach (CC BY-NC-SA); Creative Commons – Uznanie Autorstwa – Użycie niekomercyjne – Bez utworów zależnych (CC BY-NC-ND); inna otwarta licencja,*
- *data udostępnienia publikacji w sposób otwarty (przed opublikowaniem; w momencie opublikowania; po opublikowaniu – wraz z podaniem liczby miesięcy, jakie upłynęły od momentu opublikowania do momentu udostępnienia publikacji w sposób otwarty).*



## Monografie naukowe [Książki]:

### *informacje o otwartym dostępie do monografii (Open Access):*

- *sposób, w jaki monografia została udostępniona (witryna wydawcy, otwarte repozytorium, inne),*
- *wersja tekstu, która jest dostępna w sposób otwarty (oryginalna wersja autorska, ostateczna wersja autorska, ostateczna wersja opublikowana),*
- *otwarta licencja, na jakiej monografia jest dostępna (CC BY; CC BY-SA; CC BY-NC; CC BY-ND; CC BY-NC-SA; CC BY-NC-ND; inna otwarta licencja,*
- *informacja o momencie udostępnienia monografii w sposób otwarty (przed opublikowaniem; w momencie opublikowania; po opublikowaniu – wraz z podaniem liczby miesięcy, jakie upłynęły od momentu opublikowania do momentu udostępnienia publikacji w sposób otwarty).*

## Rozdziały:

### *informacje o otwartym dostępie do publikacji (Open Access):*

- *sposób, w jaki rozdział został udostępniony (witryna wydawcy, otwarte repozytorium, inne),*
- *wersja tekstu, która jest dostępna w sposób otwarty (oryginalna wersja autorska, ostateczna wersja autorska, ostateczna wersja opublikowana),*
- *otwarta licencja, na jakiej rozdział jest dostępny (CC BY; CC BY-SA; CC BY-NC; CC BY-ND; CC BY-NC-SA; CC BY-NC-ND; inna otwarta licencja,*
- *informacja o momencie udostępnienia rozdziału w sposób otwarty (przed opublikowaniem, w momencie opublikowania, po opublikowaniu – wraz z podaniem liczby miesięcy, jakie upłynęły od momentu opublikowania do momentu udostępnienia publikacji w sposób otwarty).*

Zapisy Rozporządzenia z 25 czerwca 2015 roku zostały przełożone na strukturę danych, która jest przyjmowana od instytucji naukowych przez Polską Bibliografię Naukową. Schemat danych w formacie XSD udostępniony został instytucjom naukowym i w zakresie otwartego dostępu przyjął formę przedstawioną w Tabeli 1.

**Tabela 1.** Schemat przyjmowanych przez PBN danych w zakresie otwartej nauki, stan na 21 października 2019 roku.

Element	Element podrzędny	Krotność	Opis
open-access		0..1	Informacja czy publikacja jest dostępna w "otwartym dostępie" (Open Access).
open-access	open-access-text-version	0..1	Wersja tekstu, która jest dostępna w sposób otwarty. Wartości słownikowe <sup>1</sup> : „ORIGINAL_AUTHOR” „FINAL_AUTHOR” „FINAL_PUBLISHED”
open-access	open-access-license	0..1	Otwarta licencja, na jakiej publikacja jest dostępna. Wartości słownikowe: „CC BY” „CC BY-SA” „CC BY-NC” „CC BY-ND” „CC BY-NC-SA” „CC BY-NC-ND” „OTHER”
open-access	open-access-release-time	0..1	Moment udostępnienia publikacji w sposób otwarty. Wartości słownikowe: „BEFORE_PUBLICATION” „AT_PUBLICATION” „AFTER_PUBLICATION”
open-access	open-access-months	0..1	Ilość miesięcy, które upłynęły od momentu opublikowania do momentu udostępnienia <sup>2</sup> .
open-access	open-access-mode	0..1	Sposób, w jaki publikacja została udostępniona. Wartości słownikowe <sup>3</sup> : „OPEN_JOURNAL” „OPEN_REPOSITORY” „OTHER”

<sup>1</sup> W bazie występuje w 34 rekordach nieujęta w słowniku wartość „OTHER”.

<sup>2</sup> Schemat danych zakłada, że pole powinno zawierać liczbę naturalną, jednak w danych występują trzy rekordy z wartościami ujemnymi.

<sup>3</sup> W bazie występuje w 29 721 rekordach wartość nieujęta w słowniku „PUBLISHER\_WEBSITE”.

## ŹRÓDŁA DANYCH

---

Dane z Polskiej Bibliografii Naukowej oraz dane o dyscyplinach i dziedzinach reprezentowanych przez polskich naukowców zostały udostępnione dzięki uprzejmości Ministerstwa Nauki i Szkolnictwa Wyższego przez Ośrodek Przetwarzania Informacji. Oba źródła zawierają dane według stanu na 21 października 2019 roku. Dodatkowo wykorzystane zostały dane z bazy ARIANTA dzięki uprzejmości dr Anety Drabek, według stanu danych na 29 grudnia 2019 roku oraz dane z Directory Open Access Journals, według stanu danych na 19 listopada 2019 roku.

### 4.1. Polska Bibliografia Naukowa

Analizowany w raporcie zbiór danych z Polskiej Bibliografii Naukowej zawiera 1 201 054 rekordy bibliograficzne publikacji z lat 2013–2019. Należy zauważyć, że dane były raportowane na poziomie podstawowych jednostek naukowych (np. wydziałów uczelni czy instytutów Polskiej Akademii Nauk), dlatego liczba 1,2 mln zawiera w sobie duplikaty publikacji w przypadku, gdy jedna wieloautorska publikacja została zaraportowana przez więcej niż jedną jednostkę (instytucje naukową). Każda instytucja zgłasza do PBN informacje niezależnie od siebie. PBN nie przewiduje fazy łączenia i uzgadniania, jak ma to miejsce w Finlandii czy Norwegii<sup>4</sup>. Często tytuły czasopism czy książek dla tych samych publikacji nie są tożsame, zdarzają się błędy w polach zawierających identyfikatory ISBN i ISSN. PBN w zakresie informacji o autorach wymaga tylko informacji o autorach zatrudnionych w raportującej instytucji oraz łączną liczbę autorów. Wszystko to oznacza, że nie ma możliwości automatycznego powiązania rekordów dotyczących tych samych publikacji, jeśli wszystkie rekordy odnoszące się do danej publikacji nie mają przypisanego identycznego identyfikatora na poziomie publikacji, np. DOI. W PBN jest 687 868 rekordów, które biorąc pod uwagę liczbę zadeklarowanych autorów, mogą być zduplikowane. Z tego powodu pomimo ogromnego wysiłku wkładanego przez instytucje naukowe w raportowanie danych, nie ma możliwości uzyskania z PBN informacji o liczbie rzeczywistych publikacji polskich naukowców na poziomie całego kraju. Wszystkie statystyki, również zawarte w niniejszym raporcie, są z tego powodu zaburzone i z konieczności muszą operować na poziomie rekordów bibliograficznych PBN, a nie jak należałoby oczekiwać na poziomie publikacji naukowych.

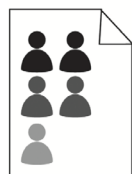
Poniższy rysunek prezentuje sytuację, w której jedna publikacja napisana przez pięciu polskich naukowców jest zaraportowana przez trzy instytucje (jednostki naukowe), w wyniku czego PBN zawiera trzy rekordy bibliograficzne dotyczące tej samej publikacji.

---

<sup>4</sup> CRISTIN, Reporting instructions (NVI), <https://www.cristin.no/english/resources/reporting-instructions/>, data dostępu: 25 kwietnia 2020 roku.

