

## II. RECENZJE I OMÓWIENIA

### METADANE DOKUMENTÓW ELEKTRONICZNYCH<sup>1</sup>

Jedna z najnowszych publikacji traktujących o metadanych ukazała się w Stanach Zjednoczonych. Książka pt. *Metadata for Digital Resources. Implementation, Systems Design and Interoperability*, autorstwa Muriel Foulonneau i Jenn Riley, opublikowana została w 2008 roku przez wydawnictwo Chandos Publishing.

Muriel Foulonneau była koordynatorem projektu CIC-OAI na Uniwersytecie Illinois w Urbana-Champaign. Wcześniej pracowała jako konsultant w systemach zarządzania treścią i jako doradca IT dla francuskiego Ministerstwa Kultury. Brała udział w projekcie Minerva we współpracy europejskich ministerstw kultury nad dygitalizacją zasobów dziedzictwa kulturowego. Znalazła się także w składzie Komisji Europejskiej jako ekspert w projektach badawczych związanych z dziedzictwem cyfrowym i z infrastrukturami badawczymi. Jest współprezesem Dublin Core Collection Working Group wchodzącej w skład Dublin Core Metadata Initiative. Od września 2006 roku należy do Centre pour la Communication Scientifique Directe CNRS we Francji, gdzie pracuje nad archiwum publikacji naukowych HAL oraz nad europejskim projektem DRIVER (Digital Repository Infrastructure Vision for European Research). Napisała wiele artykułów o bibliotekach cyfrowych i metadanych, a także jest współautorką książki *Open Archives Initiative Protocol for Metadata Harvesting*.

Jenn Riley pracuje na stanowisku bibliotekarza na Uniwersytecie Indiana w Bloomington, gdzie odpowiada za planowanie strategii metadanych dla projektów biblioteki cyfrowej i współuczestniczy w projektowaniu bibliotek cyfrowych. Przyczyniła się do wydania przewodników o metadanych *Best Practices for OAI Data Provider Implementations and Shareable Metadata* oraz *Digital Library Federation/Aquifer Implementation Guidelines for Shareable MODS Records*. Interesuje się także włączaniem tezaursów do systemów wyszukiwawczych, cyfrowymi bibliotekami muzycznymi i FRBR. Jest autorką licznych artykułów opisujących różne aspekty rozwoju zbiorów biblioteki cyfrowej, a także bloga „Inquiring Librarian” (Dociekliwy Bibliotekarz), którego posty często koncentrują się wokół intelektualnego dostępu do materiałów bibliotecznych.

Książka *Metadata for Digital Resources*, z serii Chandos Information Professional Series, została podzielona na pięć części. W części pierwszej wyjaśnione zostało przede wszystkim pojęcie metadanych. Najczęściej definiuje się je jako dane o danych. Wychodząc z założenia, że nie jest to wyczerpujące wyjaśnienie, autorki zdecydowały się przytoczyć inne definicje. Na przykład zdaniem Priscilli Caplan metadane to ustrukturyzowana informacja o zasobie infor-

---

<sup>1</sup> Recenzja książki: M. Foulonneau, J. Riley: *Metadata for Digital Resources. Implementation, Systems Design and Interoperability*, Oxford 2008.

macyjnym dowolnego typu lub formatu medialnego. Lorcan Dempsey uważa, że metadane są danymi, które uwalniają użytkownika od konieczności posiadania pełnej wiedzy o istnieniu albo cechach przedmiotów potencjalnego zainteresowania. Wyjaśnienia i rady dane w tej książce koncentrują się na metadanych dla sektora dziedzictwa kulturowego: bibliotek, muzeów, archiwów, towarzystw historycznych itd. Metadane zostały trafnie porównane do etykiet na puszkach w supermarkecie. Gdyby produkty nie miały etykiet, musielibyśmy otwierać wszystkie puszki, żeby zobaczyć, co w nich jest, musielibyśmy analizować ich zawartość, aby wiedzieć, czy substancja jest jadalna, czy nie. Dzięki etykietkom wiemy, jakie jest jedzenie, z czego jest wykonane, skąd pochodzi, ile kosztuje, wreszcie znamy datę, do której możemy bezpiecznie spożyć produkt.

