

Problemy jakości metaopisów w bibliotekach cyfrowych

Anna Domagalska¹

Streszczenie: Biblioteki cyfrowe umożliwiają korzystanie ze zbiorów za pośrednictwem Internetu każdemu użytkownikowi. Z dowolnego miejsca na ziemi i w dowolnym czasie użytkownik systemu bibliotecznego, ma wręcz natychmiastowy dostęp do poszukiwanej informacji. Twórcy bibliotek cyfrowych chcąc sprostać potrzebom użytkowników, oferują im coraz doskonalsze narzędzia do tworzenia, wyszukiwania i przeglądania zasobów online. Artykuł rozważa zagadnienie jakości metaopisów jako jednego ze składników jakości biblioteki cyfrowej. Model jakości zaproponowany przez autorkę w pracy (patrz Domagalska, 2006), opracowany w oparciu o normę ISO 11620 (patrz ISO 11620: 1998, ISO 11620: 1998/AD1:2003), zostaje rozszerzony o nowe wskaźniki jakości związane z metadanymi, które dotyczą zasobów zawartych w bibliotece cyfrowej. Autentyczne i poprawne metadane są gwarancją wiarygodności informacji, wykorzystywanej przez użytkownika np. w pracy naukowej. Teoretyczne rozważania dotyczące wpływu metadanych na jakość biblioteki cyfrowej, zostaną w przyszłości zweryfikowane w badaniach, w oparciu o rzeczywiste pomiary z wykorzystaniem różnych grup użytkowników w środowisku Cyfrowej Biblioteki Dokumentów Muzycznych realizowanego w ramach projektu MONIUSZKO.

Słowa kluczowe: biblioteki cyfrowe, jakość biblioteki cyfrowej, metadane, metaopis.

1. Wstęp

Dzięki bardzo szybkiemu rozwojowi Internetu, a w szczególności usługi World Wide Web, zaczyna urzeczywistniać się idea powszechnego i równego dostępu do informacji naukowej oraz dziedzictwa narodowego, dość często obecnie zapisanego w postaci elektronicznej. Szybki rozwój technologii informatycznych, przyczynił się do pojawienia na świecie i w Polsce projektów związanych z bibliotekami cyfrowymi. Powstały pierwsze biblioteki cyfrowe, cieszące się wielką popularnością wśród użytkowników, które zmieniły zasady i sposób dostępu do wiedzy. Ewolucja od bibliotek tradycyjnych do bibliotek cyfrowych zmieniła też pracę bibliotekarzy, a także upodobania i oczekiwania samych użytkowników. Twórcy bibliotek cyfrowych chcąc sprostać potrzebom użytkowników, systematycznie podnoszą jakość bibliotek cyfrowych. Wraz z pojawieniem się bibliotek cyfrowych pojawiło się wiele problemów, które należy rozwiązać. Jednym z nich jest problem jakości bibliotek cyfrowych. Jest on rozwiązywany w trakcie prowadzonych od 2005r. badań, w ramach projektu badawczego MONIUSZKO. Jakość w repozytoriach wiedzy

¹ Wydział Elektroniki, Telekomunikacji i Informatyki, Politechnika Gdańska, ul. Narutowicza 11/12, 80-952 Gdańsk
e-mail: {Anna.Domagalska@pg.gda.pl

związana jest przede wszystkim z jakością usług bibliotecznych. W pracy (patrz Domagalska, 2006) została przedstawiona problematyka jakości w kontekście potrzeb użytkowników, wieloznaczności definicji jakości oraz trudności związane z oceną jakości w bibliotekach cyfrowych. Niniejszy artykuł dotyczy jednej z charakterystyk drzewa jakości, a mianowicie - metadanych zasobów biblioteki cyfrowej, związanej z atrybutem wiarygodności oraz funkcjonalności. Kompletne i spójne metadane związane są z wiarygodnością zasobów biblioteki cyfrowej, a także z minimalnym czasem dostępu do poszukiwanego zasobu. Dodatkowo w artykule przedstawiono brak właściwego podejścia do problemu jakości przez amerykańską organizację National Information Standards Organization (NISO) (patrz National Information Standards Organization, 2004), a także jej brak w inicjatywie Unii Europejskiej i2010 (patrz Strona domowa - i2010, 2007) oraz w dokumentach Zespołu do Spraw Standardów Bibliotek Naukowych (ZdSBN) (patrz Strona domowa Zespołu do Spraw Standardów Bibliotek Naukowych, 2007).

