

WSPÓŁDZIAŁANIE METADANYCH

Marek Nahotko
Instytut Informacji Naukowej
i Bibliotekoznawstwa
Uniwersytet Jagielloński

Współdziałanie systemów informacyjnych, współdziałanie metadanych, semantyka, syntaktyka, pragmatyka metadanych, model współdziałania metadanych, metody osiągnięcia współdziałania

Na początek ważna uwaga terminologiczna. Przyjęło się w Polsce określać zdolność do współpracy pomiędzy systemami „interoperacyjnością”, co jest kalką z języka angielskiego, gdzie używa się terminu *interoperability*. O ile pogodziliśmy się z brakiem dobrych polskich odpowiedników dla wielu terminów informatycznych, czego przykładem jest *interfejs*, który został nie tylko powszechnie zaakceptowany jako odpowiednik angielskiego *interface*, ale nawet doczekał się spolszczonej pisowni, to stosowanie kalki *interoperacyjność* uważam za niewłaściwe z kilku powodów. Po pierwsze, termin ten w języku polskim nie jest zrozumiały, dopóki nie zostanie zdefiniowany. Po drugie, istnieje polski termin, który może być stosowany; jest nim *współdziałanie*, którego używa między innymi Kazimierz Subieta w swoim słowniku¹, został też użyty w polskiej wersji normy dotyczącej modelu OSI², a więc jest dość rozpowszechniony.

Z tych powodów w dalszej części artykułu konsekwentnie stosowałem będąc termin *współdziałanie*, w rozumieniu podanych dalej definicji.

Problem współdziałania systemów informacyjnych

Współdziałanie systemów informacyjnych ma wiele znaczeń, jest różnie postrzegane. Zasadniczą cechą współdziałania jest istnienie relacji pomiędzy systemami, polegających na komunikacji, wymianie informacji, rozpowszechnianiu i współpracy. W efekcie powstaje system współdziałających systemów³.

¹ K. Subieta: *Słownik terminów z zakresu obiektowości*. Warszawa 1999, s. 218.

² PN-T-20000:1994. Systemy przetwarzania informacji. Współdziałanie systemów otwartych (OSI). Terminologia.

³ D. Carney, J. Smith, P. Place: *Topics in interoperability: infrastructure replacement in a system of systems*. Report nr CMU/SEI2005-TN-031. Pittsburgh 2005, p. 2.

Zazwyczaj różnice w definicjach współdziałania wynikają z odmiennych opisów tych relacji i elementów systemów. Autorzy części definicji skupiają się na elementach systemowych i sprzętowych, inni na usługach dostarczania i rozpowszechniania informacji, a jeszcze inni na sposobach wykorzystania wymienianej informacji bez konieczności wykonywania dodatkowych działań. Organizacje zarządzające polityką informacyjną zwracają uwagę na aspekty procesów niezbędnych dla wymiany i wielokrotnego wykorzystania informacji⁴.

Istnieje wiele prób definiowania współdziałania systemów informacyjnych. Kilka z nich warto przytoczyć:

„Dziedzina badawcza i praktyczna zajmująca się umożliwieniem współpracy niezależnie zbudowanych (heterogenicznych) systemów, szczególnie w sieciach komputerowych”⁵.

„Współdziałanie to możliwość wymiany danych pomiędzy wieloma systemami, z różnym sprzętem i oprogramowaniem, strukturami danych i interfejsami z minimalną utratą treści i funkcjonalności”⁶.

„Współdziałanie to zdolność dwóch lub więcej systemów lub ich części do wymiany informacji i wykorzystania wymienionej informacji bez wykonywania dodatkowych prac w tych systemach”⁷.

„Współdziałanie: taka kompatybilność dwóch lub więcej systemów, że mogą one wymieniać informację i dane oraz mogą stosować wymienioną informację i dane bez wykonywania specjalnych manipulacji”⁸.

Zgodnie z tymi definicjami, współdziałanie jest cechą systemów komputerowych (informatycznych); każdy system komputerowy składa się przynajmniej z dwóch współdziałających elementów: sprzętu i oprogramowania. Takie ograniczenie nie jest jednak niezbędne, gdyż pojęcie współdziałania może być w sposób zupełnie sensowny zastosowane do systemów technicznych spoza technologii informacyjnych, na przykład do systemu kolejnictwa lub sieci elektrycznej. Natomiast specyficzną cechą współdziałania systemów informacyjnych jest istotna rola metadanych, znajdując one miejsce w każdym modelu opisującym współdziałanie tych systemów.

Zasadniczym elementem w tych definicjach jest takie rozumienie określenia „używania” lub „wykorzystywania” informacji bądź danych, które implikuje użycie wymienianych danych w sposób, który jest zgodny z celami przewidzianymi dla systemu, w którym dane powstały. W przypadku metadanych oznacza to, że interpretacje danych jako opisu rzeczy powinny być spójne. Metadane tworzone w jednym systemie i następnie przeniesione do innego systemu będą przetwarzane przez ten drugi system w sposób spójny z intencjami twórcy metadanych.

⁴ M. Manso-Callejo, M. Wachowicz, M. Bernabé-Poveda: *Automatic metadata creation for supporting interoperability levels of spatial data infrastructure*. [online]. [dostęp: 15.01.2013]. Dostępny w World Wide Web: <<http://www.gsd.org/gsdiconf/gsd11/papers/pdf/194.pdf>>.

⁵ K. Subieta: op. cit., s. 218.

⁶ NISO: *Understanding metadata*. Bethesda 2004, p. 15.

⁷ M. Nilsson: *From interoperability to harmonization in metadata standardization*. Stockholm 2010, p. 12; CC:DA (ALCTS/CCS/Committee on Cataloging: Description and Access): *Task Force on Metadata. Final report*, June 16, 2000. [online]. [dostęp: 15.01.2013]. Dostępny w World Wide Web: <<http://www.libraries.psu.edu/tas/jca/ccda/tf-meta6.html>>.

⁸ A. Taylor, D. Joudrey: *The Organization of information*. Westport 2008, p. 269.

Jarosław Pacek proponuje odróżnianie współdziałania od kompatybilności, który to termin funkcjonuje głównie w informatyce i oznacza możliwość łącznego działania sprzętu lub oprogramowania komputerowego⁹. Mikael Nilsson natomiast pisze o harmonizacji metadanych jako o współdziałaniu systemów przy wielu schematach metadanych. W tym sensie harmonizacja metadanych dotyczy zdolności do poprawnego przetwarzania kilku różnych standardów metadanych w jednym systemie informatycznym¹⁰. Może być więc traktowana jako rodzaj współdziałania metadanych, o czym mowa będzie w dalszej części artykułu.

Stefan Gradmann skupia się na współdziałaniu wybranego rodzaju systemów komputerowych – bibliotek cyfrowych¹¹. Wyróżnia sześć aspektów współdziałania tych bibliotek:

- Jednostkami współdziałającymi są tradycyjne instytucje dziedzictwa kulturowego, takie jak biblioteki, muzea, archiwa. Oferują one usługi cyfrowe w cyfrowych repozytoriach (instytucjonalnych lub nie), platformach e-nauki i e-learningu lub innych serwisach Web.

- Obiekty interakcji, jednostki, które mają być przetwarzane na zasadzie współdziałania. Możliwości w tym zakresie rozciągają się od pełnotekstowych cyfrowych obiektów informacyjnych (analogowych/digitalizowanych bądź tworzonych jako cyfrowe) po jedynie reprezentacje takich obiektów – metadane.