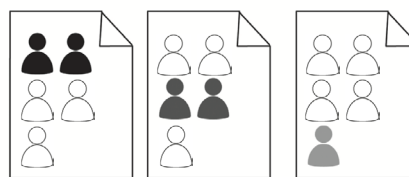
## Jedna publikacja

Pięcioro autorów z trzech różnych polskich instytucji



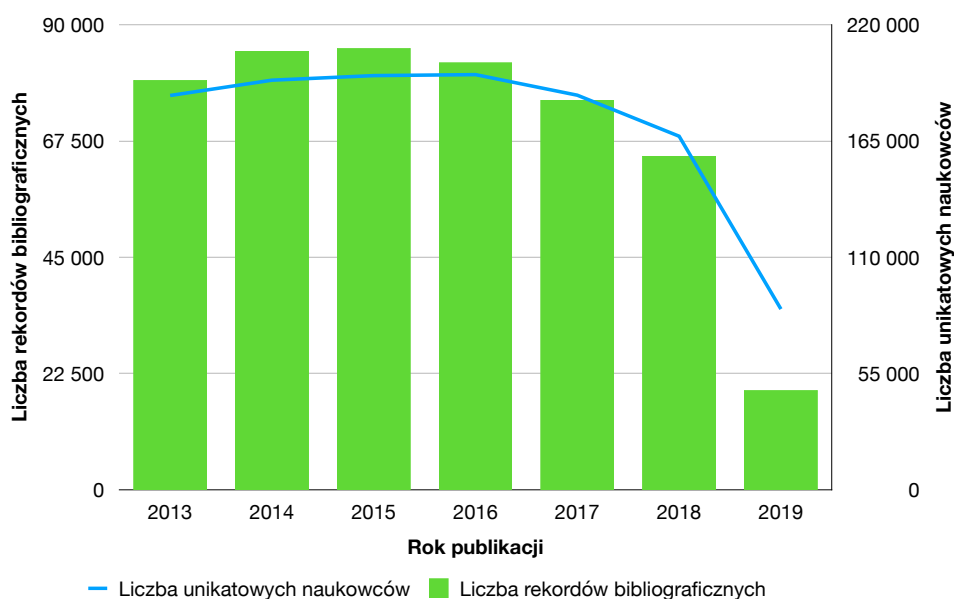
## Trzy rekordy bibliograficzne w PBN

Publikacja zaraportowana przez trzy instytucje



**Rysunek 1.** Publikacja a rekordy bibliograficzne w PBN.

Zbiór danych z PBN zawiera tylko częściowe dane za 2019 roku. Z tego powodu w dalszych analizach wykorzystane są dane tylko za lata 2013–2018, czyli 1 154 182 rekordy bibliograficzne. Rysunek 2 pokazuje liczbę rekordów (zielone słupki) zaraportowanych przez polskie instytucje naukowe oraz liczbę unikatowych naukowców (niebieska linia), których publikacje zostały zaraportowane przez instytucje. Zmniejszenie liczby publikacji w 2017 roku oraz 2018 roku może być spowodowany poluźnieniem reżimu sprawozdawczego przez instytucje oraz osiągnięciem przez naukowców z tych instytucji tzw. górki publikacyjnej w 2016 roku w związku z zamknięciem czteroletniego okresu ewaluacyjnego za lata 2013–2016.



**Rysunek 2.** Liczba rekordów bibliograficznych oraz unikatowych naukowców z polskich instytucji będących autorami tych publikacji w Polskiej Bibliografii Naukowej.

## 4.2. Dane o naukowcach

Dane o naukowcach wykorzystane w niniejszym raporcie pochodzą z państwowego rejestru prowadzonego przez Ośrodek Przetwarzania Informacji dla Ministerstwa Nauki i Szkolnictwa Wyższego, dlatego przypisanie naukowców do dyscyplin jest wiarygodne.

W zbiorze danych o reprezentowanych dyscyplinach znajduje się 106 978 rekordów odnoszących się do 91 746 unikatowych naukowców, z których 14 819 ma przypisane co najmniej dwie dyscypliny.

W niniejszym raporcie, aby przeprowadzić analizy na poziomie dziedzin i dyscypliny, wzięto pod uwagę publikacje naukowców, którzy przypisali się w 100% do jednej dyscypliny. Oznacza to, że rekordy bibliograficzne naukowców, którzy reprezentują dwie dyscypliny nie są uwzględniane w analizach prowadzonych na poziomie dziedzin i dyscyplin, gdyż w Polskiej Bibliografii Naukowej nie było – na moment otrzymania danych do analiz – przypisani rekordów bibliograficznych do dyscyplin.

W analizach, gdy prezentowane są dane zagregowane na poziomie dziedzin i dyscyplin, uwzględniono rekordy publikacyjne 60 141 naukowców, którzy:

- w latach 2013–2018 opublikowali co najmniej jedną publikację zareportowaną w PBN-ie oraz,
- zadeklarowało swoją pracę w jednej dyscyplinie (100%) oraz,
- w 2019 roku są zatrudnieni w jednym podmiocie (wykluczamy pracowników dwuetatowych, którzy zadeklarowali dwie różne dyscypliny – po jednej dyscyplinie w każdym z dwóch podmiotów).

## 4.3. Directory of Open Access Journals

W raporcie wykorzystano dane z Directory of Open Access Journals (DOAJ) aktualne na 19 listopada 2019 roku. Zbiór danych obejmuje 14 057 czasopism publikujących w otwartym dostępie i zawiera informację o tytule, ISSN, eISSN, używanej licencji oraz szczegółowe informacje związane z polityką otwartości, w tym o pobieranych opłatach.

Zetawienie danych z DOAJ z PBNem zostało oparte o połączenie tytułu czasopisma lub ISSN-u lub eISSN. W ten sposób udało się połączyć 1013 czasopism z DOAJ z rekordami bibliograficznymi artykułów opublikowanych przez polskich naukowców. Jakość wprowadzonych przez instytucje naukowe danych do PBN mogła sprawić, że mapowanie danych z dwóch baz nie jest idealne.

W analizie założono, że czasopismo ujęte w DOAJ publikuje artykuły w otwartym dostępie w całym analizowanym przez nas okresie, tj. latach 2013–2018.

## METADANE O OTWARTYM DOSTĘPIE W REKORDACH BIBLIOGRAFICZNYCH

Największy odsetek rekordów bibliograficznych w otwartym dostępie odnosi się do artykułów opublikowanych w czasopismach. Ponad 25% takich rekordów zawiera informację o sposobie, w jaki publikacja została udostępniona w otwartym dostępie. W przypadku książek jest to zaledwie 4%, a w przypadku rozdziałów 5%.

Poniższa tabela zestawia liczbę rekordów bibliograficznych z lat 2013–2018 wraz z informacją o wskazanych metadanych odnoszących się do otwartego dostępu w podziale na typy publikacji.

*Tabela 2. Liczba rekordów o otwartym dostępie w Polskiej Bibliografii Naukowej.*

	Liczba rekordów	Sposób, w jaki publikacja została udostępniona	Otwarta licencja	Embargo <sup>5</sup>	Informacja o momencie udostępnienia w sposób otwarty	Wersja tekstu, która jest dostępna w sposób otwarty
Artykuły	697 722	185 964	193 490	27 172	187 015	191 625
Książki	74 193	3 120	3 070	773	3 018	3 177
Rozdziały	382 267	19 829	21 328	4 625	19 362	20 075
<b>Łącznie</b>	<b>1 154 182</b>	<b>208 913</b>	<b>217 888</b>	<b>32 570</b>	<b>209 395</b>	<b>214 877</b>

Najczęściej wskazywanym sposobem udostępnienia publikacji jest otwarte czasopismo, jednakże dla ponad 15 tysięcy rekordów odnoszących się do artykułów wskazano otwarte repozytorium. Witryna wydawcy została wskazana jedynie w 3% przypadków.

*Tabela 3. Liczba rekordów o sposobie, w jaki publikacja została udostępniona w Polskiej Bibliografii Naukowej.*

Informacje o otwartym dostępie do publikacji	2013	2014	2015	2016	2017	2018	Łącznie
<b>Otwarte czasopismo</b>	17 173	18 347	20 473	23 118	26 465	25 740	131 316
<b>Otwarte repozytorium</b>	3 197	3 310	3 620	3 356	3 418	3 350	20 251
Artykuł	2 509	2 589	2 843	2 537	2 612	2 611	15 701
Książka	124	115	110	157	93	96	695
Rozdział	564	606	667	662	713	643	3 855

<sup>5</sup> Liczba miesięcy, jakie upłynęły od momentu opublikowania do momentu udostępnienia publikacji w sposób otwarty.

Informacje o otwartym dostępie do publikacji	2013	2014	2015	2016	2017	2018	Łącznie
<b>Inne</b>	9 253	9 751	9 860	8 597	7 191	6 370	51 022
Artykuł	7 517	7 595	7 406	6 332	5 157	4 940	38 947
Książka	314	310	363	230	130	91	1 438
Rozdział	1 422	1 846	2 091	2 035	1 904	1 339	10 637
<b>Witryna wydawcy</b>	803	929	948	1 166	1 169	1 309	6 324
Książka	146	162	162	183	138	196	987
Rozdział	657	767	786	983	1 031	1 113	5 337
<b>Łącznie</b>	<b>30 426</b>	<b>32 337</b>	<b>34 901</b>	<b>36 237</b>	<b>38 243</b>	<b>36 769</b>	<b>208 913</b>