W dziedzictwie kulturowym metadane mogą podawać tytuł artykułu, datę napisania lub opublikowania, charakter kodowania używany w określonej wersji w sieci, nazwisko posiadacza praw autorskich itp. Autorki podkreślają wzrastające znaczenie metadanych. Ich zdaniem każdego można uznać za ich twórcę. Przykładem jest bibliotekarz sporządzający opis książki czy kustosz muzeum udostępniający informacje o cyfrowym wizerunku określonego obiektu. Jest tu również mowa o zastosowaniu metadanych w bibliotece cyfrowej. Ponieważ metadane stanowią ustrukturyzowaną informację o zasobie cyfrowym i jego własnościach, pozwalają na realizowanie różnorodnych funkcji, takich jak odkrywanie treści, interpretacja, ochrona, zarządzanie, reprezentacja i ponowne użycie obiektu.

Zasób cyfrowy może być udokumentowany przez jeden lub więcej rekordów. Rekord metadanych jest przykładem zbioru elementów zgodnych z danym formatem metadanych. Rekord może być przechowywany w tym samym miejscu co obiekt, poza obiektem albo dołączony do tego obiektu. Przykładem jest dokument TEI (Text Encoding Initiative) bądź strona tworzona w HTML, zawierająca nagłówki z opisowymi metadanymi i sam tekst główny. Każdy tekst zgodny z TEI zawiera nagłówki, który nie jest substytutem strony tytułowej, ale jest interpretacją znajdujących się na niej informacji bibliograficznych sporządzonych przez twórcę. Pewne formaty plików, np. SVG, wymagają specjalnych aplikacji, aby świadczyć o obiekcie i o jego metadanych. Katalog biblioteczny, w przeciwieństwie do wspomnianych dokumentów, gromadzi rekordy metadanych, a nie same obiekty cyfrowe. Instytucje dziedzictwa kulturowego coraz częściej używają systemów takich jak CONTENTdm, żeby gromadzić w tej samej aplikacji zarówno obiekty cyfrowe, jak i rekordy metadanych. W tym przypadku rekord jest przechowywany albo w bazie danych, albo jako oddzielny plik. Zarówno ta baza danych, jak i ten plik różnią się od zasobu cyfrowego. W metadanych zawiera się wiele typów informacji. Nie są to jedynie dane analogiczne do tych, z którymi mamy do czynienia w katalogach bibliotecznych. Metadane mogą się ukazywać pod różnymi określeniami, w rozmaitych typach, formatach i pochodzić z różnorodnych źródeł. Tagowanie, adnotacje i kodowanie tekstu to też metadane. Tagowanie zostało spopularyzowane przez Internet. Tag jest terminem związanym z obiektem a relacja między tagiem a obiektem nie jest sprecyzowana. Adnotacje są danymi związanymi z obiektem cyfrowym albo jedną z jego części reprezentujących metadane dodane do zasobu. W Internecie adnotacje mogą być dodane do obiektu i przydzielone większej grupie, są tym,

czym przypisy w tekstach drukowanych. Tworzy się je, aby umożliwić interpretację tekstu. Kodowanie tekstu polega na odwzorowaniu formatu oraz semantycznej struktury dokumentu, w czym mogą być pomocne nagłówki, nazwy miejscowe i przerwy między liniami w obrębie dokumentu.

Druga część książki opisuje tworzenie metadanych (rozdziały od drugiego do piątego). Rozdział drugi skupia się na wyborze standardów metadanych dla biblioteki cyfrowej, celu tworzenia metadanych, czynnikach decydujących o wyborze standardów metadanych oraz funkcjach standardów metadanych.

Rozdział trzeci ma za temat tworzenie wytycznych stosowania metadanych. Aby proces tworzenia metadanych był konsekwentny, ważny jest plan, tematy uwzględniane w wytycznych, a także ich pisanie, testowanie i udoskonalanie. Wybór odpowiednich standardów metadanych dla planowanych projektów cyfrowych oraz ich wszechstronne przetestowanie jest niezbędnym etapem tworzenia konsekwentnych metadanych, a tym samym lepszych serwisów. Aby upewnić się, czy metadane spełniają swoje cele, konieczne jest sprawdzenie opinii ekspertów, potrzeb użytkowników, upewnienie się, przez jaki czas systemy używane do tworzenia metadanych przechowują treść i czy dostarczają ją użytkownikom. Dokumentacja wytycznych użytkownika metadanych przydaje się jako instrukcja dla osób tworzących metadane. Pisanie wytycznych nie jest łatwe. Procedura tworzenia tej dokumentacji różni się w zależności od instytucji i projektów a najtrudniejsze jest określenie ilości informacji, którą chce się zawrzeć w instrukcji.

