

Znaczenie identyfikatorów obiektów cyfrowych dla czasopism medycznych

Szerokie upowszechnianie dorobku naukowego na świecie możliwe jest poprzez zamieszczanie artykułów w czasopismach elektronicznych. Błyskawiczne docieranie do pełnych tekstów dokumentów, nieograniczony dostęp do cyfrowych archiwów i repozytoriów wiedzy medycznej, prowadzenie szczegółowych poszukiwań oraz selekcja informacji, wreszcie automatyczne zliczanie cytowań prac to niewątpliwie atuty elektronicznych baz oraz zamieszczonych w nich dokumentów cyfrowych. Niestety, obok nich coraz częściej pojawiają się także zagrożenia i trudności, z którymi skomputeryzowana nauka musi sobie poradzić. Należą do nich: problemy dublowania prac naukowych zamieszczanych w wielu różnych źródłach, idące za nimi przekłamania liczby cytowań publikacji, trudności w poprawnym zapisywaniu danych bibliograficznych czy aktywne linkowania do wiarygodnej i ostatecznej wersji dokumentu.

Jedną z wizji uporządkowania i usystematyzowania piśmiennictwa elektronicznego jest opatrywanie prac niepowtarzalnymi identyfikatorami. W ten sposób jeden numer przypisany do pracy zapewnić ma trwałe łącze do metadanych lub pełnego tekstu dokumentu, zamieszczonego najczęściej w globalnym integratorze danych dziedzinowych. W publikacjach z nauk medycznych spotkać można trzy najpopularniejsze z nich: PMID, PMCID oraz NIHMS ID. PMID to identyfikator wszystkich obiektów umieszczonych w bazie PubMed. Ponieważ pozyskanie numeru nie wymaga otwartej formy udostępniania tekstu dokumentu (*open access*), numer ten, nadany automatycznie, uwidoczniiony jest najczęściej przy metadanych artykułu, obok autora, tytułu i danych wydawcy. Według ostatnich danych identyfikator ten posiada również blisko 60 polskich czasopism medycznych indeksowanych przez PubMed. Kolejny identyfikator PMCID to numer publikacji zdeponowanej i zatwierdzonej w bazie PubMed Central, będącej owocem badań finansowanych przez państwo. Numer ten, przydzielany recenzowanym artykułom *open access* z nauk biomedycznych i przyrodniczych, jest obligatoryjny przy ubieganiu się o dotacje z amerykańskiego Narodowego Instytutu Zdrowia (NIH) [1]. Ponieważ jego nadanie poprzedza sprawdzenie kryteriów *NIH Public Access Policy*, do czasu otrzymania PMCID

deponent może posługiwać się tymczasowym numerem referencyjnym NIHMS ID, ważnym 90 dni od daty publikacji pracy w bazie PubMed Central.

Coraz bardziej popularnym i niezależnym od indeksacji dokumentu w referencyjnych bazach danych cyfrowym identyfikatorem obiektów elektronicznych jest numer DOI (*digital object identifier*). Tym unikalnym i niezbywalnym identyfikatorem można opatrywać wszystkie obiekty elektroniczne (artykuły, książki, rozdziały z książek, normy, patenty, tabele, recenzje, raporty, mapy, pliki, dane liczbowe, tabele) [2] dostępne w sieci, niezależnie od formy udostępniania treści. Do połączenia fizycznej lokalizacji obiektu cyfrowego i jego metadanych — wysyłanych w momencie aktywacji do dystrybutora numeru — DOI wykorzystuje opracowany przez CNRI *Handle System*, z przedrostkiem „10”. Proces przydzielenia numeru składa się z kilku etapów, do których zaliczamy: 1. Wybór dystrybutora DOI