- Funkcjonalny punkt widzenia na współdziałanie. Może ono dotyczyć jedynie wymiany i/lub rozpowszechniania treści cyfrowych, może także polegać na agregacji obiektów cyfrowych na jednolitym poziomie treści. Innym podejściem jest umożliwienie użytkownikom i/lub aplikacjom software'owym interakcji z wieloma bibliotekami cyfrowymi poprzez ujednoczone interfejsy (dynamiczne portale) lub umożliwienie działań w obrębie sfederowanych, autonomicznych bibliotek cyfrowych. Jeszcze inni dążą do utworzenia wspólnej architektury serwisu i/lub wspólnych definicji usług.

- Współdziałanie językowe (wielojęzyczność) może być uzyskiwane na dwa różne sposoby: przez stosowanie wielojęzycznych interfejsów użytkownika do bibliotek cyfrowych (rozwiązanie stosunkowo często stosowane) lub przez użycie dynamicznych technik wielojęzycznych dla eksploracji przestrzeni obiektów biblioteki cyfrowej. W tym drugim przypadku można wyróżnić trzy podejścia: dynamiczne tłumaczenie zapytań w celu ich przekazywania do biblioteki cyfrowej w różnych językach, dynamiczne tłumaczenie metadanych stanowiących odpowiedź na zapytania sformułowane w różnych językach i dynamiczną lokalizację treści cyfrowych.

- Projekt i punkt widzenia użytkownika. Koncepcje współdziałania osób zarządzających biblioteką cyfrową znacznie różnią się od koncepcji użytkownika końcowego. Sposób widzenia administratora technicznego będzie się bardzo różnił od postrzegania tego problemu przez użytkownika końcowego,

⁹ J. Pacek: *Bibliografia w zmieniającym się środowisku informacyjnym*. Warszawa 2010, s. 84.

¹⁰ M. Nilsson: op. cit., p. 14.

¹¹ S. Gradmann: *Interoperability. A key concept for large scale, persistent digital libraries*. [online]. [dostęp: 10.01.2013]. Dostępny w World Wide Web: <<http://www.digitalpreservationeurope.eu/publications/briefs/interoperability.pdf>>.

dostarczającego treści jako ich autor. Jeszcze inaczej współdziałanie rozumieją agregatorzy treści cyfrowych, użytkownicy metadanych i ciała zarządcze.

- Standardy techniczne umożliwiające różne rodzaje współdziałania funkcjonują obok bardziej tradycyjnych rozwiązań służących współdziałaniu metadanych bibliotecznych, na przykład protokół Z39.50 (SRU/W) lub metody zbierania metadanych za pomocą OAI-PMH; stosuje się też metody oparte na usługach Webowych (SOAP/UDDI)¹² i API tworzonych w Javie, definiowanych w JCR (JSR 170/283)¹³ lub platformach gridowych, takich jak iRods¹⁴.

Współdziałanie bibliotek cyfrowych może być postrzegane na różnych poziomach abstrakcji. Poziomy te przecinają inne, przedstawione wcześniej, podziały. Ich uporządkowanie w kolejności od najbardziej konkretnych do abstrakcyjnych pozwala na uzyskanie czterech poziomów współdziałania systemów informacyjnych, jak na rys. 1. Zauważmy, że współdziałanie metadanych uznane zostało za element syntaktycznego współdziałania bibliotek cyfrowych (poziom drugi).



Rys. 1. Rodzaje współdziałania i poziomy abstrakcji (wg S. Gradmanna)

Bardziej rozbudowany model współdziałania systemów informacyjnych przedstawił Miguel Manso-Callejo¹⁵ z zespołem¹⁶. Na podstawie literatury wyróżnili oni następujące poziomy współdziałania (zob. rys. 2):

- **techniczny**: bity i bajty, pliki, przetwarzanie, protokoły komunikacyjne na poziomie sprzętu i oprogramowania;
- **syntaktyczny**: wspólne formaty i struktury danych, języki takie jak XML;
- **semantyczny**: dotyczy znaczenia informacji, obejmuje wspólne słowniki terminów umożliwiające interpretację ich znaczenia; współdziałanie semantyczne wspomaga standardy i specyfikacje definiujące schematy wymiany informacji;

¹² SOAP – Simple Object Access Protocol; UDDI – Universal Description, Discovery and Integration.

¹³ JCR – Java Content Repository, specyfikacja API dla platformy Java, służącego do udostępniania repozytoriów treści, w tym metadanych. Zapisany jako JSR-170 i JSR-283 (JSR – Java Specification Request, standard społeczności Java).

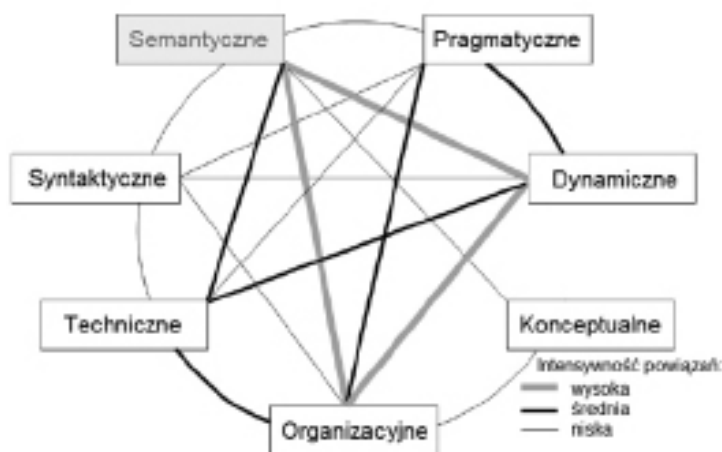
¹⁴ <https://www.irods.org/index.php/>. iRods funkcjonuje jako middleware pomiędzy użytkownikiem a zasobami sieciowymi.

¹⁵ M. Manso-Callejo, M. Wachowicz, M. Bernabé-Poveda: op. cit.

¹⁶ M. Manso, M. Wachowicz, M. Bernabé: *Towards an integrated model of interoperability for spatial data infrastructures*. „Transactions in GIS” 2009, vol. 13, no 1, p. 54.

- **pragmatyczny**: metody i procedury dla obsługi wymienianych danych, definicje interfejsów
- **dynamiczny**: nadzór transferu danych, monitoring sieci i innych systemów;
- **konceptualny**: modelowanie systemu i danych w standardowej dokumentacji, bez względu na zastosowany rodzaj modelu;
- **organizacyjny**: cele biznesowe, polityka dostępu i wykorzystania danych, modelowanie procesów, wzorce postępowania, odpowiedzialności (przynajmniej podczas oceny systemu).

Współdziałanie systemów:



Rys. 2. Poziomy współdziałania systemów informacyjnych wg zespołu Manso-Callejo i in.

Metadane są tworzone i stosowane na każdym z wymienionych poziomów współdziałania, szczególną rolę odgrywają na poziomach semantycznym, dynamicznym i organizacyjnym.

W europejskim projekcie DL.org¹⁷ w celu stworzenia pełnego obrazu problemów współdziałania systemów informacyjnych wyróżniono trzy poziomy współdziałania, zgodnie z terminologią European Interoperability Framework for pan-European eGovernment services¹⁸:

- współdziałanie organizacyjne: odnosi się do współpracy pomiędzy i w obrębie bibliotek cyfrowych jako instytucji, ich celów biznesowych i modelowania procesów. Jest to najtrudniejszy poziom współdziałania, szczególnie na poziomie danych czytelnych maszynowo i procesów zautomatyzowanych;
- współdziałanie semantyczne: dotyczy rozumienia znaczenia informacji gromadzonych w bibliotekach cyfrowych;
- współdziałanie techniczne: dotyczy połączeń, prezentacji i wymiany obiektów cyfrowych oraz problemów dostępności i bezpieczeństwa.

