

Agnieszka Brachfogel
Biblioteka Narodowa

OPEN ACCESS METADATA – nowatorski projekt NISO

Streszczenie: W artykule omówiona została inicjatywa NISO (National Information Standards Organization) opracowania standardu metadanych i znaczników dla zasobów udostępnianych w ramach wolnego dostępu, w szczególności prac naukowych. Wydawcy różnie oznaczają wolny dostęp do publikacji, co powoduje utrudnienie w korzystaniu z nich przez różnego rodzaju użytkowników: czytelników, autorów, dostawców treści itp. Dwa znaczniki: "wolny dostęp" i "licencja" powinny zdaniem twórców projektu, rozwiązać ten problem. W projekcie przedstawiono również odpowiednie zalecenia dla wydawców oraz zakres prac związanych z wprowadzeniem tych zapisów w praktyce.

Słowa kluczowe: międzynarodowe standardy, metadane, Open Access, NISO, wolny dostęp,

Idea otwartego dostępu (Open Access, OA), która znacząco wpłynęła na rozwój komunikacji i wymiany wiedzy naukowej, doczekała się kolejnego etapu standaryzacji. W ramach The National Information Standards Organization (NISO), niespełna rok temu powstała grupa robocza ds. metadanych dla wolnego dostępu (Working Group for Open Access Metadata) w celu ustalenia mechanizmu oznaczania statusu publikacji udostępnianych w ramach modelu open access.

Członkami zespołu zostały osoby reprezentujące ośrodki naukowe (University of Birmingham, American Chemical Society — ACS, wydawców (Elsevier, EDItEUR, International Association of STM Publishers), twórców sieci, baz danych i kanałów dystrybucji (Scholarly Publishing & Academic Resources Coalition — SPARC, Social Science Research Network — SSRN, Ex Libris Inc, CrossRef, JISC) oraz przedstawicieli środowiska związanego z upowszechnianiem praw autorskich i licencji wolnego dostępu (Creative Commons, Copyright Clearance Center — CCC).

Powołanie zespołu związane było z obserwowanym, znaczącym wzrostem publikacji udostępnianych w modelu OA i różnorodnością zapisu informacji o ich statusie. W praktyce spotkać bowiem można różnorodne określenia, m.in. *open access*, *increased access*, *public access* i wiele innych, oznaczających to samo: że dokument jest swobodnie dostępny. Co gorsza, różnice w terminologii pojawiają się nie tylko pomiędzy wydawcami; bywa, że ten sam wydawca stosuje różne określenia w swoich publikacjach, a także w ramach jednego czasopisma. Dodatkowo zamieszanie powodują sytuacje, w których autor zapłaciwszy wydawcy honorarium, udostępnia na łamach jego czasopisma swój artykuł w wolnym dostępie, przy czym do pozostałych artykułów dostęp jest możliwy tylko dla subskrybentów (jest to tzw. opcja mieszana — *hybrid options*).

Podobnie jest przy oznaczaniu warunków korzystania z publikacji w otwartym dostępie. Wydawcy opierają się na różnorodnych licencjach: niektórzy preferują Creative Commons (CC), a zwłaszcza Creative Commons Attribution (CC BY), inni stosują prawa własności lub ich warianty. W efekcie, współpraca pomiędzy osobami zainteresowanymi sprawą (interesariuszami) jest utrudniona, a ośrodki finansujące prace badawcze nie wiedzą, na ile dana publikacja jest zgodna z ich procedurami i strategiami. Problem ten dotyka również autorów, mających trudność w ustaleniu, które z praw nadal im przysługują i czy są one zgodne z ogólną polityką instytucji finansującej ich działalność naukową. Z kolei serwisy gromadzące dane (repozytoria, wyszukiwarki, bazy danych) nie mają jednego mechanizmu identyfikującego artykuły, które można legalnie pobierać i udostępniać jako otwarte.

Rozpoznawszy te problemy grupa robocza NISO rozpoczęła swoje prace od przygotowania ich zarysu i harmonogramu, który zaakceptowano w styczniu 2013 r. (patrz: Specification for Open Access Metadata and Indicators)¹. Poza analizą odmiennych praktyk w oznaczaniu publikacji w wolnym dostępie, zwrócono uwagę na inne projekty, które funkcjonują w obszarze OA (CrossMark, HowOpenIsIt?, Vocabularies for OA (V40A), ONIX-PL, Linked Content Coalition, Open Discovery Initiative), uznając jednocześnie za konieczne nawiązanie ścisłej współpracy z realizatorami tych projektów. Określono również odbiorców projektu, wymieniając sześć środowisk, które mogą skorzystać z zastosowania standaryzowanych metadanych OA. Są to:

- czytelnicy, którzy są zobowiązani przestrzegać praw określających zakres wykorzystania danej publikacji, byliby w prosty sposób poinformowani, czy jest ona w wolnym dostępie, czy są jakieś ograniczenia w wykorzystaniu treści itp.,
- autorzy chcący ustalić, jakie prawa zachowują do swojej pracy i na ile jest to zgodne z polityką instytucji finansującej ich działalność naukową i badawczą,
- wydawcy, którym zależy na jednoznacznym informowaniu swoich odbiorców o tym, co mogą, a czego nie mogą zrobić z udostępnioną publikacją,
- instytucje finansujące prace badawcze dążące do możliwie najszerzej promocji otwartych zasobów przez nie sponsorowanych oraz chcących lepiej kontrolować, na ile wykorzystanie publikacji jest zgodne z ich polityką,
- serwisy wyszukiwarek, baz danych abstraktowych i indeksowych oraz innych usługodawców wspierających swoich użytkowników w dotarciu do udostępnianych zasobów,
- biblioteki akademickie szukające sposobów na efektywniejsze kierowanie swoich czytelników do zasobów o nieograniczonym dostępie.

Drugi etap prac zespołu NISO polegał na opracowaniu projektu roboczego nowego standardu. Projekt ten opublikowano 6 stycznia 2014 jako „Open Access metadata and indicators. Draft”² i obecnie jest na etapie opiniowania (termin mija 4 lutego 2014 r). Tekst tego dokumentu, na tle innych standardów, jest wyjątkowo zwięzły i przej-

¹ *A proposed NISO work item: specification for open access metadata and indicators* [on-line]. Baltimore: NISO, 2013 [Dostęp 15.01.2014]. Dostępny w: http://www.niso.org/apps/group_public/download.php/9845/Open%20Access%20Metadata%20-%20Work%20Item%20for%20ballot.pdf.

rzysty. Składa się z sześciu rozdziałów, które odpowiadają celom przyjętym przez grupę roboczą wymienionym w przedmowie projektu.

Podstawowym założeniem grupy roboczej było opracowanie możliwie najprostszego zestawu metadanych. Zarekomendowano dwa podstawowe znaczniki:

a. wolny dostęp (<free_to_read>) — proste określenie statusu publikacji informujące o tym, że dany tekst jest udostępniony on-line bez opłat i innych ograniczeń. Opcjonalnie znacznik można uzupełnić o atrybut z datą początkową (start_date), gdy publikacja nie jest dostępna od razu, ale po pewnym czasie od momentu jej wydania; można też uzupełnić o drugi atrybut z datą końcową (end_date), gdy publikacja jest przez jakiś czas dostępna za darmo, a potem już odpłatnie.

b. licencja (<license_ref>) – odnośnik do URI, pod którym znajdują się warunki użytkowania danej publikacji. Nie przewidziano żadnych ograniczeń przy wyszczególnianiu licencji i ich warunków, a metadane można multiplikować w przypadku obowiązywania kilku licencji. Każdy z tych elementów może dodatkowo zawierać własną datę początkową (start_date). Nie przypisano natomiast atrybutu z datą końcową, gdyż powodowałoby to zamieszanie, np. niektóre licencje (np. CC) nie określają dat końcowych; mogłyby się też tworzyć niezamierzone „przerwy” w obowiązywaniu poszczególnych licencji.

Te dwa elementy, "wolny dostęp" i "licencja", uznano za wystarczające do jednoznacznego określania statusu danej publikacji udostępnianej w module OA oraz warunków korzystania z niej. Jeśli publikacja jest dostępna bez żadnych ograniczeń (opłat i autoryzacji), zostaje oznaczona prostym znacznikiem <free_to_read>. Natomiast jego brak może oznaczać, że dokument nie jest w wolnym dostępie lub istnieją jakieś dodatkowe warunki odnośnie jego użytkowania, które można wyjaśnić za pomocą drugiego znacznika <license_ref>.

W projekcie znalazło się również zalecenie, aby znacznik licencji był obowiązkowym elementem wszystkich metadanych tworzonych dla artykułów z czasopism lub innych prac naukowych (rozdziałów z książek, prac pochodzących z materiałów konferencyjnych itp.) Zwrócono przy tym uwagę, że nie ma możliwości automatycznego powiadomienia o aktualnych warunkach udostępniania. To użytkownik (czytelnik) powinien za każdym razem sprawdzić adres URI licencji, aby wiedzieć, które warunki go obowiązują w momencie korzystania z danego zasobu. Wspomniano również o potrzebie opracowania wykazu istniejących licencji, dzięki czemu możliwe byłoby porównanie ich zapisów i zgodności. Pozwoliłoby to na uniknięcie sytuacji, gdy ten sam dokument jest przez jedną instytucję opublikowany w wolnym dostępie, a przez drugą odpłatnie.