2. Cyfrowa Biblioteka Dokumentów Muzycznych

Cyfrowa Biblioteka Dokumentów Muzycznych (patrz Szwoch, Michalska, 2005) jest projektem tworzonym przy współpracy badaczy Politechniki Gdańskiej i bibliotekarzy z Akademii Muzycznej w Gdańsku. Jej zasoby - w szczególności dokumenty muzyczne - są częścią dziedzictwa kulturowego o dużym zróżnicowaniu, jeśli chodzi o format zasobów. Celem wspomnianego projektu MONIUSZKO jest stworzenie cyfrowej biblioteki dokumentów muzycznych, która ma zapewnić skuteczną realizację podstawowych celów biblioteki, do których zaliczyć należy: przechowywanie, indeksację, wyszukiwanie oraz prezentację jej zróżnicowanych zasobów. Dokumenty muzyczne, będące zasobami biblioteki muzycznej są jednym z ważniejszych elementów kulturowego dorobku ludzkości. Ich złożoność i specyficzna natura, powoduje trudności z ich pozyskiwaniem w systemach bibliotek cyfrowych. Wyróżnić należy następujące postaci dokumentów muzycznych: tekst (opis bibliograficzny), obraz (partytura drukowana), muzyka (partytura cyfrowa, nagranie muzyczne inaczej dokument dźwiękowy), wideo (nagranie wideo np. z koncertu). Przechowywanie dokumentów muzycznych tylko w jednej z wyżej wymienionych postaci, znacznie ogranicza możliwości i zakres ich praktycznego wykorzystania. Natomiast cyfrowe dokumenty muzyczne, zgodnie z założeniami tworzenia bibliotek cyfrowych, powinny umożliwiać rozszerzenie funkcjonalności biblioteki cyfrowej o takie możliwości jak:

- wizualizacja partytury w postaci wektorowej, cechująca się wysoką jakością niezależnie od skali obrazu;
- interaktywna praca z dokumentami muzycznymi, która umożliwi zastosowanie zdefiniowanych uprzednio reguł transformacji dokumentów takich, jak: rozbijanie partytur na poszczególne głosy, zmianę aranżacji utworu, itp.;
- analizę semantyczną przechowywanych dokumentów, np. pod kątem wyszukiwania podobieństw między utworami, czy też wykrywania plagiatów;

- tworzenie multimedialnych zapytań w postaci tekstu, dźwięku, zapisu nutowego czy obrazu;
- odtworzenie brzmienia dokumentu muzycznego bezpośrednio z elektronicznej reprezentacji partytury w dowolnej aranżacji instrumentalnej.

Aby wyżej wymienione cele cyfrowej biblioteki dokumentów muzycznych mogły być w pełni realizowane, niezbędny jest odpowiedni projekt architektury w oparciu o standard, który zawierałby w sobie różne postaci dokumentów muzycznych. Z drugiej strony, by zapewnić sprawne funkcjonowanie biblioteki cyfrowej dokumentów muzycznych, opis bibliograficzny, a z nim metadane zasobów, są niezbędne. Poruszany w niniejszym artykule problem jakości bibliotek cyfrowych (a konkretnie wpływ metadanych na jakość biblioteki cyfrowej), jest jednym z zagadnień, które są rozwiązywane w ramach prac projektu MONIUSZKO. Przewiduje się, że w przyszłości ogólny model teoretyczny jakości dotyczący bibliotek cyfrowych zostanie zweryfikowany w oparciu o cyfrową bibliotekę dokumentów muzycznych.

3. Problem jakości bibliotek cyfrowych

Jakość w bibliotekach cyfrowych związana jest przede wszystkim z jakością usług bibliotecznych. Poniżej zostanie przedstawiony problem zagadnienia jakości w trzech różnych strategiach: NISO, i2010 oraz ZdSBN.