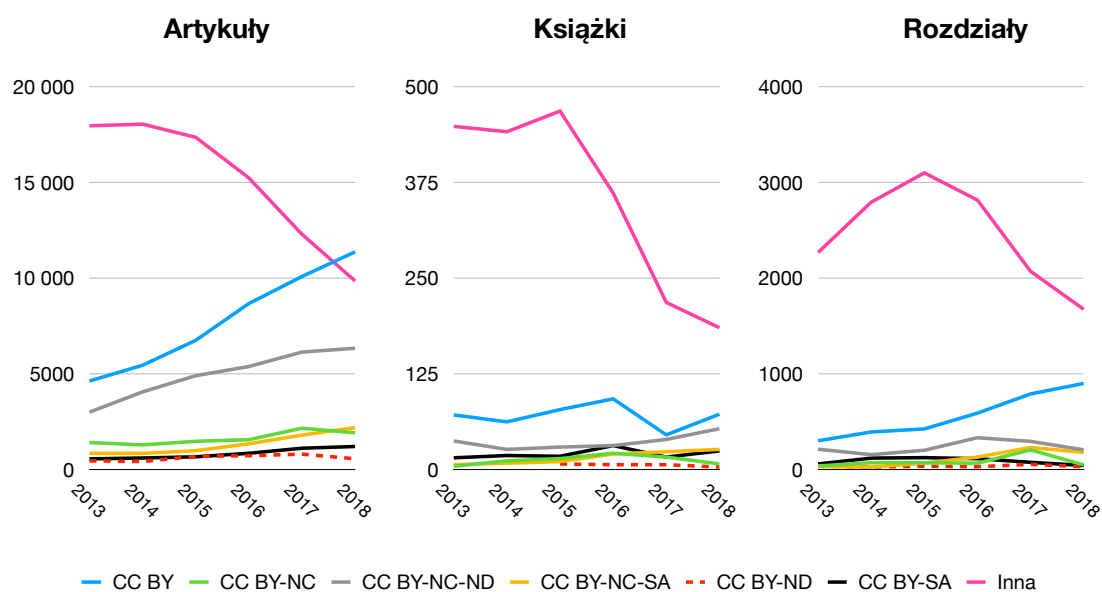
W blisko połowie przypadków wskazywano inną otwartą licencję, a w jednej czwartej przypadków wskazano licencję CC BY.

*Tabela 4. Liczba rekordów o rodzaju otwartej licencji w Polskiej Bibliografii Naukowej.*

Rodzaj licencji	2013	2014	2015	2016	2017	2018	Łącznie
<b>CC BY</b>	4 983	5 889	7 239	9 340	10 908	12 331	50 690
Artykuł	4 614	5 437	6 739	8 661	10 075	11 362	46 888
Książka	71	62	78	92	45	72	420
Rozdział	298	390	422	587	788	897	3 382
<b>CC BY-NC</b>	1 438	1 356	1 546	1 632	2 367	1 962	10 301
Artykuł	1 399	1 276	1 464	1 547	2 148	1 909	9 743
Książka	4	11	14	21	16	7	73
Rozdział	35	69	68	64	203	46	485
<b>CC BY-NC-ND</b>	3 234	4 223	5 119	5 733	6 457	6 582	31 348
Artykuł	2 987	4 044	4 891	5 372	6 126	6 326	29 746
Książka	37	26	29	31	39	53	215
Rozdział	210	153	199	330	292	203	1 387
<b>CC BY-NC-SA</b>	871	863	1 048	1 479	2 038	2 377	8 676
Artykuł	835	832	973	1 331	1 787	2 174	7 932
Książka	6	8	10	20	23	26	93
Rozdział	30	23	65	128	228	177	651
<b>CC BY-ND</b>	476	430	714	744	851	579	3 794
Artykuł	440	406	675	711	793	549	3 574
Książka	5		7	6	6	3	27
Rozdział	31	24	32	27	52	27	193

Rodzaj licencji	2013	2014	2015	2016	2017	2018	Łącznie
<b>CC BY-SA</b>	617	726	792	979	1 190	1 256	5 560
Artykuł	546	591	653	837	1 101	1 192	4 920
Książka	15	18	17	31	16	24	121
Rozdział	56	117	122	111	73	40	519
<b>Inna otwarta licencja</b>	20 664	21 268	20 918	18 402	14 570	11 697	107 519
Artykuł	17 950	18 035	17 352	15 227	12 283	9 840	90 687
Książka	448	441	468	361	218	185	2 121
Rozdział	2 266	2 792	3 098	2 814	2 069	1 672	14 711
<b>Łącznie</b>	<b>32 283</b>	<b>34 755</b>	<b>37 376</b>	<b>38 309</b>	<b>38 381</b>	<b>36 784</b>	<b>217 888</b>

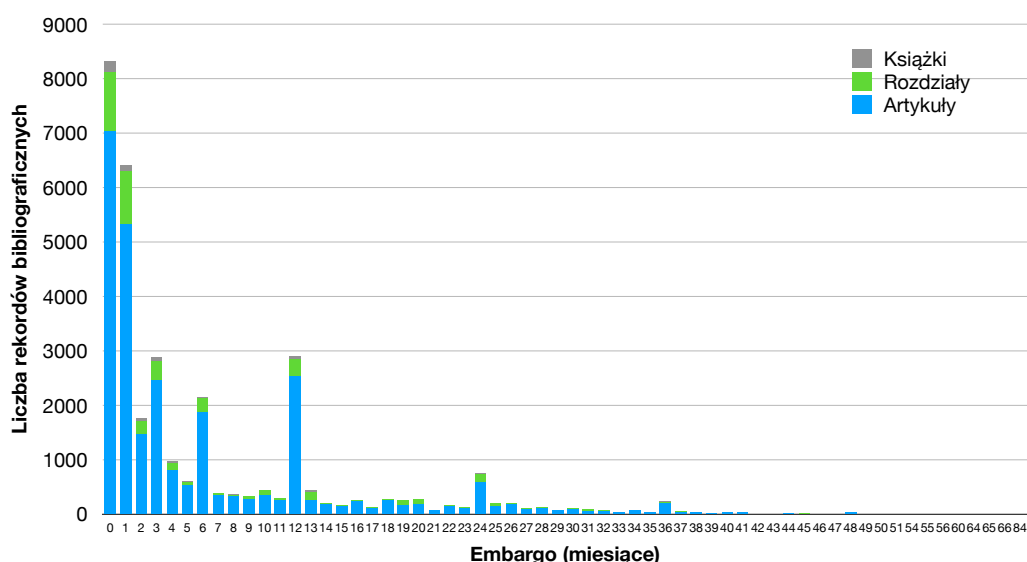
Poniższy rysunek pokazują dane z powyższej tabeli w innym przekroju, tj. według typów publikacji oraz wskazywanych licencji. Liczba wskazywanych licencji CC BY rośnie a liczba innych licencji niż CC maleje w przypadku każdego typu publikacji.



**Rysunek 3.** Liczba rekordów w Polskiej Bibliografii Naukowej z informacją o licencji.

Poniższy rysunek pokazuje liczbę rekordów bibliograficznych z lat 2013–2018 w podziale na typy publikacji, którym została przypisana metadana o liczba miesięcy, jakie upłynęły od momentu opublikowania do momentu udostępnienia publikacji w sposób otwarty. Najczęściej wskazaną informacją jest zero miesięcy, jeden miesiąc i 12 miesięcy.





Rysunek 4. Liczba rekordów w Polskiej Bibliografii Naukowej z informacją o embargo.

W ponad 75% przypadków wskazano, że publikacja została udostępniona w otwartym dostępie w momencie opublikowania. Zestawiając tę informację z metadana o liczbie miesięcy, jakie upłynęły od momentu opublikowania do momentu udostępnienia publikacji w sposób otwarty, okazuje się, że 2626 rekordów ma wskazany jako moment udostępnienia „moment opublikowania” i jednocześnie liczbę miesięcy większą od zera, co jest sprzeczne.

Tabela 5. Liczba rekordów o momencie udostępnienia publikacji w Polskiej Bibliografii Naukowej.

Moment udostępnienia publikacji	2013	2014	2015	2016	2017	2018	Łącznie
<b>Po publikacji</b>	7 189	7 314	7 251	6 900	6 609	5 340	40 603
Artykuł	6 183	6 300	6 215	6 053	5 779	4 885	35 415
Książka	142	141	145	127	61	58	674
Rozdział	864	873	891	720	769	397	4 514
<b>W momencie opublikowania</b>	20 946	22 521	25 088	26 858	33 736	33 471	162 620
Artykuł	19 057	20 314	22 390	23 900	30 211	30 097	145 969
Książka	366	343	407	400	334	369	2 219
Rozdział	1 523	1 864	2 291	2 558	3 191	3 005	14 432
<b>Przed publikacją</b>	815	999	1 048	1 209	981	1 120	6 172
Artykuł	741	871	972	1 093	899	1 055	5 631
Książka	21	41	20	17	15	11	125
Rozdział	53	87	56	99	67	54	416
<b>Łącznie</b>	<b>28 950</b>	<b>30 834</b>	<b>33 387</b>	<b>34 967</b>	<b>41 326</b>	<b>39 931</b>	<b>209 395</b>

W przeważającej większości wskazano, że wersja tekstu będąca w otwartym dostępie jest ostateczną opublikowaną wersją. Jednakże dla 5139 rekordów wskazano, że ostateczna wersja opublikowana została udostępniona przed publikacją.