¹⁷ DL.org Digital Library Interoperability, Best Practices and Modelling Foundations [online]. [dostęp: 2.05.2013]. Dostępny w World Wide Web: <<http://www.dlorg.eu/>>.

¹⁸ European Commission: *European interoperability framework for pan-European e-government services*. Luxembourg 2004, p. 16.

Jak określono w DELOS Digital Library Manifesto¹⁹, biblioteka cyfrowa jest skomplikowanym układem, w którym można wyróżnić trzy odrębne konceptualnie systemy: biblioteki cyfrowe (DL), systemy bibliotek cyfrowych (DLS) i systemy zarządzania biblioteką cyfrową (DLMS). W modelu DELOS przedstawione są współdziałające elementy systemów informacyjnych. Wyróżnione zostały tam:

- Współdziałanie treści: jest to najczęstsza forma współdziałania, wymaga uzgodnienia danych prymarnych, schematów, metadanych.
- Współdziałanie funkcjonalne: budowa serwisów dla tworzenia złożonych przepływów danych, zwykle ma charakter syntaktyczny.
- Współdziałanie użytkownika: gromadzenie, wymiana i integracja informacji o użytkownikach, takich jak profile, preferencje, prawa dostępu.
- Współdziałanie jakościowe: służy ocenie systemów zewnętrznych i optymalizacji własnych.
- Współdziałanie na poziomie polityk: polityki wyrażane jako zasady i uregulowania dotyczące treści, funkcji i użytkowników²⁰.

Jak widać, współdziałanie systemów to skomplikowane i wielopoziomowe zagadnienie. Na jednym z poziomów zawsze uwzględniane są metadane, które wykorzystywane są na wszystkich innych poziomach systemu informacyjnego.

Współdziałanie metadanych

Jak wynika z poprzedniej części artykułu, metadane i ich współdziałanie są istotną częścią współdziałania systemów informacyjnych. Często nie jest łatwo wyznaczyć granicę pomiędzy tym, co należy do współdziałania metadanych, a tym, co dotyczy współdziałania systemów. Muriel Foulonneau i Jenn Riley²¹ piszą np. o współdziałaniu technicznym – dzięki któremu komunikaty przesyłane pomiędzy systemami w wyniku stosowania odpowiednich protokołów mogą być wzajemnie zrozumiałe dla oprogramowania tych systemów – jako o warunku do współdziałania semantycznego i syntaktycznego metadanych. Tak rozumiane współdziałanie techniczne będą jednak traktował nie jako współdziałanie metadanych, lecz systemów informacyjnych. Zastosowanie definicji współdziałania systemów informacyjnych, przedstawionych wcześniej do problemów metadanych pozwala na przedstawienie następującej definicji współdziałania metadanych:

„Współdziałanie metadanych to zdolność dwóch lub więcej systemów lub ich składników do wymiany danych opisowych o rzeczach oraz do interpretowania tych danych opisowych, które podlegały wymianie w sposób spójny z interpretacją twórcy danych”²².

¹⁹ L. Candela [et al.]: *Setting the foundations of digital libraries. The DELOS Manifesto*. „D-Lib Magazine” 2007, vol. 13, no 3/4. [online]. [dostęp: 15.01.2013]. Dostępny w World Wide Web: <<http://www.dlib.org/dlib/march07/castelli/03castelli.html>>.

²⁰ L. Candela [et al.]: *The DELOS digital library reference model. Foundations for digital libraries*. Ver. 0.98 [online]. [dostęp: 15.01.2013]. Dostępny w World Wide Web: <<http://www.delos.info/ReferenceModel>>.

²¹ M. Foulonneau, J. Riley: *Metadata for digital resources: implementation, systems design and interoperability*. Oxford 2008, p. 139.

²² M. Nilsson: op. cit., p. 13.

Najważniejszym elementem przedstawionej definicji jest wymiana informacji lub danych, co implikuje stosowanie wymienianych danych zgodnie z intencjami twórców systemu, w którym dane powstały. W przypadku metadanych istnieje wymóg spójności interpretacji danych jako opisu „rzeczy”. Oznacza to, że metadane utworzone przez bibliotekarza w jednym systemie, a następnie przeniesione do innego systemu, będą przetwarzane w tym drugim systemie w sposób spójny z intencjami twórcy metadanych²³.

Możliwość współdziałania metadanych traktowana jest jako jedna z najważniejszych cech umożliwiających ich stosowanie, obok prostoty, modularności, wielokrotnego użycia (*reusability*) i rozszerzalności²⁴. Cechy te powinny stanowić podstawowe wyznaczniki projektowania oraz wdrażania systemów i projektów metadanych. W publikacji IFLA²⁵ podkreślana jest rola współdziałania zarówno w systemach tradycyjnych, jak i elektronicznych oraz stosowania standardów w celu osiągnięcia współdziałania. Rzeczywiście, znakomita większość standardów stosowanych w związku z tworzeniem i wykorzystywaniem metadanych w taki czy inny sposób służy osiągnięciu ich współdziałania.

Decyzje co do stosowania metadanych podejmowane są na wielu etapach tworzenia systemów informacyjnych, na przykład bibliotek cyfrowych. Część z nich podejmowana powinna być jeszcze zanim powstanie pomysł utworzenia jakiegokolwiek systemu. Niektórzy autorzy wskazują na potrzebę stosowania w skali globalnej jednego JIW (języka deskryptorowego lub jhp), wspólnej khw oraz tworzenia pojedynczego rekordu dla każdego dzieła, który mógłby być wykorzystywany przez wszystkie istniejące i nowo powstające systemy informacyjne²⁶. Z punktu widzenia współdziałania metadanych najprostszym przypadkiem jest podjęcie decyzji o stosowaniu przez wszystkie współpracujące biblioteki wspólnego schematu metadanych, takiego jak MARC lub Dublin Core, co zapewnia wysoki poziom spójności tworzonych danych. Nawet stosowanie tego samego schematu nie zapewnia jednak pełnego współdziałania, o czym świadczy potrzeba harmonizacji formatu MARC pod koniec XX w., w wyniku której powstał MARC 21, oraz problemy polskich bibliotek stosujących oprogramowanie dLibra i format Dublin Core²⁷. W tym drugim przypadku problemy ze współdziałaniem modyfikowanego standardu spowodowały próby tworzenia profilu aplikacyjnego PLMET, który miałby ujednoclić stosowane struktury metadanych poprzez uwzględnienie wszystkich niezbędnych modyfikacji²⁸. Jak więc widać, mimo iż stosowanie wspólnego schematu metadanych jest konceptualnie prostym

²³ Można także mówić o spójności metadanych w ramach jednego systemu, co oznacza stosowanie jednolitych zasad tworzenia metadanych w systemie w dłuższym okresie czasu i nie jest traktowane w artykule jako problem związany ze współdziałaniem.

²⁴ M. Zeng, J. Qin: *Metadata*. New York 2008, p. 268.

²⁵ IFLA: *Sharing of bibliographic information and resources*. [online]. [dostęp: 15.01.2013]. Dostępny w World Wide Web: <<http://archive.ifla.org/VII/d4/pub/InteroperabilityStandards.pdf>>.

²⁶ I. Tolkoff: *The path toward global interoperability in cataloging*. „Information Technology and Libraries” 2010, vol. 29, no 1, p. 31.