3.1. Strategia National Information Standard Organization

W związku z tworzeniem w ostatnich latach dużej ilości bibliotek cyfrowych oraz braku stosowania zasad, za które wiele inicjatyw było krytykowanych, w 2004 roku amerykańska organizacja National Information Standards Organization (NISO) wydała przewodnik tworzenia bibliotek cyfrowych, który ma realizować dwa zasadnicze cele:

- przedstawienie głównych składników i czynności związanych z tworzeniem dobrych zasobów cyfrowych,
- przedstawienie zasad identyfikacji, organizacji i stosowania istniejącej wiedzy i źródeł w celu promowania lokalnych działań służących tworzeniu i zarządzaniu dobrymi zasobami cyfrowymi.

Zastosowanie określenia dobre zasoby wymaga wyjaśnienia. Co innego oznacza bowiem dobre w początkowym etapie digitalizacji zbiorów, a co innego w okresie realizacji projektów dojrzałych. Projekty uznane za dojrzałe, powinny zaspokajać określone potrzeby odpowiednich grup użytkowników. W tym przypadku dobre zasoby są użyteczne i dostępne dla właściwie określonych grup użytkowników. Dodatkowo autorzy przewodnika NISO określają czynniki, które wpływają na ocenę bibliotek cyfrowych jako „dobrych”. Zaliczają do nich: zasoby (zorganizowane grupy obiektów), obiekty (zdigitalizowane materiały), metadane (informacja dotycząca obiektów), projekty (inicjatywy służące tworzeniu lub zarządzaniu zasobami). Z punktu widzenia celu artykułu, najistotniejszy jest problem jakości

metaopisów w bibliotekach cyfrowych, dlatego też poniżej zostały wymienione zasady tworzenia dobrych metadanych: *Zasada 1:* Dobre metadane powinny być odpowiednio dobrane do materiałów w zasobach, do użytkowników tych zasobów oraz planowanego, aktualnego i możliwego przyszłego wykorzystania obiektu cyfrowego. *Zasada 2:* Dobre metadane ułatwiają współdziałanie. *Zasada 3:* Dobre metadane stosują kontrolę autorytarną i narzędzia kontroli treści, takie jak słowniki kontrolowane, spełniające wymagania użytkowników co do opisu treści obiektów i łączenia ze sobą obiektów podobnych. *Zasada 4:* Dobre metadane zawierają jasno określone warunki i zasady użytkowania obiektu cyfrowego. *Zasada 5:* Dobre metadane ułatwiają długoterminowe zarządzanie obiektami w zasobach. *Zasada 6:* Rekordy dobrych metadanych także są obiektami, powinny więc być tworzone zgodnie z zasadami tworzenia dobrych obiektów, takimi jak autorytatywność, autentyczność, archiwizacja, niezmienność i jednoznaczna identyfikowalność. Zdaniem autorów NISO, powyższe czynniki, którym przyporządkowano odpowiednie zasady (w artykule wymieniono tylko zasady dotyczące tworzenia dobrych metadanych) są konieczne i wystarczające w celu stworzenia 'dobrego' projektu informacyjnego, jakim jest każda biblioteka cyfrowa. W przewodniku całkowicie został pominięty problem jakości biblioteki cyfrowej. Autorzy argumentują to tym, iż jakość biblioteki związana jest tylko z jakością usług.

3.2. Strategia europejska - I2010

W 2005r. członkowie Komisji Europejskiej odpowiedzialni za Społeczeństwo Informacyjne i Środki Przekazu zapoczątkowali inicjatywę, mającą na celu ponowne ożywienie wkładu technologii informacyjnych i komunikacyjnych tak, aby zgodnie z priorytetami Lizbońskimi przyczyniły się do wzrostu ekonomicznego i stworzenia nowych i lepszych miejsc pracy. Inicjatywa ta jest określona mianem „i2010”. Opiera się na trzech podstawowych filarach:

- po pierwsze, na wspólnej przestrzeni informacyjnej, stworzeniu nowoczesnych, zorientowanych na rynek ram dla gospodarki cyfrowej i stymulowaniu dostępności zawartości cyfrowej;
- po drugie, inwestycjach w badania i innowacje technologii informacyjnych i komunikacyjnych;
- po trzecie, koncentracji na europejskim społeczeństwie informacyjnym. Komisja Europejska od roku 2005 opowiada się za utworzeniem europejskiej biblioteki wirtualnej (Strategia i2010 - Biblioteki cyfrowe), mającej służyć udostępnieniu wszystkim zainteresowanym europejskim dorobkiem kulturowym i naukowym.