*Tabela 6. Liczba rekordów o wersji tekstu, która jest dostępna w sposób otwarty w Polskiej Bibliografii Naukowej.*

Wersja publikacji	2013	2014	2015	2016	2017	2018	Łącznie
<b>Ostateczna wersja autorska</b>	767	766	859	864	1 444	1 588	6 288
Artykuł	693	692	757	771	1 239	1 445	5 597
Książka	17	9	16	9	32	32	115
Rozdział	57	65	86	84	173	111	576
<b>Ostateczna wersja opublikowana</b>	28 754	30 563	32 670	33 474	39 396	38 064	202 921
Artykuł	25 754	27 225	28 896	29 665	35 110	34 296	180 946
Książka	542	531	574	536	406	422	3 011
Rozdział	2 458	2 807	3 200	3 273	3 880	3 346	18 964
<b>Oryginalna wersja autorska</b>	567	748	1 097	1 355	983	917	5 667
Artykuł	506	647	993	1 198	897	840	5 081
Książka	5	15	14	11	3	3	51
Rozdział	56	86	90	146	83	74	535
<b>Inna<sup>6</sup></b>						1	1
Artykuł						1	1
<b>Łącznie</b>	30 088	32 077	34 626	35 693	41 823	40 570	214 877

Metadane o otwartym dostępie publikacji analizowane krzyżowo pokazują, że występują błędne zestawienia, które powinny być niemożliwe do zaraportowania (np. udostępnienie w otwartym dostępie ostatecznej wersji publikacji przed jej publikacją).

<sup>6</sup> Według regulacji prawnych nie można było wskazać „innej” wersji.

## ARTYKUŁY W CZASOPISMACH UJĘTYCH W DOAJ

Polscy naukowcy opublikowali swoje artykuły w 1013 czasopismach ujętych w DOAJ w latach 2013–2018, z czego 169 czasopism ma przypisanych co najmniej 100 rekordów bibliograficznych.

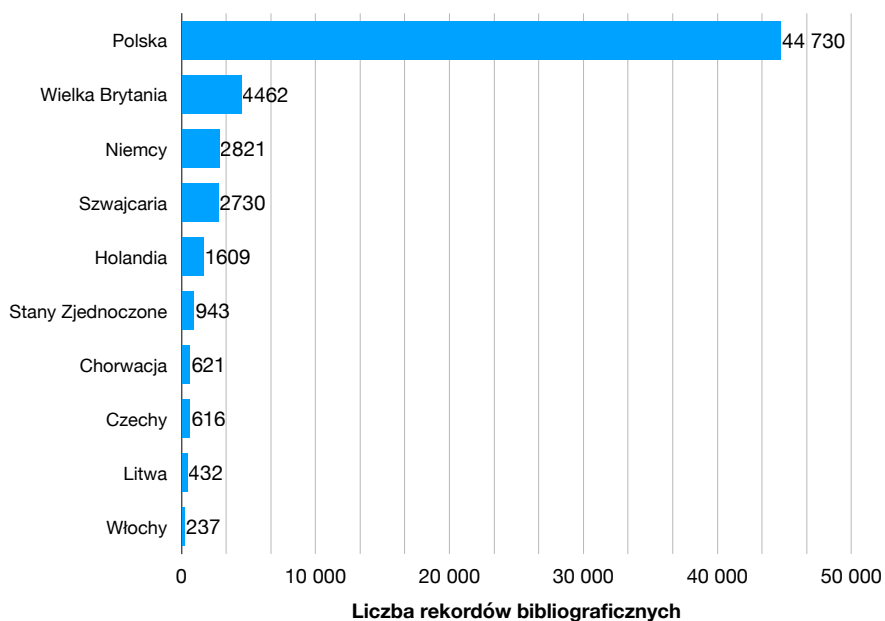
Poniższa tabela prezentuje czasopisma ujęte w DOAJ z największą liczbą rekordów bibliograficznych w PBN z lat 2013–2018.

*Tabela 7. Najczęściej występujące czasopisma w Polskiej Bibliografii Naukowej ujęte w DOAJ.*

Tytuł czasopisma	Liczba rekordów bibliograficznych w PBN
Journal of Education, Health and Sport	1718
Journal of High Energy Physics	1472
Scientific Papers of Silesian University of Technology. Organization and Management Series	1304
Physics Letters B	1216
Annals of Agricultural and Environmental Medicine	1088
European Physical Journal C: Particles and Fields	999
Postępy Higieny i Medycyny Doświadczalnej	892
Molecules	892
Ginekologia Polska	864
BioMed Research International	821
Archives of Foundry Engineering	807
Journal of Civil Engineering, Environment and Architecture	766
Archives of Medical Science	689
Endokrynologia Polska	645
Acta Universitatis Lodziensis. Folia Oeconomica	633
Advances in Dermatology and Allergology	568
Przegląd Spawalnictwa	560
Family Medicine & Primary Care Review	542
Inżynieria Ekologiczna	477
Przegląd Dermatologiczny	445
Medycyna Pracy	430
Budownictwo i Architektura	412
Folia Cardiologica	398
Materials	395
Cardiology Journal	381

Polscy naukowcy opublikowali większość swoich artykułów w polskich czasopismach ujętych w DOAJ, na drugim miejscu znalazły się natomiast czasopisma brytyjskie. Łącznie zostały zidentyfikowane 54 kraje wydawców, z czego do 20 krajów zostało przypisanych co najmniej 100 rekordów bibliograficznych.

Poniższy wykres pokazuje najczęściej występujące kraje według liczby rekordów bibliograficznych w Polskiej Bibliografii Naukowej, do których są przypisani wydawcy czasopism ujętych w DOAJ.



*Rysunek 5. Kraje wydawców czasopism ujętych w DOAJ występujących w Polskiej Bibliografii Naukowej.*

## ODSETEK PUBLIKACJI W OTWARTYM DOSTĘPIE WEDŁUG TYPÓW PUBLIKACJI

---

Najwyższy odsetek artykułów w otwartym dostępie jest obserwowany w dziedzinie nauk medycznych i nauk o zdrowiu (41,4%), a rozdziałów (7,67%) i książek (6,07%) w dziedzinie nauk teologicznych.

Poniższa tabela przedstawia publikacje opublikowane oraz opublikowane w otwartym dostępie w podziale na typ publikacji, dyscyplinę i dziedzinę. Odsetek jest liczony na poziomie każdego naukowca osobno a następnie agregowany na poziom dziedziny i dyscypliny.

Tabela 8. Odsetek otwartych publikacji w Polskiej Bibliografii Naukowej.

Dziedzina/Dyscyplina	Artykuły			Książki			Rozdziały			Wszystkie publikacje		
	N	w tym OA	% OA	N	w tym OA	% OA	N	w tym OA	% OA	N	w tym OA	% OA
<b>Dziedzina nauk humanistycznych</b>	48 383	10 545	22,1%	25 663	842	2,86%	182 088	16 902	3,41%	256 134	5 515	8,82%
Archeologia	2 648	555	24,4%	1 106	42	3,13%	8 864	960	3,33%	12 618	363	8,94%
Filozofia	5 161	1 329	21,4%	1 800	116	5,27%	10 159	2 063	5,83%	17 120	618	12,87%
Historia	9 156	1 845	18,0%	6 337	72	1,09%	38 413	2 593	1,49%	53 906	676	5,59%
Językoznawstwo	11 243	2 728	25,7%	5 605	229	3,98%	44 700	4 275	4,13%	61 548	1 318	10,98%
Literaturoznawstwo	12 608	2 748	23,5%	7 428	302	2,29%	56 633	4 855	3,21%	76 669	1 805	8,08%
Nauki o kulturze i religii	4 356	967	23,5%	1 860	57	3,10%	14 094	1 623	4,72%	20 310	599	10,34%
Nauki o sztuce	3 211	373	10,8%	1 527	24	1,25%	9 225	533	1,25%	13 963	136	3,82%
<b>Dziedzina nauk inżynierijno-technicznych</b>	183 706	40 885	23,6%	12 754	877	2,28%	259 299	59 802	5,08%	455 759	18 040	16,93%
Architektura i urbanistyka	3 320	878	25,9%	1 216	40	2,42%	15 390	1 445	4,57%	19 926	527	12,55%
Automatyka, elektronika i elektrotechnika	24 666	6 228	25,0%	1 091	42	1,26%	36 287	7 231	2,64%	62 044	961	15,47%
Informatyka techniczna i telekomunikacja	14 690	3 225	23,0%	1 686	231	3,27%	35 177	5 304	4,55%	51 553	1 848	13,90%
Inżynieria biomedyczna	3 494	868	27,3%	191	5	0,55%	6 033	1 065	6,70%	9 718	192	19,33%
Inżynieria chemiczna	9 966	1 826	18,6%	303	11	1,02%	5 949	2 182	3,24%	16 218	345	13,91%
Inżynieria lądowa i transport	29 046	4 804	18,4%	2 466	73	1,86%	35 761	6 632	4,22%	67 273	1 755	13,14%
Inżynieria materiałowa	25 396	3 472	15,9%	808	15	0,59%	40 168	4 452	1,66%	66 372	965	10,74%
Inżynieria mechaniczna	40 976	9 837	24,6%	2 400	113	1,85%	42 657	14 126	6,63%	86 033	4 176	19,66%
Inżynieria środowiska, górnictwo i energetyka	32 152	9 747	31,2%	2 593	347	4,90%	41 877	17 365	8,77%	76 622	7 271	25,48%