²⁷ J. Potęga, A. Wróbel: *The Dublin Core Metadata Element Set. Ver. 1.1 a potrzeby i oczekiwania bibliotekarzy cyfrowych – analiza przypadków*. W: *Polskie biblioteki cyfrowe 2009*. Materiały z konferencji. Pod red. C. Mazurek, M. Stroiński, J. Węglarz. Poznań 2010, s. 77.

²⁸ M. Mielnicki [et al.]: *Agregacja metadanych w skali kraju – kierunki rozwoju Federacji Bibliotek Cyfrowych*. Warszawa 2012, s. 75.

rozwiązaniem, bywa że nie do końca zdaje egzamin. Przy tym nie zawsze jest możliwe lub praktyczne, szczególnie w heterogenicznym środowisku sieci rozległych, gdzie może prowadzić do izolacji (jak w przypadku formatu MARC, stosowanego wyłącznie w aplikacjach bibliotecznych). Według Kima Veltmana²⁹ poszukiwania pojedynczego, poprawnego ontologicznie rozwiązania w zakresie metadanych nie uwzględnia pragmatycznej rzeczywistości dominującej w wielu instytucjach. Brak takiej możliwości występuje na przykład w przypadku, gdy użytkowane jest środowisko obsługujące różne społeczności, gdy zasoby są już opisane przy pomocy różnych wyspecjalizowanych schematów. Taka sytuacja może wystąpić, gdy polscy bibliotekarze zechcą współpracować z muzealnikami w zakresie wymiany danych. W tym drugim środowisku stosowane są schematy metadanych inne niż dla zasobów bibliotecznych i ich rozwój następuje w innym kierunku niż w bibliotekach³⁰. Ponieważ standaryzacja przez unifikację możliwa jest tylko na początku lub na pierwszych etapach budowy repozytoriów, przed implementacją różnych schematów przez twórców projektów, w przypadku przyszłej współpracy twórców bibliotek cyfrowych i e-muzeów trzeba będzie szukać innych rozwiązań.

Marcia Zeng i Jian Qin wyróżniły kilka możliwych przypadków postępowania w trakcie prac, od utworzenia schematu do jego wdrożenia w pojedynczych bibliotekach cyfrowych lub połączonych repozytoriach³¹:

- schemat metadanych został utworzony i zastosowany w jednym lub kilku projektach;
- wzięto pod uwagę elementy metadanych z kilku schematów. Utworzono profil aplikacyjny na podstawie kilku schematów; następnie zestaw elementów metadanych wyspecyfikowanych w profilu aplikacyjnym zastosowano do tworzenia rekordów w określonym projekcie;
- dokonano wymiany lub integracji zawartości dwóch lub więcej baz danych, zawierających rekordy metadanych, w oparciu o porównanie znaczenia elementów odpowiednich schematów metadanych;
- rekordy z istniejących zasobów metadanych zostały pobrane lub włączone do jednolitego repozytorium metadanych. Przed pobraniem w zasobach tych stosowane były różne schematy metadanych lub tworzone były własne profile aplikacyjne.

Każde z tych sposobów postępowania ma wyraźnie określony cel, a wysiłki służące zapewnieniu współdziałania metadanych pojawić się mogą w każdym miejscu.

W ostatnich latach rozwiązanie problemu współdziałania różnych schematów metadanych w różnych aplikacjach staje się coraz bardziej istotne. Prace w tym zakresie mogą być przedstawiane na różnych poziomach i z różnych punktów widzenia:

- Poziom schematu metadanych: prace skupiają się na elementach schematów, które są niezależne od zastosowań. Projektant schematu powinien

²⁹ K. Veltman: *Syntactic and semantic interoperability: new approaches to knowledge and the semantic web*. „New Review of Information Networking” 2001, vol. 7, no 1, p. 175.

³⁰ W. Klenczon: *Międzynarodowe standardy opisu obiektów muzealnych jako podstawa katalogowania zbiorów*. Warszawa 2011, s. 43.

³¹ M. Zeng, J. Qin: op. cit., p. 268.

zapewnić współdziałanie metadanych tworzonych przy jego pomocy jeszcze przed rozpoczęciem tworzenia rekordów. W efekcie tych prac zwykle powstają: poходne zestawy elementów (wyprowadzanie nowego schematu z elementów istniejącego: tworzenie nowych wersji, tłumaczenie na inne języki, przeróbki dla dostosowania do lokalnych potrzeb, np. MODS wyprowadzony z MARC 21); struktura nadrzędna (szkielet pozwalający na integrację różnych schematów, np. OAIIS dla metadanych archiwalnych); tablice przejścia (mapowanie elementów, semantyki i syntaktyki z jednego schematu metadanych do innego; pozwalają na efektywną konwersję danych w jednym standardzie metadanych do innego, dzięki czemu heterogeniczne zasoby mogą być przeszukiwane jednocześnie na podstawie jednego zapytania – mapowanie może odbywać się bezpośrednio pomiędzy dwoma schematami lub za pośrednictwem tzw. schematu przełącznikowego, do którego mapowana jest dowolna liczba schematów metadanych; mapowanie może oznaczać utratę części semantyki); profile aplikacyjne (łączenie elementów pochodzących z kilku schematów, dostosowanie wybranego schematu do określonych potrzeb, jak DC-LAP,³² lub dołączanie własnych elementów do istniejącego schematu, co oznacza utworzenie i zarządzanie nową przestrzenią nazw); oraz rejstry elementów metadanych (zawierają dane dotyczące schematów metadanych w celu ułatwienia wielokrotnego stosowania w różnych schematach istniejących, wcześniej zastosowanych terminów; mogą dotyczyć jednego lub wielu schematów, projektu, dziedziny lub formatu).

- Poziom rekord metadanych: prace służą integracji rekordów metadanych poprzez mapowanie elementów zgodnie ze znaczeniem (semantyką) tych elementów. Stosowane są w przypadku, gdy rekordy metadanych już istniały przed powstaniem potrzeby współdziałania systemów. Wówczas za późno jest na współdziałanie na poziomie schematu, co wymusza prace na istniejących rekordach współdziałających baz danych. Efektem zazwyczaj jest zbiór skonwertowanych rekordów (problemem jest utrata lub przekłamanie danych, szczególnie podczas konwersji

wirtualna baza danych, do której metadane są pobierane z różnych źródeł, ułatwia współdziałanie dzięki tworzeniu jednolitego i wiarygodnego narzędzia do udostępniania danych. Zapewnienie współdziałania na poziomie repozytorium może być związane z takimi działaniami jak pobieranie metadanych (np. przy pomocy protokołu OAI-PMH), obsługa wielu formatów, agregacja, usługi przejścia (ang. *crosswalk*), mapowanie oparte na wartościach metadanych służące wyszukiwaniu w wielu zasobach jednocześnie i mapowanie współwystępowania oparte na wartościach metadanych. Podstawową decyzją, którą należy podjąć, związaną ze współdziałaniem na poziomie repozytorium, jest to, czy metadane z różnych źródeł będą zachowywać własny, pierwotny format metadanych. Jeżeli nie, to trzeba określić sposoby konwersji lub integracji wszystkich rekordów metadanych do formatu przyjętego jako standardowy. Jeśli tak, to trzeba zapewnić możliwość prowadzenia jednoczesnego wyszukiwania we wszystkich zasobach. Współdziałanie na tym poziomie osiągane jest poprzez zbieranie (ang. *harvesting*) metadanych (np. z użyciem protokołu OAI-PMH), stosowanie wielu formatów bez konwersji rekordów (np. DLESE Collection System), agregację elementów metadanych i ich wartości z różnych źródeł, tablice przejścia, mapowanie wartości metadanych dla wyszukiwania wielu kolekcji (np. haseł przedmiotowych w MACS) i mapowanie współwystępowania wartości metadanych (mapowanie zawartości słowników na podstawie współwystępowania tych wartości elementów w rekordach)³⁴.