Podjęta na rzecz bibliotek cyfrowych inicjatywa ma sprawić, by europejskie zasoby informacji można było łatwiej i w bardziej interesujący sposób wykorzystywać w Internecie. W celu wykorzystania potencjału technologii cyfrowych dla rozpowszechniania i ułatwienia dostępu do informacji uwzględnione zostały trzy główne aspekty:

- dostęp online, który jest warunkiem wstępnym dla maksymalizacji korzyści, jakie informacje mogą przynieść obywatelom, naukowcom i przedsiębiorstwom;
- digitalizacja zbiorów analogowych, służąca ich szerszemu wykorzystaniu w społeczeństwie informacyjnym;
- konserwacja i przechowywanie, które zapewnią przyszłym pokoleniom dostęp do materiałów cyfrowych oraz zapobiegną utracie cennych zasobów.

Generując powyższe aspekty nie można pominąć funkcji jakie, metadane spełniają w opisie bibliotecznym. Zostaną one wymienione w rozdziale czwartym, który poświęcony jest metadanom. Jednak podobnie jak w przypadku strategii NISO tak i w strategii i2010 również została pominięta jakość biblioteki cyfrowej.

3.3. Strategia zespołu do spraw standardów bibliotek naukowych

W Polsce od 2002r. został powołany Zespół ds. Standardów dla Bibliotek Naukowych. Jego głównym zadaniem jest zbieranie danych na temat bibliotek naukowych, ich zasobów itp., które mają służyć do realizacji projektu „Analiza funkcjonowania polskich bibliotek naukowych w Polsce”. Głównym celem tego projektu jest wypracowanie metodologii i kryteriów pomocnych do oceny polskich bibliotek naukowych. Analiza danych ma na celu opracowanie metody oceny funkcjonowania bibliotek naukowych w Polsce oraz przygotowanie zestawu norm (standardów) dotyczących funkcjonowania bibliotek naukowych w Polsce. Z analizy danych zebranych od bibliotek uczestniczących w projekcie wynika, że jest w nich stosunkowo mało informacji na temat zbiorów cyfrowych. W chwili obecnej brak jakichkolwiek wniosków dotyczących metod oceny funkcjonowania bibliotek naukowych. Ponadto pomysł zespołu raczej nie spotkał się ze zbyt dużym uznaniem, gdyż tylko kilka bibliotek wypełniło ankietę (początkowy projekt zespołu), która w przyszłości ma służyć do dalszych analiz i porównań. Strategia Zespołu podobnie jak wymienione wcześniejsze dwie strategie NISO i i2010, pominęły jakość biblioteki cyfrowej. Pojawia się więc pytanie dlaczego tak się stało? Czy jest to problem, który rzeczywiście należy pominąć?

4. Metadane

Jednym z najtrudniejszych aspektów środowiska cyfrowego jest identyfikacja zasobów dostępnych w Web. Istnienie wyszukiwawczych metadanych opisowych zwiększa szansę, że zdigitalizowane treści zostaną odnalezione i wykorzystane we właściwym celu. Metadane jako ustrukturyzowana informacja dotycząca obiektu, mającą na celu umożliwienie jego wyszukania, opisu, wykorzystania, administrowania i zarządzania, mogą być dołączone na dowolnym etapie funkcjonowania obiektu cyfrowego. Zawierają charakterystykę cyfrowych obiektów informacyjnych (w przypadku bibliotek cyfrowych charakterystykę ich różnorodnych zasobów), która podczas pracy użytkownika z systemem wyszukiwawczym biblioteki cyfrowej, służy

efektywnemu oraz trafnemu wyszukiwaniu zasobów, których poszukuje użytkownik. Przekłada się to na jakość biblioteki cyfrowej, gdyż usatysfakcjonowany klient, nie tylko ponownie wybierze daną bibliotekę, ale poleci ją również znajomym. Właściwy opis metadanych jest szczególnie ważny w przypadku wielkich zasobów informacji Internetu (a takimi zasobami, na pewno są zasoby bibliotek cyfrowych), poza tym metadane wpływają na zarządzanie zasobami oraz ich wartościowanie. Najczęściej wyróżnia się trzy rodzaje metadanych:

- metadane opisowe, które pozwalają użytkownikowi odnaleźć obiekt, odróżnić od siebie różne obiekty i zrozumieć przedmiot lub treść obiektu;
- metadane administracyjne, które pomagają zarządzającym zasobami kontrolować obiekty dla takich celów, jak zarządzanie plikami, obsługa praw autorskich i archiwizacja;
- metadane strukturalne, które opisują relacje pomiędzy obiektami, np. związki pomiędzy artykułami, zeszytami i tomami czasopism lub stronami i rozdziałami książek.

Wykorzystanie szeroko dostępnych metadanych opisowych jest sposobem na maksymalizację udostępniania dla aktualnych użytkowników i przyciągnięcie nowych. Ze względu na sposób wykorzystania oraz twórcę, wyróżnia się dwa rodzaje metadanych:

- z punktu widzenia twórcy treści dokumentu elektronicznego, gdy metadane dotyczą informacji bibliograficznej o dokumencie;
- z punktu widzenia dostawcy serwisów internetowych, gdy metadane opisują elementy niezbędne do wyszukiwania informacji, np. dane o formatach, w którym źródło jest dostępne.

Niewątpliwie można powiedzieć, że tworzenie i utrzymanie metadanych wpływa na poprawienie jakości danych, polepszenie interakcji z systemem, wspomaganie procesu integracji systemu, wspomaganie utrzymania systemu, analiz i projektowania, automatyzację wielu czynności administracyjnych oraz zwiększenie elastyczności. Ponieważ przedmiotem zainteresowania artykułu jest problem jakości metaopisów w bibliotekach cyfrowych, należy wymienić funkcje, jakie metadane spełniają w zastosowaniach bibliotecznych, a w szczególności:

- dostarczają opisu bibliograficznego i charakterystykę treściową dokumentu elektronicznego;
- umożliwiają wyszukiwanie danych;
- zabezpieczają dane przed dostępem określonych grup użytkowników;
- lokalizują i określają dostępność dokumentu elektronicznego;
- dostarczają informacji o interpretacji danych;

- umożliwiają użytkownikowi wybór, w podjęciu decyzji, które wystąpienie danych ma być wyszukiwane (w sytuacji, gdy dostępnych jest kilka formatów);
- dostarczają informacji o warunkach wykorzystania danych z punktu widzenia prawa;
- informują o historii danych: o ich pierwotnym źródle oraz wszystkich kolejnych jego transformacjach;
- zawierają informacje o sposobie kontaktu, np. z właścicielem danych;
- określają relacje z innymi źródłami (np. linki z poprzednimi i kolejnymi wersjami, pochodne zestawy danych, inne kolejne zestawy danych, inne dane lub programy, które powinny być używane w połączeniu z tymi wyszukanyymi);
- kontrolują zarządzanie danymi (np. wymogi archiwizacji, autoryzacja usunięcia danych).