Dziedzina/Dyscyplina	Artykuły			Książki			Rozdziały			Wszystkie publikacje		
	N	w tym OA	% OA	N	w tym OA	% OA	N	w tym OA	% OA	N	w tym OA	% OA
<b>Dziedzina nauk medycznych i nauk o zdrowiu</b>	124 986	31 881	25,1%	42 687	2 202	3,89%	420 692	56 981	5,48%	588 365	22 898	13,45%
Nauki farmaceutyczne	30 714	9 610	32,5%	7 111	525	4,40%	56 817	15 440	7,29%	94 642	5 305	19,59%
Nauki medyczne	6 023	1 840	29,3%	1 464	121	5,47%	12 523	3 532	8,71%	20 010	1 571	19,01%
Nauki o kulturze fizycznej	3 701	945	22,9%	2 172	86	2,75%	17 259	2 084	4,58%	23 132	1 053	10,52%
Nauki o zdrowiu	2 763	557	23,1%	1 380	30	2,63%	12 754	1 023	4,01%	16 897	436	9,83%
<b>Dziedzina nauk rolniczych</b>	6 176	1 109	19,8%	3 544	174	4,81%	35 324	2 185	3,87%	45 044	902	9,13%
Nauki leśne	26 866	8 282	32,1%	5 840	430	5,08%	75 473	14 200	8,61%	108 179	5 488	18,14%
Rolnictwo i ogrodnictwo	21 518	3 285	17,1%	11 295	444	2,92%	132 158	9 785	3,52%	164 971	6 056	6,84%
Technologia żywności i żywienia	5 159	1 081	19,9%	2 151	181	6,47%	15 953	1 796	4,97%	23 263	534	11,46%
Weterynaria	11 597	2 605	20,8%	6 337	151	2,11%	50 226	4 015	3,13%	68 160	1 259	8,13%
Zootechnika i rybactwo	498	166	33,8%	173	-	0,00%	2 813	172	0,16%	3 484	6	6,84%
<b>Dziedzina nauk ścisłych i przyrodniczych</b>	9 971	2 401	25,2%	1 220	60	2,77%	9 392	2 749	3,95%	20 583	288	17,28%
Astronomia	3 971	722	17,1%	1 794	62	6,07%	12 564	1 728	7,67%	18 329	944	11,01%
Informatyka	3 971	722	17,1%	1 794	62	6,07%	12 564	1 728	7,67%	18 329	944	11,01%
Matematyka	1 538	210	13,7%	956	35	1,58%	4 693	379	1,76%	7 187	134	9,01%
Nauki biologiczne	64	2	3,8%	18	-	0,00%	113	2	0,00%	195	-	3,41%
Nauki chemiczne	502	58	9,3%	337	-	0,00%	2 067	130	1,53%	2 906	72	6,83%
Nauki fizyczne	972	150	17,0%	601	35	2,57%	2 513	247	2,04%	4 086	62	10,69%
Nauki o Ziemi i środowisku	832 769	241 277	27,1%	99 701	5 098	2,56%	1 084 098	306 315	4,27%	2 016 568	59 940	19,29%

Dziedzina/Dyscyplina	Artykuły			Książki			Rozdziały			Wszystkie publikacje		
	N	w tym OA	% OA	N	w tym OA	% OA	N	w tym OA	% OA	N	w tym OA	% OA
<b>Dziedzina nauk społecznych</b>	124 986	31 881	25,1%	42 687	2 202	3,89%	420 692	56 981	5,48%	588 365	22 898	13,45%
Ekonomia i finanse	30 714	9 610	32,5%	7 111	525	4,40%	56 817	15 440	7,29%	94 642	5 305	19,59%
Geografia społeczno- ekonomiczna i gospodarka przestrzenna	6 023	1 840	29,3%	1 464	121	5,47%	12 523	3 532	8,71%	20 010	1 571	19,01%
Nauki o bezpieczeństwie	3 701	945	22,9%	2 172	86	2,75%	17 259	2 084	4,58%	23 132	1 053	10,52%
Nauki o komunikacji społecznej i mediach	2 763	557	23,1%	1 380	30	2,63%	12 754	1 023	4,01%	16 897	436	9,83%
Nauki o polityce i administracji	6 176	1 109	19,8%	3 544	174	4,81%	35 324	2 185	3,87%	45 044	902	9,13%
Nauki o zarządzaniu i jakości	26 866	8 282	32,1%	5 840	430	5,08%	75 473	14 200	8,61%	108 179	5 488	18,14%
Nauki prawne	21 518	3 285	17,1%	11 295	444	2,92%	132 158	9 785	3,52%	164 971	6 056	6,84%
Nauki socjologiczne	5 159	1 081	19,9%	2 151	181	6,47%	15 953	1 796	4,97%	23 263	534	11,46%
Pedagogika	11 597	2 605	20,8%	6 337	151	2,11%	50 226	4 015	3,13%	68 160	1 259	8,13%
Prawo kanoniczne	498	166	33,8%	173	-	0,00%	2 813	172	0,16%	3 484	6	6,84%
Psychologia	9 971	2 401	25,2%	1 220	60	2,77%	9 392	2 749	3,95%	20 583	288	17,28%
<b>Dziedzina nauk teologicznych</b>	3 971	722	17,1%	1 794	62	6,07%	12 564	1 728	7,67%	18 329	944	11,01%
Nauki teologiczne	3 971	722	17,1%	1 794	62	6,07%	12 564	1 728	7,67%	18 329	944	11,01%
<b>Dziedzina sztuki</b>	1 538	210	13,7%	956	35	1,58%	4 693	379	1,76%	7 187	134	9,01%
Sztuki filmowe i teatralne	64	2	3,8%	18	-	0,00%	113	2	0,00%	195	-	3,41%
Sztuki muzyczne	502	58	9,3%	337	-	0,00%	2 067	130	1,53%	2 906	72	6,83%
Sztuki plastyczne i konserwacja dzieł sztuki	972	150	17,0%	601	35	2,57%	2 513	247	2,04%	4 086	62	10,69%
<b>Łącznie</b>	<b>832 769</b>	<b>241 277</b>	<b>27,1%</b>	<b>99 701</b>	<b>5 098</b>	<b>2,56%</b>	<b>1 084 098</b>	<b>306 315</b>	<b>4,27%</b>	<b>2 016 568</b>	<b>59 940</b>	<b>19,29%</b>



## ZGODNOŚĆ METADANYCH PBN I DOAJ

Polska Bibliografia Naukowa zawiera metadane o otwartym dostępie dotyczące publikacji naukowych reprezentowanych w bazie przez rekordy bibliograficzne. DOAJ zawiera metadane o otwartym dostępie czasopism naukowych. Dane w DOAJ pochodzą bezpośrednio od czasopism naukowych. W świecie idealnym, w przypadku poprawnego wprowadzenia danych do PBN metadane z obu źródeł, tj. PBN i DOAJ, powinny być w pełni zgodne. Celem niniejszej części raportu jest:

1. sprawdzenie jakości wprowadzanych do PBN danych przez instytucje naukowe, zakładając, że dane zapisane w DOAJ są poprawne,
2. oszacowanie, ile rekordów ma nieoprawnie wprowadzone dane o otwartym dostępie,
3. oszacowanie, ile rekordów powinno być oznaczonych jako otwarty dostęp, a nie jest.

Punktem wyjścia analizy było zmapowanie rodzajów licencji w danych z DOAJ oraz danych z PBN. Obie bazy danych wskazują ogólną licencję Creative Commons (CC) bez wskazywania wersji (np. licencja CC BY zawiera w sobie zarówno licencję CC BY 3.0 jak i CC BY 4.0).

W DOAJ oprócz wariantów licencji CC wskazywana jest również domena publiczna (Public Domain) oraz własna licencja wydawcy (Publisher's own licence). Obie wartości metadanych zostały w analizie uznane za zgodne z metadanymi z PBN, gdy w PBN-ie wskazana była inna licencja (Inna otwarta licencja). Szczegółowe mapowanie przedstawia poniższa tabela.

Informacje o licencji w PBN-ie są przypisane na poziomie pojedynczego rekordu bibliograficznego, podczas gdy w DOAJ dane o licencji przypisane są na poziomie całego czasopiśma.

*Tabela 9. Mapowanie typów licencji DOAJ na typy licencji PBN.*

Directory of Open Access Journals	Polska Bibliografia Naukowa
CC BY	CC BY
CC BY-NC	CC BY-NC
CC BY-NC-ND	CC BY-NC-ND
CC BY-NC-SA	CC BY-NC-SA
CC BY-ND	CC BY-ND
CC BY-SA	CC BY-SA
Public Domain	Inna otwarta licencja
Publisher's own license	Inna otwarta licencja

W analizie zgodności metadanych uwzględnione zostały rekordy bibliograficzne z Polskiej Bibliografii Naukowej z lat 2013–2018 opublikowane w czasopismach ujętych w Directory of Open Access Journals. Zostało przyjęte założenie, że czasopismo ujęte w DOAJ publikowało artykuły w otwartym dostępie w całym okresie objętym analizą.