Bernhard Haslhofer i Wolfgang Klas do wyżej wymienionych poziomów (rekordy-schemat-baza danych) dodali jeszcze czwarty poziom modelu bazy danych, na którym funkcjonują uniwersalne języki modelowania³⁵. Według autorów problemy ze współdziałaniem metadanych wynikają z pojawiania się niejednorodności na tych poziomach. Współdziałanie metadanych może być osiągnięte przez eliminację lub ominięcie tych niejednorodności. Wyróżnili oni trzy sposoby osiągania współdziałania: 1) uzgodnienia na poziomie schematu metadanych; 2) wprowadzenie i uzgodnienie wspólnego modelu bazy danych (repozytorium); 3) uzgodnienie niejednorodności strukturalnych i semantycznych na poziomie rekordu oraz schematu i bazy danych. W tab. 1. przedstawione zostały dostępne techniki uzyskiwania współdziałania metadanych z podziałem na wspomniane trzy kategorie.

³⁴ M. Zeng, L. M. Chan: *Metadata interoperability and standardization – a study of methodology part II. Achieving interoperability at the record and repository levels*. „D-Lib Magazine” 2006, vol. 12, no 6. [online]. [dostęp: 15.01.2013]. Dostępny w World Wide Web: <<http://www.dlib.org/dlib/june06/zeng/06zeng.html>>.

³⁵ B. Haslhofer, W. Klas: *A survey of techniques for achieving metadata interoperability*. „ACM Computing Survey” 2010, vol. 42, no 2, p. 11.

Narzędzia służące współdziałaniu metadanych³⁶

	Uzgodnienie schematu		Uzgodnienie modelu bazy danych	Uzgodnienia na trzech poziomach	
Baza danych	Hybrydowy system metadanych (np. MPEG-7, TV-Anytime)	Standardowy język (np. OWL, UML, XMLS)	Model metadanych (np. MOF) Abstrakcyjny model metadanych (np. Abstrakcyjny model DCMI)	Mapowanie języka	
Schemat		Standardowy schemat metadanych (np. DC, TEI, MODS)	Globalny model konceptualny (np. CIDOC-CRM, FRBR)	Struktura nadrzędna (np. MPEG-21, METS, OAIS)	Tablice przejścia, mapowanie schematów
			Profil aplikacyjny (np. DCAP)		
Rekordy		Schemat kodowania wartości (np. normy ISO)	Słownik kontrolowany (np. LCSH, UKD, KABA)	Rekordy KHW (np. LoC Authorities)	Transformacje na poziomie rekordów

Według Stuarta Weibela³⁷ wymiana metadanych wymaga porozumienia na trzech poziomach:

- **Semantyki:** znaczenie nadawane stwierdzeniom budowanym przy pomocy metadanych. Semantyka jest domeną ludzi. Uzgadnianie semantyki polega na budowie uzgodnionej struktury elementów metadanych (takiej jak DCMES).

- **Syntaktyki:** metody kodowania elementów metadanych w taki sposób, aby można je było przesyłać pomiędzy komputerami, a po przesłaniu odkodowywać i automatycznie przetwarzać, w tym wyświetlać dla użytkowników-ludzi. W trakcie tego przetwarzania semantyka nie powinna ulegać zmianom. W tym celu stosowane są standardy (takie jak RDF i XML).

- **Struktur:** Niezawodność syntaktyki opiera się na istnieniu jednoznacznej struktury danych. Jej specyfikacją jest model danych. Wymienić można niektóre decyzje niezbędne do uzyskania dobrze ustrukturyzowanych metadanych:

- określenie granic dla zestawu stwierdzeń (rekordów metadanych);
- powtarzalność elementów metadanych, ograniczenia powtarzalności;
- struktura i sposób zapisu nazw własnych, w tym głównie osobowych;
- sposób obsługi zagnieżdżania danych;
- sposób zapisu dat;
- sposób identyfikacji schematów kodowania (standardów) dla powyższych;

³⁶ Wg B. Haslhofer, W. Klas: op. cit., p. 19.

³⁷ S. Weibel: *Metadata: Semantics; structure; syntax*. [online]. [dostęp: 15.01.2013]. Dostępny w World Wide Web: <<http://weibel-lines.typepad.com/weibelines/2008/02/metadata-semant.html>>.

- sposób identyfikacji przez schemat kodowania (taki jak LCSH lub UKD) wartości metadanych, z którego pobierane mają być te wartości. Czy schemat ma być opcjonalny, czy obowiązkowy;
- identyfikacja wartości metadanych przez odnośnik (URI) lub ciąg znaków (literal).

Rys. 3. Poziomy współdziałania metadanych Dublin Core
(wg M. Nilssona, T. Bakera, P. Johnstona)

W podobny sposób współdziałanie metadanych przedstawione zostało w dokumentach DCMI (rys. 3). Wyróżniono tam cztery poziomy współdziałania metadanych Dublin Core³⁸:

1. Uzgodnione definicje terminów: 15 elementów Dublin Core stanowi zestaw (słownik) pojęć, zawierający definicje w języku naturalnym. Stanowią one podstawę ujednoczenia rozumienia struktury pomiędzy ludźmi, czyli nieformalnego współdziałania, dla którego nie ma potrzeby stosowania URI. Przyjęło się, że nieformalne współdziałanie z Dublin Core oznacza stosowanie jego 15 elementów, chociaż koncepcja nieformalnego współdziałania może funkcjonować w odniesieniu do dowolnych terminów metadanych, mających definicje w języku naturalnym.

2. Współdziałanie semantyki formalnej: opiera się na precyzyjnym i poprawnym stosowaniu formalnej semantyki RDF wyrażonej w modelu danych grafu RDF i słownikach opartych na RDF, takich jak DCMI Metadata Terms. „Semantyka” w tym sensie nie oznacza stosowania poprawnych definicji w języku naturalnym (tak jak to było tradycyjnie rozumiane w społeczności Dublin Core). Oznacza ona formalnie stwierdzone relacje pomiędzy terminami oraz zasady stosowania tych stwierdzeń w celu automatycznego uzyskiwania wniosków (wnioskowania logicznego). Bez względu na pierwotnie zastosowany format kodowania wyrażenie można nazwać „współdziałającym semantycznie”, jeśli

3. Współdziałanie syntaktyczne zestawu opisów³⁹: stanowi podstawę dla walidacji i wymiany rekordów metadanych w oparciu o DCMI Abstract Model. Metadane ustrukturyzowane zgodnie z DCMI Abstract Model mogą być uważane za współdziałające zgodnie z tym modelem. DCMI Abstract Model jest usytuowany nadrzędnie w stosunku do syntaktyki RDF, dostarczając:

- a) określenie zestawu opisów, który jako jednostka ograniczona i identyfikowalna odpowiada terminowi „rekordu” stosowanemu w bibliotekarstwie,
- b) określenie opisu, również jednostki ograniczonej i identyfikowalnej, nawet gdy włączona jest do rekordu,
- c) konwencje dla odróżnienia „reprezentacji” opisywanego źródła od jego „identyfikacji” za pomocą URI,
- d) konwencje dla „reprezentacji” jednej wartości źródła za pomocą wielu ciągów wartości.

