

Wpływ otwartego dostępu na cytowalność publikacji w polskich repozytoriach instytucjonalnych

Jan Kowalski¹, Anna Nowak²

¹ Uniwersytet Warszawski, Wydział Dziennikarstwa, Informacji i Bibliologii | ORCID: 0000-0002-1825-0097

² Politechnika Wrocławska, Biblioteka Główna | ORCID: 0000-0001-5109-3700

Streszczenie

Cel. Artykuł analizuje zależność między udostępnieniem publikacji naukowej w otwartym repozytorium instytucjonalnym a liczbą cytowań uzyskanych w okresie trzech lat od daty publikacji. Metoda. Badaniem objęto 4 812 artykułów z lat 2019–2022, zdeponowanych w dwunastu polskich repozytoriach opartych na oprogramowaniu DSpace. Dane cytowań pozyskano z baz Scopus i Web of Science, a metadane repozytoryjne poprzez interfejs OAI-PMH oraz REST API. Zastosowano model regresji ujemnej dwumianowej z kontrolą dziedziny, punktacji MNiSW i typu dostępu. Wyniki. Publikacje dostępne w modelu otwartym uzyskały średnio o 31% więcej cytowań ($IRR = 1,31$; $p < 0,001$). Efekt był najsilniejszy w naukach społecznych i humanistycznych, a najslabszy w naukach medycznych. Kompletność metadanych okazała się istotnym predyktorem widoczności. Wnioski. Deponowanie prac w repozytorium instytucjonalnym przynosi mierzalną korzyść cytowości, jednak jej skala zależy od jakości opisu bibliograficznego i konsekwentnego stosowania identyfikatorów trwałych.

Słowa kluczowe: otwarty dostęp · bibliometria · repozytoria instytucjonalne · DSpace · cytowalność · metadane

1. Wprowadzenie

Postulat otwartego dostępu do wyników badań finansowanych ze środków publicznych towarzyszy polityce naukowej od ponad dwóch dekad. W Polsce jego instytucjonalnym wyrazem stały się repozytoria uczelniane, budowane w większości na oprogramowaniu DSpace, oraz obowiązek raportowania dorobku do systemu POL-on. Mimo rosnącej liczby zdeponowanych prac, wciąż brakuje systematycznych badań nad tym, czy i w jakim stopniu obecność w repozytorium przekłada się na mierzalną widoczność publikacji.

Dotychczasowe ustalenia międzynarodowe nie są jednoznaczne. Część autorów wskazuje na wyraźną przewagę cytowości prac otwartych, inni tłumaczą ją efektem autoselekcji — badacze chętniej udostępniają teksty, które sami uznają za wartościowe. Niniejsze opracowanie stara się rozdzielić te dwa mechanizmy na materiale polskim.

2. Materiał i metoda

Próbie zbudowano na podstawie zbiorów dwunastu repozytoriów udostępniających punkt końcowy OAI-PMH. Metadane pobrano w schemacie Dublin Core, a następnie uzupełniono o dane z REST API DSpace w wersji 7 i nowszych, co pozwoliło uzyskać informacje o pakietach plików i faktycznej dostępności pełnych tekstów. Rekordy dopasowano do rekordów Scopus poprzez DOI, a w razie jego braku — metodą porównania tytułu, roku i nazwisk autorów.

Charakterystyka próby	Liczba	Udział
Artykuły ogółem	4 812	100,0%
Pełny tekst otwarty	2 967	61,7%
Wyłącznie metadane	1 845	38,3%
Z identyfikatorem DOI	4 401	91,5%
Z identyfikatorem ORCID autora	2 118	44,0%

Tab. 1. Struktura badanej próby (opracowanie własne)

3. Wyniki

Model regresji ujemnej dwumianowej wykazał istotny statystycznie efekt otwartego dostępu na liczbę cytowań. Po uwzględnieniu dziedziny, roku publikacji i punktacji ministerialnej współczynnik IRR wyniósł 1,31 przy przedziale ufności 95% [1,22; 1,41]. Oznacza to, że publikacja dostępna w pełnym tekście uzyskuje przeciętnie o niemal jedną trzecią cytowań więcej niż porównywalna praca opisana wyłącznie metadanymi.

Analiza w podziale na dziedziny ujawniła znaczne zróżnicowanie. W naukach społecznych efekt sięgał 47%, w humanistycznych 39%, natomiast w naukach medycznych — gdzie dominują inne kanały dystrybucji preprintów — jedynie 12% i na granicy istotności.

4. Dyskusja i wnioski

Uzyskane rezultaty wpisują się w nurt badań potwierdzających korzyść cytowalnościową otwartego dostępu, jednocześnie wskazując na jej warunkowy charakter. Kluczową rolę odgrywa jakość opisu bibliograficznego: rekordy pozbawione DOI lub identyfikatorów autorskich są znacznie trudniejsze do odnalezienia przez agregatory, co osłabia potencjalny efekt widoczności niezależnie od dostępności pliku.

Praktyczna rekomendacja dla instytucji prowadzących repozytoria brzmi zatem: automatyzacja deponowania powinna obejmować nie tylko przesłanie pliku, lecz także kontrolę kompletności metadanych na etapie ich tworzenia. Integracja repozytorium z uczelnianym systemem obiegu dokumentów pozwala osiągnąć ten cel przy minimalnym obciążeniu autorów.