Poniższa tabela pokazuje, że zgodność podanych danych dotyczy tylko ponad 35% rekordów.

*Tabela 10. Zgodność typów licencji w PBN i DOAJ.*

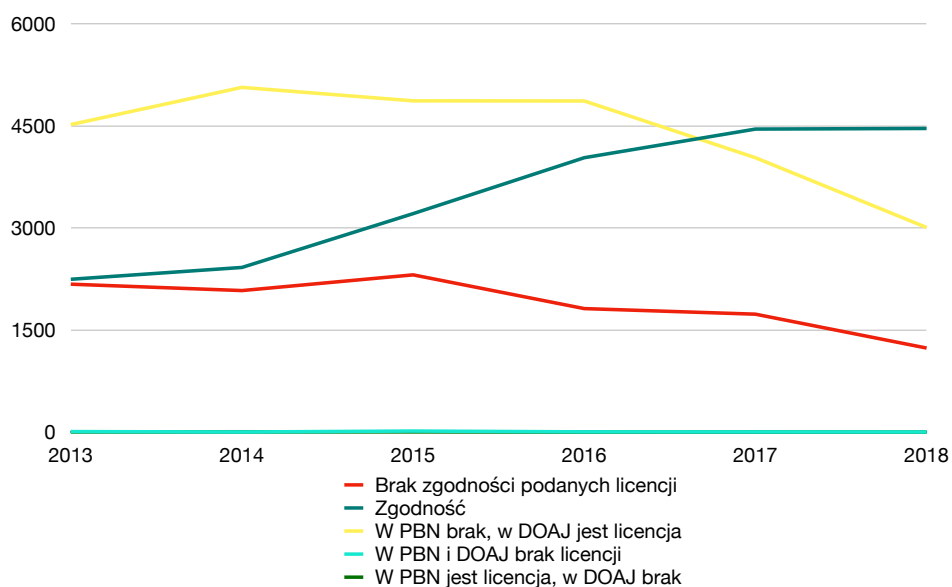
Zgodność	Liczba rekordów	Odsetek rekordów
W PBN brak, w DOAJ jest licencja	26369	45,00%
Zgodność	20833	35,55%
Brak zgodności podanych licencji	11354	19,37%
W PBN i DOAJ brak licencji	35	0,06%
W PBN jest licencja, w DOAJ brak	13	0,02%
<b>Łącznie</b>	<b>58604</b>	<b>100,00%</b>

Analiza przypadków, gdy w PBN nie wskazano licencji, a w DOAJ wskazano pokazuje, że dotyczy to najczęściej licencji CC BY-NC-ND.

*Tabela 11. Niezgodność typów licencji w PBN i DOAJ w podziale licencję raportowane w PBN.*

Rodzaj licencji DOAJ	Liczba rekordów w PBN
CC BY-NC-ND	10 037
CC BY	7 520
CC BY-NC	3 628
CC BY-NC-SA	2 187
CC BY-SA	1 203
CC BY-ND	1 059
Inna otwarta licencja	735

Poniższy rysunek pokazuje, że liczba zgodnych rekordów systematycznie wzrasta, a liczba niezgodnych maleje w kolejnych latach.



**Rysunek 6.** Liczba rekordów według zgodności metadanych w danym roku.

Analiza rekordów ze zgodnymi metadanyami pokazuje, że zgodność dotyczy najczęściej licencji CC BY oraz CC BY-NC-ND.

**Tabela 12.** Liczba metadanych zgodnych w podziale na rodzaj licencji.

Rodzaj licencji	Liczba rekordów w PBN
CC BY	7 417
CC BY-NC-ND	6 480
CC BY-NC-SA	3 751
CC BY-NC	1 638
CC BY-ND	776
Inna otwarta licencja	439
CC BY-SA	332

Natomiast niezgodność rekordów pokazuje, że taka sytuacja dotyczy najczęściej wskazania w PBN innej licencji, podczas gdy w DOAJ jest podana konkretna licencja CC. W wierszach przedstawiono dane zebrane w PBN, a w kolumnach dane zebrane w DOAJ.

**Tabela 13.** Liczba metadanych niezgodnych w podziale na rodzaj licencji.

Rodzaj licencji	CC BY	CC BY-NC	CC BY-NC-ND	CC BY-NC-SA	CC BY-ND	CC BY-SA	Inna otwarta licencja	Łącznie niezgodności
CC BY		415	352	60	38	52	81	998
CC BY-NC	121		190	772	4	4	49	1140
CC BY-NC-ND	156	535		89	14	37	6	837
CC BY-NC-SA	24	19	389		1	49	1	483
CC BY-ND	10	55	187	5		2	4	263
CC BY-SA	37	16	37	29	2		4	125
Inna otwarta licencja	1650	1095	3323	1003	187	250		7508
<b>Łącznie</b>	<b>1998</b>	<b>2135</b>	<b>4478</b>	<b>1958</b>	<b>246</b>	<b>394</b>	<b>145</b>	<b>11354</b>

Poniższa tabela przedstawia instytucje naukowe z najwyższym odsetkiem zgodnych metadanych o licencji w PBN i w DOAJ. Uwzględniono tylko instytucje naukowe, które raportowały do PBN co najmniej 100 rekordów bibliograficznych dotyczących artykułów opublikowanych w czasopismach ujętych w DOAJ.

**Tabela 14.** Instytucje naukowe najdokładniej raportujące metadane o licencjach.

Instytucja naukowa	Liczba rekordów	Odsetek zgodności
Instytut Fizyki Jądrowej im. Henryka Niewodniczańskiego Polskiej Akademii Nauk	919	91,99%
Instytut Medycyny Pracy im. prof. dr. med. Jerzego Nofera	199	88,44%
Wyższa Szkoła Gospodarki Euroregionalnej im. Alcide De Gasperi w Józefowie	123	86,62%
Politechnika Śląska	1525	79,76%
Gdański Uniwersytet Medyczny	502	74,26%
Uniwersytet Medyczny im. Piastów Śląskich we Wrocławiu	1205	71,85%
Warszawski Uniwersytet Medyczny	1201	71,74%
Śląski Uniwersytet Medyczny w Katowicach	1097	71,65%
Akademia Wychowania Fizycznego Józefa Piłsudskiego w Warszawie	206	70,79%
Akademia Morska w Szczecinie	141	66,51%

Instytucja naukowa	Liczba rekordów	Odsetek zgodności
Uniwersytet Przyrodniczy w Lublinie	294	64,05%
Instytut Medycyny Doświadczalnej i Klinicznej im. Mirosława Mossakowskiego Polskiej Akademii Nauk	128	63,05%
Uniwersytet Gdański	503	62,72%
Uniwersytet Medyczny w Białymstoku	443	61,87%
Akademia Wychowania Fizycznego i Sportu im. Jędrzeja Śniadeckiego w Gdańsku	141	61,84%
Instytut Kardiologii im. Prymasa Tysiąclecia Stefana Kardynała Wyszyńskiego	107	61,49%
Politechnika Białostocka	301	60,20%
Politechnika Wrocławska	464	58,07%
Uniwersytet Medyczny w Łodzi	797	56,52%
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie	209	54,29%
Politechnika Lubelska	343	52,61%
Uniwersytet Jagielloński w Krakowie	1067	52,28%
Politechnika Częstochowska	411	50,43%
Politechnika Krakowska im. Tadeusza Kościuszki	224	50,34%
Uniwersytet Przyrodniczy we Wrocławiu	246	47,58%

Poniższa tabela przedstawia rodzaje instytucji naukowych sprawozdające dane do Polskiej Bibliografii Naukowej według liczby rekordów bibliograficznych oraz zgodności metadanych. Szczegółowe zestawienia na poziomie instytucji naukowej zawiera Załącznik ZGODNOŚĆ.