4. Współdziałanie profili zestawu opisów: profile aplikacyjne dostarczają nieformalnych modeli oraz wyrażonych w XML ograniczeń strukturalnych na zestaw opisów. Profile aplikacyjne dla zachowania współdziałania muszą być zgodne z językiem opisu profili aplikacyjnych.

Jak widać, współdziałanie metadanych jest nie mniej skomplikowane od współdziałania systemów informacyjnych. W przedstawionych charakterystykach uwzględniane są różne cechy metadanych, w szczególności związane z praktyką ich stosowania, ale w niedostatecznym stopniu uwzględniają cechy metadanych jako języka.

Poziomy metadanych i współdziałania

Za pomocą metadanych tworzone są teksty odwzorowujące wybrane cechy formalne opisywanego źródła. Języki służące budowie (kodowaniu) tych tekstów wyrażają własności źródeł traktowanych jako nośnik informacji. Teksty te wyrażane są w języku opisu formalnego, wyspecjalizowanym w funkcji metainformacyjnej i umożliwiającym realizację funkcji wyszukiwawczej tego języka⁴⁰. Rekordy metadanych są więc komunikatami przekazywanymi za pomocą wyspecjalizowanych języków opisu formalnego, w których część elementów służy wyrażaniu własności treści dokumentu przy użyciu innych języków (opisu rzeczowego). Tak jak każdy język, język opisu formalnego określany jest przez reguły syntaktyczne, semantyczne i pragmatyczne⁴¹.

³⁹ Zwróćmy uwagę na specyficzną terminologię stosowaną w DCMI: źródło (*resource*) oznacza opisywany dokument, opis (*description*) i zestaw opisów (*description set*) są odpowiednikami rekordu (przy czym record to description set zakodowany w jakimś języku kodowania, np. HTML, XML), element i cecha (*element, property*) to atrybut opisywanego źródła – pole rekordu, natomiast wartość (*value*) jest wartością przyjmowaną przez cechę; opis składa się z par cecha-wartość.

⁴⁰ *Słownik encyklopedyczny informacji, języków i systemów informacyjno-wyszukiwawczych*. Oprac. B. Bojar. Warszawa 2002, s. 38.

⁴¹ W. Babik: *Lingwistyczne podstawy języków informacyjno-wyszukiwawczych*. W: *Informacja naukowa. Rozwój – metody – organizacja*. Pod red. Z. Żmigrodzkiego, W. Babika, D. Pietruch-Reizes. Warszawa 2006, s. 193.

Syntaktyka metadanych dotyczy głównie technicznych założeń poszczególnych schematów metadanych. Szczegółowe specyfikacje elementów, tworzone na poziomie semantycznym, tutaj wzbogacane są o opis sposobu ich kodowania oraz tworzenia powiązań między poszczególnymi elementami. Nie dotyczy ona znaczenia lub celowości schematu, ani nie jest przeznaczona dla jego użytkowników-ludzi; wykorzystywana jest raczej przez algorytmy komputerowe podczas przetwarzania danych. Poszczególne elementy muszą być jednoznacznie etykietowane i identyfikowane przy pomocy URI. Do tego celu służą języki definiowania schematu, które dostarczają podstawowych jednostek (ang. *primitives*) do opisu schematu (takich jak klasy, atrybuty, relacje)⁴², na przykład OWL i UML. Na tym poziomie wykorzystywane są takie standardy jak języki kodowania danych HTML, XML, RDFS i protokoły wymiany metadanych Z39.50 i OAI-PMH⁴³. Ponieważ komputery muszą „rozumieć” te języki, podstawowe jednostki są nie tylko konstruktami syntaktycznymi, ale posiadają własną semantykę. Język definiowania schematu pełni więc funkcję metajęzyka w stosunku do języka metadanych, służąc jego opisowi⁴⁴.

Syntaktyka metadanych bywa także odmiennie rozumiana, na przykład w cytowanym materiale IFLA⁴⁵ współdziałanie syntaktyczne rozumiane jest m.in. jako oznaczenie elementów danych jako obowiązkowe/opcjonalne i ich powtarzalność. W rozumieniu przyjętym przeze mnie w tym artykule cechy te należą do pragmatyki metadanych, gdyż w żaden sposób nie wynikają z potrzeb kodowania metadanych, natomiast wynikają z praktyki ich stosowania w systemie informacyjnym.

Wielu twórców schematów metadanych skupiało się na budowie struktury zestawu elementów, bez szczegółowego opisu sposobów ich kodowania, przynajmniej w początkowych wersjach. W ten sposób powstawał m.in. Dublin Core. Istnieją jednak takie schematy metadanych, które od początku tworzone były jako aplikacja określonego schematu kodowania, czego przykładem mogą być schematy EAD i TEI, dla których początkowo utworzono SGML DTD, a później zastosowano schematy kodowania w XML.

Semantyka dotyczy głównie definicji znaczeń i funkcjonalności elementów metadanych. W efekcie definiowania znaczeń powstaje schemat metadanych, czyli zestaw elementów o precyzyjnie zdefiniowanej semantyce. Semantyka schematu jest definiowana przez znaczenie jego elementów. Na tym poziomie określane jest znaczenie etykiet i sposoby znaczeniowego wiązania metadanych, na przykład przy pomocy ontologii. W różnych schematach metadanych mogą być stosowane różne nazwy dla oznaczenia tych samych lub podobnych pojęć; często jako przykład podawane jest stosowanie różnych elementów, o różnych nazwach, takich jak ‘Autor’, ‘Twórca’ lub ‘Kompozytor’ na ozna-

⁴² B. Haslhofer, W. Klas: op. cit., p. 7.

⁴³ H. Ding: *Challenges in building semantic interoperable digital library system*. [online]. [dostęp: 15.01.2013]. Dostępny w World Wide Web: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.198.8018>>.

⁴⁴ B. Bojar: *O metainformacji i metajęzyku*. „Zagadnienia Informacji Naukowej” 1976, nr 2(29), s. 57.

⁴⁵ IFLA: op. cit., p. 2.

czenie odpowiedzialności⁴⁶. I odwrotnie – elementy metadanych nazywane tymi samymi terminami w różnych systemach mogą oznaczać bardzo różną zawartość tych elementów, co zdarza się w bibliotekach stosujących system dLibra i modyfikowany na różne sposoby Dublin Core. Może to oczywiście powodować powstawanie błędów podczas używania systemów, zarówno podczas wyszukiwania jak i wymiany danych, a więc w sytuacjach, gdy potrzebne jest współdziałanie systemów. Rozpoznawanie znaczenia, podobieństw i różnic znaczeniowych jest domeną ludzi, bez ich udziału (ścisłego zdefiniowania znaczeń i jednoznacznego zapisu tych definicji na poziomie syntaktycznym) oprogramowanie komputerowe nie jest w stanie samodzielnie funkcjonować na tym poziomie. Semantyka metadanych dotyczy więc głównie ich struktur oraz znaczenia umieszczonych w niej elementów. Struktura wyrażana jest więc podziałem schematu metadanych na elementy, dzielone zwykle na bardziej szczegółowe części. Znaczenie wszystkich elementów struktury jest jednoznacznie definiowane. Istnieje wiele takich struktur stosowanych w różnych kontekstach społeczności (biblioteki, archiwa, muzea, serwisy społecznościowe), opisywanych obiektów (obiekty tekstowe, graficzne) lub procesów (wyszukiwanie informacji, ochrona praw, działalność komercyjna). Najbardziej oczywistymi przykładami struktur metadanych funkcjonujących w bibliotekach (tradycyjnych i cyfrowych) są struktury opisane w MARC 21 i Dublin Core.