Rozdział czwarty omawia tworzenie samych metadanych, skupia się na narzędziach i metodologiach. Opisane zostały: proces tworzenia metadanych, projektowanie pracy nad tworzeniem metadanych, interfejsy użytkownika, tworzenie metadanych w XML, korzyści z nauki technologii XML oraz kontrola jakości tworzenia metadanych.

W rozdziale piątym autorki skoncentrowały się na praktycznym wdrażaniu strategii metadanych. Omówiono kryteria doboru personelu tworzącego metadane, rolę specjalisty od metadanych, problemy integrowania pracy w większej infrastrukturze, a także zwrócono uwagę na kwestie finansowe.

Część trzecia dotyczy projektowania systemów i składa się z dwóch rozdziałów. W rozdziale szóstym opisano funkcje pełnione przez bibliotekę cyfrową. Autorki zdefiniowały bibliotekę cyfrową jako zestaw narzędzi, który czyni dostępnymi zasoby cyfrowe i ich metadane. W centrum projektu systemu umieściły użytkownika. Nie chodzi tutaj tylko o użytkowników końcowych, takich jak klienci, goście i badacze, ale także o personel instytucjonalny. Na początku należy zdefiniować, kim są aktualni i potencjalni użytkownicy oraz zdecydować, którą z tych grup wesprzeć. Wśród funkcji bibliotek cyfrowych autorki wyróżniły identyfikowanie użytkowników i udostępnianie informacji, zarządzanie obiektami cyfrowymi, ochronę tych obiektów oraz identyfikowanie i wspieranie funkcji systemów nieznanych. Rozdział zamykają rozważania na temat celów, jakim powinny służyć metadane.

Rozdział siódmy zestawia metadane, które kierują wyszukiwaniem. Na początku podkreśla rolę badań użytkowników. Następnie analizuje metadane według różnych aspektów, m. in. jakości metadanych oraz poziomu szczegółowości. Omówiono także korzystanie z wyszukiwarek, przeglądarek, zaawanso-

wane wyszukiwanie, oraz wyświetlanie wyników. Część czwartą autorki przeznaczyły na przedstawienie współdziałania (ang. *interoperability*) metadanych. Rozdział ósmy stanowi próbę definicji współdziałania. Autorki odpowiadają na pytanie, dlaczego należy dzielić zbiory. Są instytucje, które wypracowały wiele serwisów online dla podkreślenia wartości swoich zbiorów. Takie serwisy są implementowane lokalnie. Wiele bibliotek oferuje strony internetowe, żeby użytkownik miał dostęp do katalogu komputerowego. Pozostałe serwisy można budować we współpracy z innymi instytucjami. W tym wypadku treść cyfrowa z jednej instytucji może być dostępna poprzez wiele różnych typów aplikacji. Dzięki temu zwiększa się krąg odbiorców indywidualnych. Następnie zostały wyliczone i omówione zbiory, a także serwisy. Serwis to aplikacja, która pozwala na wypełnienie dowolnego zadania, bądź funkcji biblioteki cyfrowej, zbiór to zestaw cyfrowych lub fizycznych zasobów, określony odpowiednimi zasadami. Najważniejsze części współdziałania to: współdziałanie techniczne, współdziałanie treściowe i współdziałanie organizacyjne. Współdziałanie systemów to ich porozumiewanie się między sobą. Aby umożliwić interakcję pomiędzy różnymi systemami, konieczne jest zdefiniowanie wspólnego języka oraz odpowiednie przygotowanie zbiorów do użytku przez systemy złożone. Współdziałanie dotyczy nie tylko dokumentowania całej wysyłanej informacji, gdyż efektywna komunikacja jest możliwa tylko wtedy, gdy inny system jest w stanie zinterpretować tę informację. Pracownicy informacji, którzy dzielą zbiory w sieci, muszą dostosować lokalne metadane do potrzeb innych systemów, aby umożliwić interakcje z zasobami w innym kontekście. Metadane odgrywają dużą i ważną rolę we współdziałaniu systemów, a tym samym w tworzeniu nowych serwisów w Internecie.