Niezbędne jest stworzenie kontroli autorytatywnej metadanych, gdyż zawierają one informacje świadczące o pochodzeniu, integralności i autentyczności obiektu. Tak zapisana informacja o metadanych, powinna zawierać identyfikację tworzącej ją instytucji oraz określenie standardów kompletności i jakości przyjętych podczas tworzenia. Instytucja powinna dostarczyć informację wystarczającą do weryfikacji metadanych przez użytkownika, w tym sposób tworzenia (manualny czy automatyczny) oraz wykorzystane standardy i słowniki. W większości przypadków za błędne metadane odpowiedzialny jest człowiek (w sytuacji, gdy tworzy opis bibliograficzny) lub maszyna (np. w przypadku generacji już istniejących metadanych). Ostatnio bardzo często do metadanych stosowane są te same techniki, które wykorzystuje się do weryfikacji integralności i autentyczności dokumentów cyfrowych (np. cyfrowe podpisy). Metadane w zastosowaniu do dokumentów World Wide Web wykorzystywane są do wyszukiwania informacji i służą opracowaniu dokumentów elektronicznych. Ze względu na to, iż każde środowisko (np. bibliotekarze, archiwiści, muzealnicy, wydawcy) stara się dostosować elementy metadanych do własnych potrzeb, istnieje wiele standardów metadanych służących opisowi różnych obiektów cyfrowych. Te różnorodne schematy są jednak na tyle spójne, że umożliwiają współdziałanie. Najpopularniejszymi standardami metadanych są: Dublin Core (patrz Nahotko, 2002; Strona domowa - Dublin Core (2007)), Marc21 (patrz Lont, Woźniak, 2006), MPEG-7 (patrz Skarbek, Keller, 2002) oraz inne. Przegląd dotychczasowych badań związanych z metadanymi, pozwala na wyróżnienie podstawowych założeń dotyczących wykorzystania metadanych, z których część jest realizowana, a część jest jeszcze w trakcie dopracowywania. Na uwagę zasługują jeszcze nie rozwiązane zagadnienia, takie jak: modułowość, rozszerzalność, sztywność i wielokulturowość. W Polsce używany jest Dublin Core i Marc21. Te dwa formaty służące do opisu metadanych zostaną omówione w dalszej części artykułu. Ponadto zostanie przedstawiony najlepszy z istniejących obecnie standard do opisu cyfrowych dokumentów muzycznych, a mianowicie MPEG-7. Jest on, nie tylko zdaniem autorki, standardem godnym polecenia w przypadku tworzenia

Tabela 1. Schemat elementów Dublin Core.

Zawartość (Content)	Własność intelektualna (Intellectual Property)	Dookreślenie (Instantiation)
Tytuł (ang. Title)	Twórca (ang. Creator)	Data (ang. Date)
Opis rzeczowy (ang. Subject)	Wydawca (ang. Publisher)	Typ (ang. Type)
Opis (ang. Description)	Współtwórca (ang. Contributor)	Format (ang. Format)
Źródło (ang. Source)	Własność (ang. Rights)	Identyfikator (ang. Identifier)
Język (ang. Language)		
Relacja (ang. Relation)		
Miejsce i Czas (ang. Coverage)		

opisów bibliograficznych zewnętrznych obiektów multimedialnych, które opisują muzykę.

4.1. Standard Dublin Core

Standard ten spełnia dominującą rolę, jako standard metadanych dla prostego opisu zasobów. Ułatwia także współpracę pomiędzy instytucjami używającymi różnych standardów metadanych. Format ten został stworzony przez bibliotekarzy i obecnie jest najczęściej wykorzystywanym formatem opisu dokumentów elektronicznych. Cechy, które charakteryzują Dublin Core to:

- prostota - nawet niezbyt wprawny użytkownik może się nim posługiwać
- spójność - dostarcza te same kategorie metadanych dla różnych typów dokumentów
- jednolitość - międzynarodowy charakter formatu pozwala na stosowanie go na całym świecie
- elastyczność - można tworzyć w nim zarówno proste, jak i bardzo złożone opisy
- dopasowanie - DC może pracować w różnych środowiskach i narzędziach.

Podstawą Dublin Core jest zestaw piętnastu elementów metadanych, które zostały wymienione w Tabeli nr 1.

Szczegółowy opis i definicja elementów Dublin Core w języku polskim znajduje się (patrz Nahotko, 2002) . Wszystkie elementy Dublin Core są opcjonalne i powtarzalne.

Przykładowe metadane tego artykułu w schemacie Dublin Core (HTML) wyglądają następująco:

```
<HEAD>
<TITLE>Problemy jakości metaopisów w bibliotekach cyfrowych</TITLE>
<META name="DC.title"
  content="Problemy jakości metaopisów w bibliotekach cyfrowych">
<META name="DC.creator" content="Domagalska,Anna">
<META name="DC.subject"
  content="biblioteki cyfrowe, jakość, NISO, metadane ">
<META name="DC.description" content="Biblioteki cyfrowe umożliwiają ... ">
<META name="DC.publisher" content="KKTPD 2007, PP">
<META name="DC.date" content="2007-09-24">
</HEAD>
```

4.2. Standard MARC21

Format Marc21 służy do przechowywania opisu bibliograficznego pozycji wydawniczych. Jedną z jego specyfikacji odnosi się do opisu wydawnictw muzycznych, lecz samej muzyki format ten nie opisuje. Niemniej jednak informacje bibliograficzne stanowią ważny element muzycznego dokumentu cyfrowego. Ponadto implementacje tego formatu są wykorzystywane w wielu systemach bibliotecznych, m. in.: w systemie Virtua VTLIS, Horizon i innych. Jednak żaden z wyżej wymienionych standardów nie jest wystarczający do opisu zasobów cyfrowej biblioteki muzycznej.