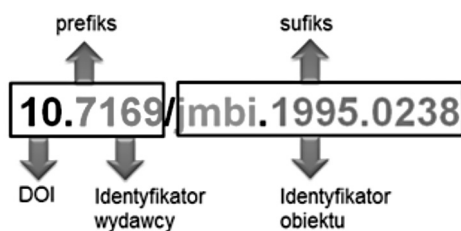
Decyzja o nadawaniu numeru DOI wymaga od wydawcy czasopisma rozpoznania oferty kilku dostępnych na rynku dystrybutorów numerów DOI (*DOI Registration Agency*), nad którymi kontrolę pełni International DOI Foundation. Najpopularniejszymi firmami rejestrującymi numery dla wydawnictw akademickich są: CrossRef, mEDRA czy R.R. Bowker. Nawiązując współpracę:

- instytucja ubiegająca się o członkostwo musi opublikować treści naukowe online, w pełnym tekście. Nawet w przypadku czasopisma zamkniętego dystrybutor ma prawo kontroli pełnego tekstu publikacji elektronicznej.
- instytucja musi przygotować i przesłać standaryzowane metadane obiektu elektronicznego. Korzystać przy tym może zarówno z własnych rozwiązań informatycznych lub, jeśli takich nie posiada, z aplikacji dostępnych na stronach dystrybutorów.
- w instytucji posiadającej wiele jednostek — umowę członkowską podpisuje centrala. Umowa ta jest standaryzowana dla wszystkich członków społeczności, co ułatwia dopełnianie formalności.
- obowiązkowym elementem współpracy jest uiszczanie składki członkowskiej, obliczanej według przydzielonej puli numerów DOI. Składkę tę uiszcza się co roku. Do-

datkową opłatę stanowi koszt nadania pojedynczego numeru.

2. Utworzenie struktury numeru

Tworzenie struktury numeru polega na uzupełnieniu prefiksu oraz stworzeniu sufiksu tak, aby w sposób logiczny i jednoznaczny identyfikował obiekt elektroniczny w sieci. Charakterystyczny dla numerów DOI jest początek przedrostka, rozpoczynający się od liczby 10, po którym pojawia się identyfikator wydawcy (*Publisher ID*) [3]. Numer ten otrzymuje wydawca w momencie podpisania deklaracji i wniesieniu rocznej opłaty członkowskiej. Po ukośniku tworzy się zapis przyrostka. Musi być on unikalny, zwięzły, logiczny, a równocześnie pozbawiony znaków diakrytycznych czy wielkich liter. Jedną z propozycji jest stworzenie węzłów, składających się ze skrótu czasopisma, numeru woluminu czy zeszytu oraz kolejnego numeru artykułu. Inną propozycją, powszechną szczególnie przy czasopismach posadowionych na odpowiednich platformach czasopism elektronicznych, jest powiązanie skrótu czasopisma z numerem systemowym artykułu, co eliminuje konieczność szerokiego rozbudowywania zapisu.



3. Aktywacja numeru DOI

Aby numer DOI identyfikował obiekt elektroniczny, konieczna jest jego aktywacja. Dokonuje się tego, wysyłając dystrybutorowi plik XML, zawierający jak największą liczbę metadanych. Będą one zindeksowane zarówno w bazie dystrybutora, a także, dzięki nawiązanym przez niego umowom partnerskim z wydawcami baz czy firmami tworzącymi multiwyszukiwarki, w integracjach i systemach informatycznych. Plik XML musi zawierać zarówno informacje o czasopiśmie (system sprawdzi prawa własności czasopisma po Publisher ID), dane osobowe autora, tytuł i słowa kluczowe pracy, utworzony numer DOI, jak i — obowiązkowo — aktualne łącze do pełnego tekstu. Może ono kierować do samego pliku dokumentu bądź do strony z informacjami bibliograficznymi, z której można dostać się do pliku. Aktywacja numeru trwa od kilku do kilkunastu minut, po których wydawca otrzymuje stosowne potwierdzenie na skrzynkę mailową.