Rozdział dziewiąty traktuje o współdziałaniu i przeszukiwaniu zasobów. Wyjaśniono ogólne zasady przeszukiwania zasobów, zaprezentowano przeszukiwanie zasobów w innych powiązanych systemach. Przedstawiono najważniejsze wyszukiwarki, portale i mechanizmy automatycznego przeszukiwania.

Rozdział dziesiąty zajmuje się technicznym współdziałaniem. Omówiono tu scentralizowane i zdecentralizowane indeksy wyszukiwawcze, segregację danych, identyfikowanie i lokalizowanie zasobów oraz transformację rekordów metadanych.

W rozdziale jedenastym omówiono treściowe współdziałanie podzielnych metadanych, podano przykłady modeli segregacji i używania metadanych z różnorodnych źródeł, wyjaśniono, jak tworzyć podzielne metadane.

W części piątej autorki nakreśliły przyszłość metadanych. Powstanie generacja zautomatyzowanych metadanych. Coraz więcej ich będzie w Web 2.0, czyli w Internecie tworzonym przez użytkowników. Instytucje dziedzictwa kulturowego powinny zdefiniować strategię zarządzania metadanymi, aby efektywnie włączyć się w prace tworzenia środowiska informacyjnego, w którym poruszają się użytkownicy.

Książka *Metadata for Digital Resources* powstała na podstawie badań własnych, a także przy wykorzystaniu źródeł. Jako materiał źródłowy posłużyły czasopisma anglojęzyczne i francuskojęzyczne oraz dokumenty elektroniczne, których adresy internetowe podano w przypisach.

Metadane w pewnych formach zostały wykorzystane w instytucjach dziedzictwa kulturowego, wdrożone w określonych aplikacjach na przykład w katalo-

gach bibliotecznych. Coraz większą uwagę zwraca się na sposób, w jaki metadane wpływają na używalność, zarządzanie i ochronę cyfrowych zasobów.

Metadane to temat szeroki. Rozumiane jako dane o danych mogą znaleźć się w różnego rodzaju dokumentach, jednak dzisiaj terminu tego używa się głównie w odniesieniu do dokumentów elektronicznych. Znaczenie metadanych stale wzrasta. Ich elastyczność i modułowość są kluczowymi komponentami ewolucji serwisów online. Modułowość pozwala twórcom schematów metadanych tworzyć nowe połączenia elementów na bazie schematów utworzonych wcześniej. Typy metadanych, o których wspomina praca, mają realizować potrzeby użytkowników i różnych typów aplikacji, umożliwiać interakcję z zasobami cyfrowymi.

Książka *Metadata for Digital Resources* pokazuje jak można myśleć o metadanych przy budowaniu zbiorów w środowisku cyfrowym. Przeznaczona jest dla bibliotekoznawców zajmujących się metadanymi. Poleca się ją również bibliotekarzom prowadzącym prace nad katalogowaniem dokumentów, a także pracownikom muzeów w celu doskonalenia techniki przygotowania ekspozycji. Wszyscy ludzie pracujący w instytucjach dziedzictwa kulturowego powinni sięgnąć po tę lekturę.

Autorki dobrze zobrazowały problem. Napisały o nowościach w sporządzaniu opisów dokumentów. W pracy słusznie ograniczyły się do zasobów cyfrowych, ponieważ nadeszła era dygitalizacji. Bibliotekarze coraz intensywniej starają się sprowadzać gromadzone dokumenty do postaci elektronicznej.

Układ pracy jest starannie przemyślany. Rozdziały są podzielone na podrozdziały, dzięki czemu czytanie staje się łatwiejsze. Tekst został wzbogacony o zamieszczone w ramach przykłady opisów dokumentów, a także o ilustracje. Pozwala to na dokładniejsze zapoznanie się z amerykańskimi instytucjami dziedzictwa kulturowego.

W książce autorki operują skrótami bez rozwijania ich, (np. RDF, DTD i TEI) co dla laika, który mógłby sięgnąć po tę pracę może stanowić pewną trudność.

Na pewno powstanie w niedalekiej przyszłości wiele publikacji na temat metadanych. Jest to ciekawe zagadnienie i powinno być rozwijane przez współczesnych naukowców. Powstaje coraz więcej bibliotek cyfrowych, należy więc szukać nowych sposobów opisu książek i czasopism elektronicznych.

*Michał Grzechnik*