4.3. Standard MPEG-7

MPEG-7 jest językiem opisu zewnętrznych obiektów multimedialnych opisujących muzykę. Kluczową rolę w pracach nad standardem MPEG-7 odgrywa kwestia zapewnienia kompatybilności z innymi standardami opisu metadanych. Standard MPEG-7 może znaleźć zastosowanie wszędzie tam, gdzie podstawową rolę odgrywa informacja multimedialna. Dlatego też istnieje wiele różnych obszarów działalności, gdzie można go wykorzystać. Należą do nich między innymi:

- biblioteki cyfrowe, edukacja na odległość (np. katalogi obrazów, słowniki muzyczne, katalogi obrazowania biomedycznego);
- edycja materiałów multimedialnych (np. zindywidualizowane serwisy wiadomości, multimedialne narzędzia autorskie);
- kultura (np. muzea, galerie sztuki);
- katalogi multimedialne (np. informatory turystyczne, geograficzne systemy informacyjne, katalogi produktów firmowych);
- selekcja medium nadawczego (np. program radiowy, kanał telewizyjny);
- dziennikarstwo (np. wyszukiwanie przemówień danego polityka na podstawie jego nazwiska, głosu lub zdjęcia);

- handel elektroniczny (np. katalogi w sieci, wykazy sklepów elektronicznych, zindywidualizowany serwis reklamowy);
- monitorowanie i nadzór wizyjny (np. kontrola ruchu drogowego, transport naziemny, nieinwazyjne pomiary w warunkach trudnego dostępu);
- służby śledcze (np. rozpoznawanie cech charakterystycznych człowieka);
- rozrywka domowa (np. systemy zarządzania prywatnymi kolekcjami multimedialnymi z możliwością zmiany zawartości, w tym edycja amatorskich filmów wideo, wyszukiwanie gier, karaoke);
- robotyka (inteligentne interfejsy robotów, np. automatyczna selekcja obiektów na podstawie ich przykładowych obrazów).

Podstawowym celem standardu MPEG-7 jest stworzenie metod opisu, indeksacji i klasyfikacji obiektów multimedialnych. W tym celu konieczne jest stworzenie efektywnych metod wyszukiwania poszczególnych elementów, które mogłyby odbywać się na podstawie bardzo wielu różnych cech, między innymi na podstawie: metadanych (takich, jak: autor, producent, copyright), semantyki (obiekty, zdarzenia, ludzie), podziałów (regiony, segmenty), cech (tekstury, kolory, kształty). Niemniej jednak bezpośrednie wykorzystanie standardu MPEG-7 do opisu dokumentów muzycznych nie jest możliwe, gdyż nie pozwala na opis bibliograficzny tego dokumentu oraz na dołączenie tak ważnej postaci cyfrowego dokumentu muzycznego, jakim jest partytura cyfrowa. Wykorzystanie standardu MPEG-7 jest możliwe, gdyż język definicji deskryptorów pozwala na rozszerzenie zdefiniowanych w standardzie deskryptorów oraz schematów.

5. Zakończenie

Biblioteki cyfrowe są dziedziną nadal wymagającą dużych nakładów badawczych, w związku z tym, obecnie prowadzonych jest wiele prac związanych z zagadnieniem bibliotek cyfrowych. Istniejące rozwiązania wraz z rozwojem technologii, zapewne będą stopniowo ulepszone. Nowe wyzwania dla bibliotek stanowią coraz większe rzesze użytkowników, stopniowo zwiększające się zasoby informacji, a także zarządzanie i wyszukiwanie rozproszonych danych. Zagadnienie jakości, nie wiadomo dlaczego jest pominięte przez strategie światowe dotyczące problemu bibliotek cyfrowych (i2010, NISO). Autorzy NISO pomijają czynniki wpływające na jakość usług, gdyż przyjmują, że warunkiem usług dobrej jakości są wartościowe zasoby, obiekty i metadane (czynniki wymienione w przewodniku na temat tworzenia dobrych bibliotek cyfrowych). Wydaje się, że w ostatnim czasie o jakości (w większości przypadków życia codziennego, o jakości produktów, a w odniesieniu do bibliotek cyfrowych o jakości usług) mówi się bardzo dużo. Problem jakości metaopisów w bibliotekach cyfrowych jest jednym z zagadnień, poruszanych podczas rozważania jakości bibliotek cyfrowych. Aby spełnić funkcje metadanych w zastosowaniu bibliotecznym, należy weryfikować pracę człowieka w oparciu o odpowiedni do tego celu standard. Ponadto dodatkowym problemem do rozważenia jest zagadnienie związane z miarami dotyczącymi metadanych, do których należy zaliczyć:

- adekwatność opisów;
- ich kompletność;
- rozszerzalność opisów (w sensie dodawania do opisu bibliograficznego dodatkowych istotnych informacji o zasobie);
- kompatybilność metadanych.

Analiza wyżej wymienionych problemów to badania, jakie będą prowadzone w przyszłości i temat na kolejny artykuł.

Literatura

- DOMAGALSKA, A (2006) *Model jakości bibliotek cyfrowych*. Zeszyty Naukowe Wydziału ETI Politechniki Gdańskiej, Technologie Informacyjne Tom 10, Gdańsk.
- ISO 11620: 1998, ISO 11620: 1998/AD1:2003 : *Information and Documentation. Library performance indicators* Pr PN-ISO 11620 *Informacja i dokumentacja. Wskaźniki funkcjonalności bibliotek*.
- LONT, M., WOŹNIAK, E.(2006) *Opis bibliograficzny dokumentów w formacie Marc21 w systemie HORIZON*.
- NAHOTKO, M. [PRZEKŁAD] (2002) *Dublin Core Metadata Element Set, Version 1.1: Reference Description [online]*
<http://ebib.oss.wroc.pl/standard/dc.html>
- NATIONAL INFORMATION STANDARDS ORGANIZATION (2004) *A Framework of Guidance for Building Good Digital Collections [on-line]*. Wyd. 2. Bethesda: NISO Framework Advisory Group
<http://www.niso.org/framework/framework2.pdf>
- SKARBK, W., KELLER, T. (2002) *MPEG-7 w archiwizacji multimedialnych*. Konferencja MISSI'2002. Wrocław
- STRONA DOMOWA - DUBLIN CORE (2007) *Dublin Core [on-line]*
<http://purl.org/dc>
- STRONA DOMOWA - I2010 (2007) *i2010 [on-line]*
http://ec.europa.eu/information_society/activities/digital_libraries/index_en.htm
- STRONA DOMOWA - ZESPÓŁ DO SPRAW STANDARDÓW BIBLIOTEK NAUKOWYCH (2007) *Zespół do Spraw Standardów Bibliotek Naukowych [on-line]*
<http://ssk2.bu.amu.edu.pl/standaryzacja/>
- SZWOCH, M., MICHALSKA, A. (2005) *Interaktywna biblioteka cyfrowa dokumentów muzycznych*. Inforbazy'05, Bazy danych dla nauki, Red. A. Nowakowski, Warszawa: SIMP

SZWOCH, M., SZWOCH, W. (2005) *Formaty zapisu cyfrowych dokumentów muzycznych*. Zeszyty naukowe WETI PG, Konferencja Technologie Informacyjne, Gdańsk

The problems quality of meta descriptions in Digital Libraries

The digital libraries allow every single user for the access to the collections through Internet. From any place on the Earth and in any time, the library system user has immediate access to wanted information. The authors of digital libraries willing to met the users needs offer more and more advanced tools for creating, searching and viewing the online collections. The paper discusses the quality of meta descriptions as one of the digital library quality components. The quality model proposed by the author in (look Domagalska, 2006), based on ISO 11620 (look ISO 11620: 1998, ISO 11620: 1998/AD1:2003) is enhanced with new quality indicators related to meta data which refers to collection included in digital library. The authentic and correct meta data are guarantee of information reliability, used by user in i.e. research work. Theoretical discussion of meta data influence on digital library quality are to be verified in future research, basing on real measures with the use of various groups of users in environment.