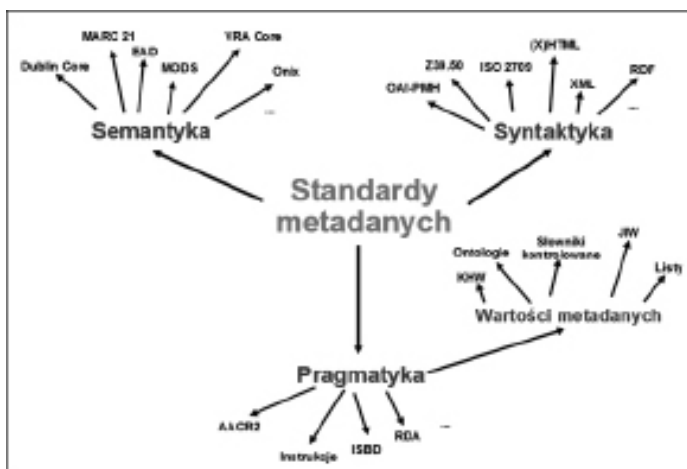
Praktyczne stosowanie metadanych o określonej semantyce wymaga skodyfikowania ich pragmatyki, czyli tzw. zasad performancji, zasad funkcjonowania metadanych w komunikacji pomiędzy ich użytkownikami. Dla każdego zdefiniowanego elementu metadanych ustalane są zasady tworzenia wartości elementu. Są to zasady wyboru treści wartości metadanych (np. zasady identyfikacji tytułu właściwego), zasady reprezentacji wartości (np. zasady dotyczące stosowania dużych liter lub podawania określenia czasu) oraz możliwe wartości elementu (np. wymóg pobierania wartości elementu z kontrolowanego słownika). Stosowanie tych reguł pozwala na tworzenie poprawnych par element (cecha)-wartość, których kombinacje tworzą opis (rekord) metadanych dla określonego obiektu informacyjnego.

Pragmatyka metadanych stanowi zbiór kontekstów pragmatycznych dla semantycznych schematów metadanych. Kontekst pragmatyczny zbudowany jest z kontekstu wspólnego oraz wielu kontekstów indywidualnych⁴⁷. Wspólny kontekst jest definiowany przez ogólnie akceptowane pojęcia i definicje konceptualne, stanowiące przedmiot zainteresowania społeczności stosującej metadane, interakcje komunikacyjne, w których pojęcia te są definiowane i używane, oraz przez zestaw parametrów wspólnego kontekstu (odpowiednie cechy pojęć, wspólne cele, sytuacje komunikacyjne). Każdy członek społeczności posiada także własny kontekst indywidualny, kształtowany przez zainteresowania oraz parametry kontekstu indywidualnego.

⁴⁶ P. Miller: *Interoperability: what is it and why should I want it?* „Ariadne” 2000, nr 24. [online]. [dostęp: 1.02.2013]. Dostępny w World Wide Web: <<http://www.ariadne.ac.uk/issue24/interoperability/>>.

⁴⁷ G. Richmond: *Interoperability as desideratum, problem, and process*. W: A. de Moore, S. Polovina, H. Delugach (ed.) *Proceedings of the 1st ICSS Conceptual Structures Tool Interoperability Workshop*. Aalborg 2006, p. 4.

Decyzja o szerszym (internetowym) rozpowszechnianiu metadanych i źródeł informacji ma wpływ na funkcjonowanie zainteresowanych organizacji (mogąc oznaczać dla nich utratę kontroli lub własności), pracowników (którzy mogą nie posiadać umiejętności obsługi nowych struktur metadanych, zarówno z punktu widzenia ich semantyki, jak i syntaktyki oraz zaspokajania nowych potrzeb użytkowników) i użytkowników końcowych. Globalne rozpowszechnianie źródeł i ich metadanych implikuje także problemy prawne, związane np. z publikowaniem danych osobowych. Problemy związane ze współdziałaniem metadanych rosną, gdy dotyczą skali międzynarodowej. Mogą one dotyczyć problemów językowych, i ogólnie, kulturowych w sytuacji, gdy praktyki, oczekiwania i wymogi różnią się w poszczególnych krajach i kulturach. Problemy te rozwiązywane są między innymi poprzez stosowanie wspólnych zbiorów zasad tworzenia wartości metadanych, niezależnych od ich syntaktyki, a ściśle związanych z przyjętymi regułami semantycznymi, takich jak ISBD, AACR i RDA oraz ich licznych modyfikacji dostosowujących rozwiązania przyjęte na poziomie międzynarodowym do kontekstu lokalnego. Wymienionym standardom towarzyszą standardy związane z tworzeniem lub wyborem z zestawu wartości metadanych. Łącznie opisują one reguły pragmatyki schematu metadanych. Współdziałanie metadanych na tym poziomie przez niektórych autorów bywa także nazywane strukturalnym⁴⁸.



Rys. 4. Trzy aspekty współdziałania metadanych⁴⁹

Na rys. 4. przedstawione są przykładowe standardy stosowane dla zapewnienia współdziałania metadanych na trzech omówionych poziomach. Lista ta ma charakter przykładowy, w każdej części może być uzupełniana o inne standardy. Zestaw ten dotyczy tzw. metadanych ustrukturyzowanych, a więc tworzonych w ramach działalności intelektualnej przez ludzi. Metadane tworzone automatycznie przez oprogramowanie wyszukiwarek funkcjonują w inny sposób.

⁴⁸ S. Weibel: op.cit.

⁴⁹ Por. M. Nahotko: *Struktury danych. W: Biblioteki cyfrowe*. Pod red. M. Janiak, M. Krakowskiej, M. Próchnickiej. Warszawa 2012, s. 364.

Wielka różnorodność rozwiązań dotyczących wszystkich poziomów metadanych może powodować pytania o przyczyny tej heterogeniczności oraz o możliwości jej usunięcia w przypadku, gdyby nie wynikała z przesłanek merytorycznych, a na przykład jedynie z historycznych zaszłości. Na pytania te odpowiedź daje zastosowanie podejścia konstruktywistycznego⁵⁰, które uznaje, że znaczenie jest konstruowane i kształtowane przy aktywnym zaangażowaniu obserwatora/badacza. Znaczenie otaczającego świata jest narzucane przez nas, nie funkcjonuje niezależnie od obserwatora. Istnieje więc wiele sposobów na strukturalizację świata i wiele znaczeń lub perspektyw dla każdego wydarzenia lub pojęcia. W takim razie istnieje wiele interpretacji każdego obiektu bibliograficznego tworzonych przez poszczególne osoby, grupy, w różnych krajach i regionach geograficznych. Tworzenie metadanych jest działalnością kulturalną w szerokim tego słowa znaczeniu, a kultura zawsze wiąże się z różnorodnością. Bardzo utrudnia to, jeśli nie uniemożliwia, tworzenie obiektywnych definicji obiektów. Stąd systemy organizacji informacji, w tym standardy metadanych, powinny posiadać zdolność do odzwierciedlenia różnych interpretacji rzeczywistości. Brak jednorodności w zakresie standardów metadanych ma więc podobne przyczyny jak brak jednego języka naturalnego.