Ponieważ numer DOI jest stały i niezbywalny, wszelkie zmiany lokalizacji pliku czy danych dokumentu wymagają dokonania korekty w wyeksportowanych metadanych. W tym celu należy ponownie wysłać plik XML, aktualizując informacje w bazie dystrybutora. W konsekwencji ten sam numer DOI — po zmianie miejsca przechowywania pliku — linkował będzie do nowego miejsca na serwerze.

4. Zapis numeru

Zaleca się, aby identyfikatory były dołączone do źródeł zarówno drukowanych, jak i elektronicznych. Zapis DOI zazwyczaj wprowadza się na pierwszej stronie artykułu w czasopiśmie elektronicznym oraz w podstronie z metadanymi.

Korzyścią z nadawania numerów DOI jest ich trwałość, zapewniająca stałe łącze do cyfrowych artykułów pomimo zmian w obrębie metadanych lub lokalizacji plików [4]. Równocześnie, dzięki gromadzeniu przez dystrybutorów standaryzowanych metadanych, ułatwione jest współdziałanie z innymi źródłami danych. Wreszcie możliwa jest także konfiguracja dodatkowych funkcji i usług, np. wzajemne linkowanie do artykułów z bibliografii załącznikowej, zarządzanie danymi niezależnie od formatów i rodzaju platform. Jedną z podstawowych funkcjonalności *CrossRef* jest np. *Reference Linking*, polegające na umożliwieniu linkowania z bibliografii załącznikowej (*listy referencji*) bezpośrednio do dokumentów, które są cytowane przez autora pracy, oraz *Cited-by Linking* — bezpłatna i opcjonalna usługa umożliwiająca linkowanie do dokumentów, które cytują dany artykuł.

Identyfikatory DOI stanowią także dużą wartość dla wydawnictwa i naukowca. Dzięki indeksacji numerów i pochodnych od nich metadanych przez bazy i multiwyszukiwarki (Google) zwiększa się widoczność czasopisma oraz cytowalność artykułów naukowych. Wpływa to w znaczący sposób na prestiż wydawcy, punktację periodyku oraz — wielokrotnie — na pozytywną decyzję dotyczącą zindeksowania zasobu przez międzynarodowe, dziedzinowe bazy danych.

Mnogość publikacji elektronicznych dostępnych w internecie może doprowadzić użytkowników do chaosu informacyjnego. Użytkownicy globalnej sieci z coraz większą trudnością potrafią rozpoznać naukowy charakter artykułów oraz ocenić ich przydatność, stąd identyfikatory obiektów cyfrowych, których nadawanie poprzedza dokładna weryfikacja techniczna i merytoryczna, są jedną z gwarancji tej wiarygodności. Równocześnie ich zapis jest spójny i logiczny, w odróżnieniu od rozbudowanych i po krótkim czasie nieaktywnych linków html, prowadzących do fizycznej lokalizacji dokumentów.

Mgr Piotr Karwasiński

Biblioteka Uniwersytecka

Uniwersytet im. Adama Mickiewicza w Poznaniu

e-mail: karwas@amu.edu.pl

Piśmiennictwo

1. Digital Identifiers of Scientific Literature (PMID, PMCID, NIHMS, DOI) and How to Use Them. <http://www.biosciencewriters.com/Digital-identifiers-of-scientific-literature-PMID-PMCID-NIHMS-DOI-and-how-to-use-them.aspx>.
2. Harris Pat, and Ted Koppel. Technology and Standards: Chaos: the Sici Emerges, Cicada-Like, After Eight Years of Dormancy — Sometimes Standards Take Time to Come into Their Own. As Ted Explains the Sici Is Enjoying a Renaissance. Against the Grain. 16.6 (2004): 88. Print.
3. Publication Manual of the American Psychological Association. Washington, DC: American Psychological Association, 2010. Print.
4. Berinstein P. DOI: A New Identifier For Digital Content. Searcher 6.1 (1998): 72. Business Source Complete. Web. 30 Sept. 2014.