² [Open access metadata and indicators](http://www.niso.org/apps/group_public/download.php/12047/rp-22-201x_OA_indicators_draft_for_comments.pdf). Draft [on-line]. Baltimore:NISO, 2014 [Dostęp 15.01.2014]. Dostępny w: http://www.niso.org/apps/group_public/download.php/12047/rp-22-201x_OA_indicators_draft_for_comments.pdf.

Szczegółowe zalecenia dotyczące prezentacji treści licencji skierowano do wydawców. Zwrócono przede wszystkim uwagę na to, aby informacje o warunkach udostępniania były napisane zrozumiałym językiem. Dzięki temu użytkownik nie będzie miał wątpliwości, co może zrobić z udostępnianą zawartością oraz jakie spoczywają na nim obowiązki. Teksty licencji powinny być dostarczane w postaci poprawnie zbudowanych stron HTML lub plików PDF.

Przedstawiono również ogólne zalecenia dotyczące dystrybucji znaczników OA. Zachęcono wydawców oraz usługodawców zajmujących się rozpowszechnianiem danych, aby włączyli je do swoich struktur metadanych, uwzględniając je także w metadanych kanałów komunikacji (np. RSS). Znaczniki te powinny też stać się integralnym elementem m.in. znaczników HTML META oraz plików PDF z wbudowanymi metadanymi bibliograficznymi itp. W dalszej perspektywie pozostają prace nad implementacją znaczników OA w istniejących strukturach metadanych, jak ONIX, RDF, OAI-PMH czy Dublin Core (DC).

Na uwagę zasługuje również rozdział piąty omawianego projektu, w którym przedstawiono dziewięć konkretnych przypadków użycia i oceny przydatności metadanych OA. Każdy przypadek opisany jest w szablonie składającym się z trzech kategorii informacji: lista podmiotów zaangażowanych w dany proces wykorzystania metadanych, opis kontekstu i roli metadanych oraz odpowiedź na pytanie, czy te metadane zaspokajają określoną w założeniu potrzebę oraz w jaki sposób jest lub powinno być to realizowane. Poszczególne przypadki analizują użyteczność metadanych OA z różnych perspektyw poczynając od użytkownika szukającego źródeł w otwartych zasobach oraz chcącego się dowiedzieć o warunkach korzystania z publikacji w całości (np. artykułu) lub fragmentu (np. abstraktu), po repozytoria dążące do pokazania wolnych zasobów będących w ich posiadaniu, a skończywszy na przypadku instytucji chcącej wykorzystać metadane OA do weryfikowania procesu zastosowania otwartego dostępu w odniesieniu do twórczości naukowej. W pięciu pierwszych przypadkach użycie metadanych OA może w pełni rozwiązać dany problem. W pozostałych czterech sytuacjach metadane mogą okazać się przydatne, przy czym wymagałaby zastosowania jeszcze dodatkowych narzędzi.

Warto wspomnieć, że grupa robocza przymierzała się też do opracowania odpowiednich znaków graficznych (ikon, logów) dla statusu publikacji i licencji, jednak w dalszej części projektu uznała to za wykraczające poza zakres działań, do których została powołana. Zapewne będzie to z czasem rozwiązane przez samych zainteresowanych upowszechnianiem zasobów w OA.

Wydaje się, że projekt przedstawiony przez grupę roboczą NISO, dzięki swojej prostocie, może wkrótce stać się realnie obowiązującym standardem. Bez wątpienia sprzyja temu fakt zaangażowania się w niego przedstawiciele instytucji naukowych, organizacji i środowisk kształtujących przestrzeń informacyjną i komunikacyjną (wydawcy, wyszukiwarki, promotorzy OA), a także instytucji pośredniczących między

dystrybutorami a końcowymi użytkownikami (np. bibliotekami i ich stowarzyszeniami). Warunkiem powodzenia jest, aby zapisy tego standardu zostały wprowadzone na stałe do struktur metadanych tworzonych przez wydawców oraz innych dostawców treści. Korzystne wydają się zalecenia dotyczące uwidocznienia informacji o statusie publikacji i warunkach ich użytkowania, dzięki czemu znacznie skróci się czas poświęcany przez użytkowników na odnajdywanie zasobów OA oraz formalności związane z legalnym ich wykorzystaniem.

Bibliografia:

1. *A proposed NISO work item: specification for open access metadata and indicators* [on-line]. Baltimore: NISO, 2013 [Dostęp 15.01.2014]. Dostępny w: http://www.niso.org/apps/group_public/download.php/9845/Open%20Access%20Metadata%20-%20Work%20Item%20for%20ballot.pdf.
2. *Open access metadata and indicators. Draft* [on-line]. Baltimore: NISO, 2014 [Dostęp 15.01.2014]. Dostępny w: http://www.niso.org/apps/group_public/download.php/12047/rp-22-201x_OA_indicators_draft_for_comments.pdf.