**Tabela 15.** Zgodność metadanych o licencjach w PBN i DOAJ wg rodzaju instytucji naukowych.

Rodzaj instytucji naukowej	Brak zgodności podanych licencji		W PBN brak, w DOAJ jest licencja		W PBN i DOAJ brak licencji		W PBN jest licencja, w DOAJ brak		Zgodność		Łączna liczba rekordów	%
	Liczba rekordów	%	Liczba rekordów	%	Liczba rekordów	%	Liczba rekordów	%	Liczba rekordów	%		
Artystyczne	5	10,87%	37	80,43%		0,00%		0,00%	4	8,70%	46	100,00%
Ekonomiczne	165	10,42%	1275	80,54%	24	1,52%		0,00%	119	7,52%	1583	100,00%
Instytuty badawcze	933	20,35%	2461	53,68%		0,00%		0,00%	1191	25,98%	4585	100,00%
Instytuty PAN	565	19,11%	860	29,09%		0,00%		0,00%	1531	51,79%	2956	100,00%
Medyczne	2582	24,21%	1993	18,69%		0,00%		0,00%	6089	57,10%	10664	100,00%
Morskie	37	10,45%	107	30,23%		0,00%		0,00%	210	59,32%	354	100,00%
Pedagogiczne	136	17,22%	431	54,56%		0,00%		0,00%	223	28,23%	790	100,00%
Przyrodnicze	490	15,88%	1768	57,29%	2	0,06%		0,00%	826	26,77%	3086	100,00%
Techniczne	1653	14,90%	5232	47,16%	2	0,02%		0,00%	4208	37,93%	11095	100,00%
Teologiczne		0,00%	11	78,57%		0,00%		0,00%	3	21,43%	14	100,00%
Uniwersytety	3555	20,35%	9593	54,90%	5	0,03%	3	0,02%	4317	24,71%	17473	100,00%
Wojskowe	109	17,52%	355	57,07%		0,00%		0,00%	158	25,40%	622	100,00%
Wychowania fizycznego	289	23,36%	535	43,25%		0,00%		0,00%	413	33,39%	1237	100,00%
Zawodowe		0,00%	4	100,00%		0,00%		0,00%		0,00%	4	100,00%
Inne	835	20,46%	1693	41,48%	2	0,05%	10	0,25%	1541	37,76%	4081	100,00%
<b>Łącznie</b>	<b>11354</b>	<b>19,38%</b>	<b>26355</b>	<b>44,98%</b>	<b>35</b>	<b>0,06%</b>	<b>13</b>	<b>0,02%</b>	<b>20833</b>	<b>35,56%</b>	<b>58590</b>	<b>100,00%</b>

## ZGODNOŚĆ METADANYCH W OTWARTYM MIEJSCU PUBLIKACJI

W przypadku wszystkich artykułów opublikowanych w czasopismach ujętych w DOAJ pole o otwartym miejscu publikacji powinno wskazywać otwarte czasopismo. Analiza wykazała, że zgodność występuje w przypadku tylko 43% rekordów.

*Tabela 16. Zgodność metadanych PBN i DOAJ o otwartym miejscu publikacji.*

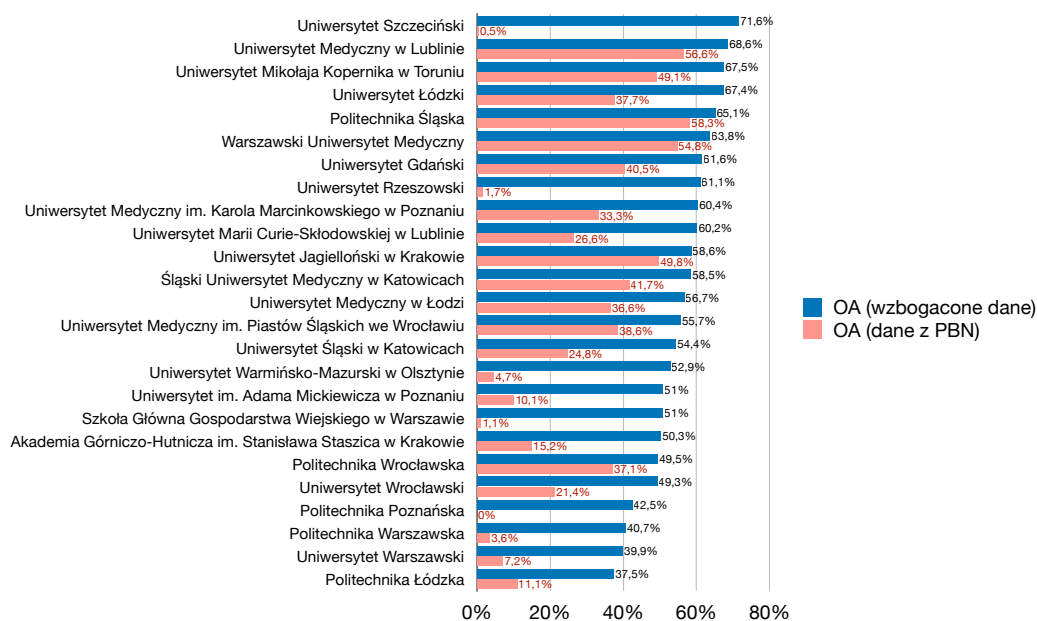
Liczba rekordów		
Zgodność, tj. wskazano „otwarte czasopismo”	25 220	43,0%
Brak metadanych o otwartym czasopiśmie, a artykuł opublikowany w czasopiśmie z DOAJ	27 947	47,7%
Wskazany „Inne” miejsce publikacji	2 789	4,8%
Wskazane „Otwarte repozytorium”	2 648	4,5%
<b>Łącznie</b>	<b>58 604</b>	<b>100%</b>

## POZIOM OTWARTOŚCI INSTYTUCJI NAUKOWYCH

W celu zbadania poziomu otwartości polskich instytucji naukowych oraz dyscyplin naukowych rozszerzone zostały metadane z PBN o metadane zawarte w DOAJ oraz w bazie ARIANTA. Baza ARIANTA zawiera informacje bibliograficzne o wszystkich polskich czasopiśmie naukowych, w tym informacje o otwartym dostępie.

W ten sposób uzupełnione zostały nie tylko informacje o wszystkich polskich czasopiśmie naukowych, co jest niezwykle istotne w przypadku badania praktyk publikacyjnych w naukach humanistycznych i społecznych, ale również uwzględnione zostały zagraniczne czasopisma otwarte. Należy podkreślić, że nawet tak wzbogacony zbiór nie zawiera kompletu informacji, tj. nie wszystkie funkcjonujące na świecie czasopisma naukowe są ujęte w DOAJ. Co więcej, praktyki niektórych dyscyplin naukowych, jak np. fizyki czy matematyki oparte na preprintach i otwartych repozytoriach, sprawiają, że analiza otwartości praktyk publikacyjnych w oparciu o dane dotyczące czasopiśmie naukowych nie jest wyczerpująca.

Dzięki procedurze wzbogacania danych z PBN można porównać dla każdej instytucji i dyscypliny otwartość na poziomie rekordów bibliograficznych dotyczących artykułów z czasopiśmie naukowych ze zbiorem rekordów wzbogaconym dzięki wykorzystaniu dodatkowych źródeł. Poniższy rysunek zestawia instytucje naukowe z największym odsetkiem rekordów bibliograficznych odnoszących się do artykułów naukowych z lat 2013–2018 w otwartym dostępie oraz pokazuje różnice w otwartości tych artykułów według dwóch zbiorów danych (PBN oraz wzbogacone dane). Jak widać, dla niektórych instytucji (jak np. Uniwersytetu Warmińsko-Mazurskiego w Olsztynie) deklarowany w PBN-ie poziom otwartości jest ponad dziesięciokrotnie niższy od tego uzyskanego dzięki wzbogaceniu metadanych z dodatkowych źródeł.



Rysunek 7. Instytucje naukowe z największym odsetkiem rekordów w otwartym dostępie w Polskiej Bibliografii Naukowej.



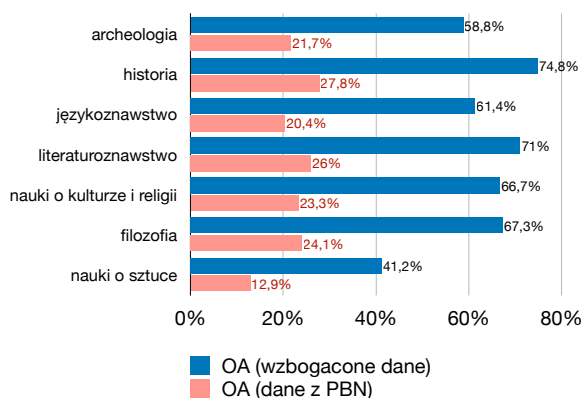
Poniższa tabela wskazuje liczbę rekordów bibliograficznych przypisanych do artykułów z lat 2013–2018 wraz z porównaniem odsetek otwartości według PBN-u i wzbogaconego zbioru danych. Szczegółowe informacje dla każdej instytucji zawiera Załącznik OA.

*Tabela 17. Rekordy dotyczące artykułów naukowych według rodzaju instytucji naukowej.*

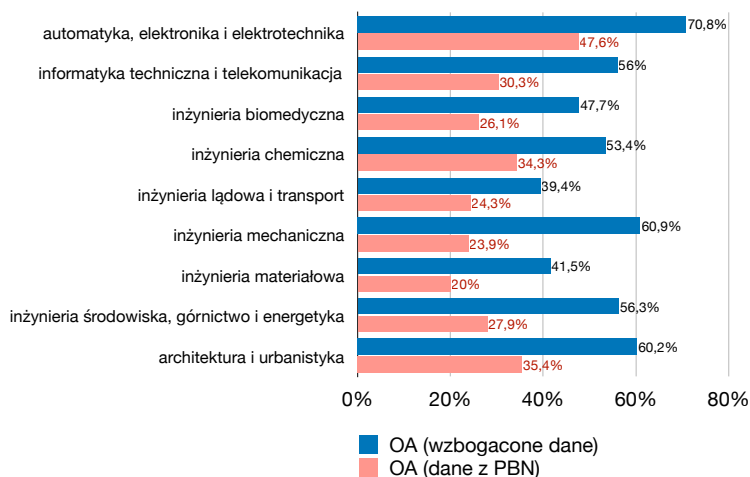
Rodzaj instytucji naukowej	Liczba publikacji	Liczba publikacji w OA (wzbogacone dane)	Odsetek OA (wzbogacone dane)	Liczba publikacji w OA (PBN)	Odsetek publikacji w OA (PBN)
Artystyczne	1209	544	45,0%	104	8,6%
Ekonomiczne	15409	11699	75,9%	2955	19,2%
Instytuty badawcze	41125	23689	57,6%	13907	33,8%
Instytuty PAN	35284	15386	43,6%	11042	31,3%
Medyczne	51861	31942	61,6%	22377	43,1%
Morskie	1979	1540	77,8%	943	47,7%
Pedagogiczne	10779	7064	65,5%	3207	29,8%
Przyrodnicze	28714	16041	55,9%	6895	24,0%
Techniczne	101934	54660	53,6%	30003	29,4%
Teologiczne	204	164	80,4%	39	19,1%
Uniwersytety	173446	98830	57,0%	42797	24,7%
Wojskowe	6607	3823	57,9%	1463	22,1%
Wychowania fizycznego	5375	4047	75,3%	2435	45,3%
Zawodowe	14	13	92,9%	0	0,0%
Inne	35952	22852	63,6%	10826	30,1%
<b>Łącznie</b>	<b>509892</b>	<b>292294</b>	<b>57,3%</b>	<b>148993</b>	<b>29,2%</b>

## POZIOM OTWARTOŚCI DYSCYPLIN NAUKOWYCH

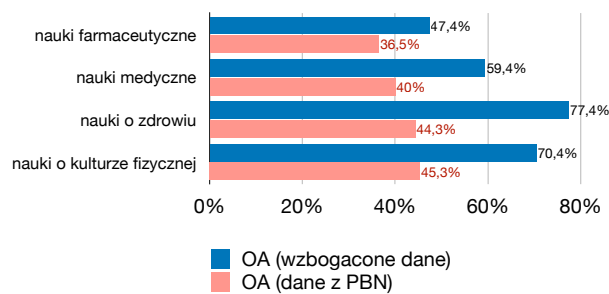
Analogicznie do analizy otwartości instytucji naukowych, przeprowadzona została analiza otwartości dyscyplin naukowych opartą o porównanie danych z PBN z danymi ze wzbogaconego zbioru danych. Poniżej przedstawiamy zestawienia odsetków artykułów w poszczególnych dziedzinach i dyscyplinach naukowych.



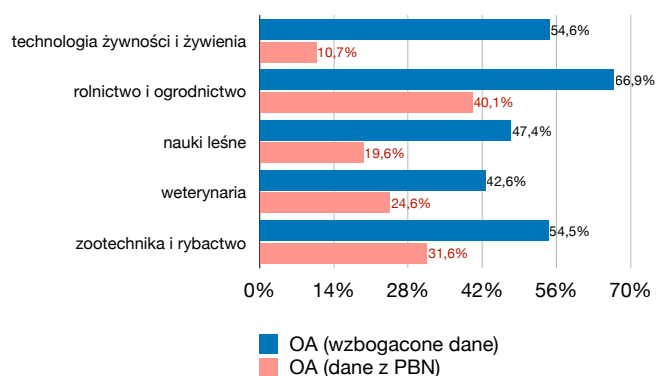
**Rysunek 8.** Rekordy dotyczące artykułów naukowych z dziedziny nauk humanistycznych w Polskiej Bibliografii Naukowej.



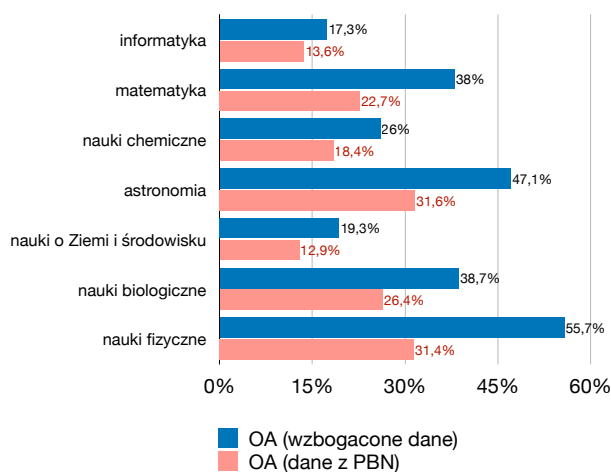
**Rysunek 9.** Rekordy dotyczące artykułów naukowych z dziedziny nauk inżynierjno-technicznych w Polskiej Bibliografii Naukowej.



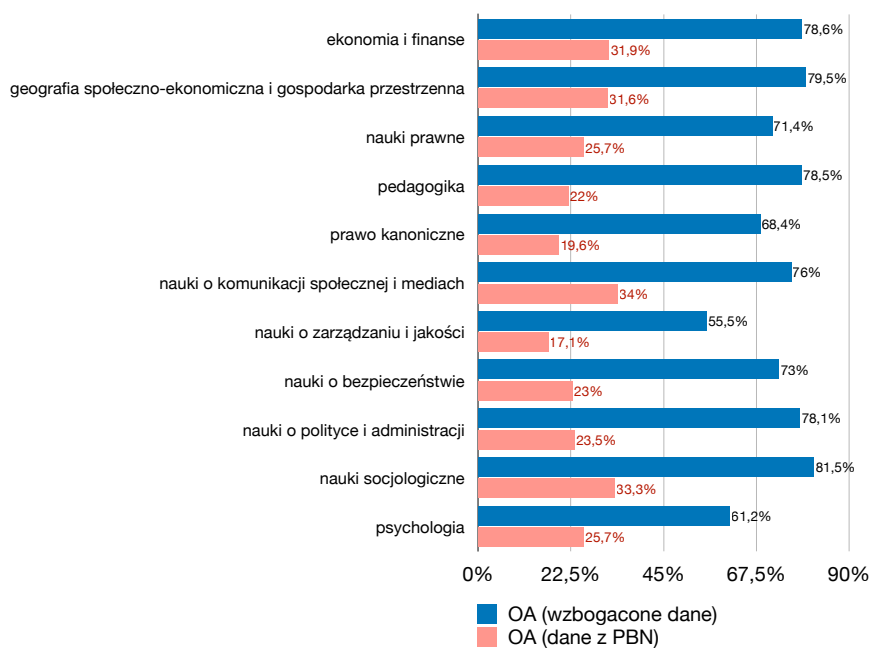
**Rysunek 10.** Rekordy dotyczące artykułów naukowych z dziedziny nauk medycznych i o zdrowiu w Polskiej Bibliografii Naukowej.



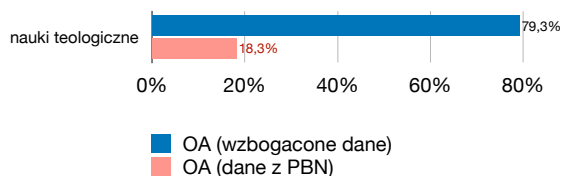
**Rysunek 11.** Rekordy dotyczące artykułów naukowych z dziedziny nauk rolniczych w Polskiej Bibliografii Naukowej.



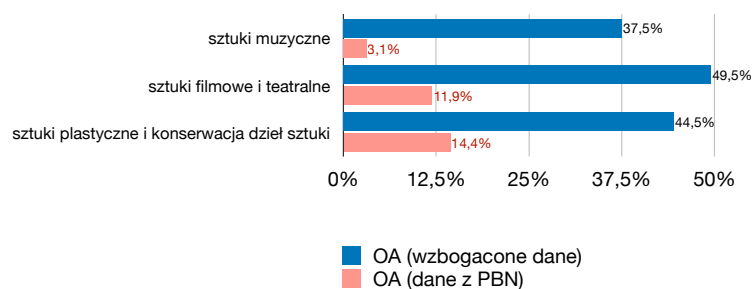
**Rysunek 12.** Rekordy dotyczące artykułów naukowych z dziedziny nauk ścisłych i przyrodniczych w Polskiej Bibliografii Naukowej.



**Rysunek 13.** Rekordy dotyczące artykułów naukowych z dziedziny nauk społecznych w Polskiej Bibliografii Naukowej.



**Rysunek 14.** Rekordy dotyczące artykułów naukowych z dziedziny nauk teologicznych w Polskiej Bibliografii Naukowej



**Rysunek 15.** Rekordy dotyczące artykułów naukowych z dziedziny sztuki w Polskiej Bibliografii Naukowej.

We wszystkich dziedzinach poziom otwartości poszczególnych dyscyplin istotnie wzrasta po uwzględnieniu dodatkowych źródeł danych. Dla przykład otwartość historii rośnie z 27,8% do 74,8%, a nauk teologicznych z 18,3% do 79,3%. Wzbogacenie pokazuje w lepszy sposób, jaka jest rzeczywista skala praktyk publikacyjnych w zakresie otwartej nauki.