Jeżeli skazani jesteśmy na różnorodność (heterogeniczność) standardów metadanych, musimy wypracować metody zapewnienia ich współdziałania w stopniu umożliwiającym współdziałanie systemów informacyjnych. Sposób dochodzenia do współdziałania metadanych na poziomach: semantycznym, syntaktycznym i pragmatycznym można przedstawić za pomocą uproszczonego modelu, opisanego poniżej. Załóżmy, że posiadamy trzy zestawy rekordów metadanych, różniących się między sobą we wszystkich trzech, wymienionych aspektach (tab. 2):

Tab. 2

Trzy niewspółdziałające zestawy metadanych

Baza danych 1	Baza danych 2	Baza danych 3
260\$c 245\$a 650	Data Wykonał Przedmiot	Data publikacji Twórca Deskryptor

Pomimo podobieństw struktur przedstawionych w tab. 3, dostrzegalnych szczególnie dla człowieka, najlepiej posiadającego wiedzę w zakresie stosowania struktur metadanych (katalogowania), są one na tyle różne, że nie umożliwiają zastosowania zautomatyzowanych narzędzi ich wspólnego przetwarzania. Przy pomocy tablic konwersji można ujednoczyć struktury, na przykład stosując w tym celu dodatkową strukturę, do której sprowadza się wszystkie struktury pierwotnie występujące; w tym celu często stosowana jest struktura schematu Dublin Core (tab. 3):

⁵⁰ G. Alemu, B. Stevens, P. Ross: *Towards a conceptual framework for user-driven semantic metadata interoperability in digital libraries: a social constructivist approach*. „New Library World” 2012, vol. 113, no 1/2, p. 40.

Tab. 3

Współdziałanie semantyczne

Baza danych 1	Baza danych 2	Baza danych 3
Data Twórca Opis	Data Twórca Opis	Data Twórca Opis

Pozostaje nadal do rozwiązania problem braku współdziałania na poziomie wartości metadanych, a więc formy danych (pragmatyki) zapisanych na poszczególnych poziomach struktury schematu (semantyki):

Tab. 4

Brak współdziałania pragmatycznego

Baza danych 1	Baza danych 2	Baza danych 3
Data: 2013 Twórca: Kowalski, Jan Opis: poradniki	Data: 21 maja 2013 Twórca: Jan Kowalski Opis: wskazówki	Data: 2013.05.21 Twórca: Kowalski J. Opis: przewodniki

Współdziałanie na poziomie pragmatyki metadanych uzyskano stosując słowniki zawierające dozwolone wartości metadanych (słowniki JIW, KHW), instrukcje i normy (np. dotyczące zasad tworzenia dat, stosowania skrótów itp.):

Tab. 5

Współdziałanie pragmatyczne

Baza danych 1	Baza danych 2	Baza danych 3
Data: 2013 Twórca: Kowalski, Jan Opis: przewodniki	Data: 2013.05.21 Twórca: Kowalski, Jan Opis: przewodniki	Data: 2013.05.21 Twórca: Kowalski, J. Opis: przewodniki

Zauważmy, że pełne uzgodnienie pragmatyki nie jest możliwe ze względu na braki w wartościach metadanych. W przykładowej bazie danych 1 brak pełnej daty (istnieje tylko data roczna), w bazie 3 brak imienia twórcy (istnieje tylko inicjał). W tym drugim przypadku istnieje możliwość dalszego ujednoczenia danych przy pomocy kartotek wzorcowych lub autorytatywnych. Podobnie ujednoczenie elementu opisu rzeczowego (element 'Opis') wymaga zastosowania słowników odpowiednich JIW oraz narzędzi umożliwiających ich współdziałanie.

Po ujednoczeniu metadanych na poziomie semantyki i pragmatyki pozostaje problem współdziałania na poziomie syntaktyki, czyli ujednoczenia schematu kodowania metadanych, co pozwala na ich scalenie w jedną, fizyczną bazę danych (tab. 6):

Zapewnienie współdziałania syntaktycznego

Baza danych 1	Baza danych 2	Baza danych 3
ISO 2709	(X)HTML	XML
K o n w e r s j a		



Tabele 5-9 obrazują bardzo uproszczony przypadek ujednolicania metadanych. W rzeczywistości, jak widzieliśmy wcześniej, współdziałanie nie ogranicza się do tak jednoznacznych przypadków i bywa, że jednocześnie prowadzone są działania mające zapewnić współdziałanie na kilku poziomach. Na każdym z nich niezbędne jest podejmowanie wielu szczegółowych decyzji. Tak może być na przykład w przypadku konwertowania wartości elementów opisu rzeczowego, przedstawionego przez Agnieszkę Brachfogel⁵¹ jako tzw. fasetowanie, czyli rozdzielania elementów hasła rzeczowego w jednym formacie na kilka równorzędnych wartości elementów schematu docelowego. Wówczas zawartość jednego pola 650 formatu MARC 21 może być konwertowana na kilka elementów Dublin Core, na przykład: 'Temat', 'Zasięg przestrzenny', 'Zasięg czasowy' i 'Rodzaj'. W takim przypadku konwersja następuje jednocześnie na poziomie semantycznym (struktury pól/elementów metadanych) i pragmatyki (słowniki kontrolowane stosowanych JIW).

W tab. 2. zestawione zostały metody uzyskiwania współdziałania metadanych, wyróżnione przez Marcję Zeng i Lois Chan⁵² oraz Bernharda Haslhofera i Wolfganga Klasa⁵³ we wspomnianych wcześniej publikacjach, w odniesieniu do trzech wyróżnionych przeze mnie poziomów metadanych: semantycznego, syntaktycznego i pragmatyki.

⁵¹ A. Brachfogel: *Terminy metadanych DCMI i możliwości ich wykorzystania w opisie rzeczowym*. „Zagadnienia Informacji Naukowej” 2010, nr 2(96), s. 72.

⁵² M. Zeng, L. M. Chan: op.cit.; L. M. Chan, M. Zeng: *Metadata interoperability and standardization – a study of methodology part I. Achieving interoperability at the schema level*. „D-Lib Magazine” 2006, vol. 12, no 6. [online]. [dostęp: 15.01.2013]. Dostępny w World Wide Web: <<http://www.dlib.org/dlib/june06/chan/06chan.html>>.

⁵³ B. Haslhofer, W. Klas: op. cit.

Poziomy metadanych i poziomy współdziałania (w nawiasach przykłady zastosowań)

		Poziomy współdziałania		
		Schematy	Rekordy	Repozytoria (bazy danych)
Poziomy metadanych	Semantyka	Zestawy pochodne (MODS); Profile aplikacyjne; Mapowanie definicji elementów; Tablice przejścia; Struktury nadrzędne (CDWA, OAIS); Rejestry metadanych	Konwersja; Agregacja (METS, RDF)	Agregacja rekordów; Tablice przejścia
	Syntaktyka	Zestawy pochodne (MARCXML)	Konwersja; Mapowanie języków kodowania	Wiele formatów (DLESE); Zbieranie (OAI-PMH); Agregacja rekordów; Tablice przejścia
	Pragmatyka	Zestawy pochodne (DC w różnych językach); Mapowanie zasad pragmatyki	Konwersja; Schematy kodowania; Słowniki kontrolowane, w tym KHW	Mapowanie wartości metadanych opisu rzeczowego (MACS)

Jak wynika z tego zestawienia, na poziomie schematu metadanych najważniejszą rolę odgrywa współdziałanie semantyczne, na poziomie rekordu – pragmatyczne, natomiast na poziomie baz danych – syntaktyczne. Najwięcej metod zapewniających współdziałanie metadanych funkcjonuje na poziomie semantyki i pragmatyki. Zapewne wynika to z różnorodności standardów stosowanych na obu tych poziomach. Rozwiązania na tych poziomach współdziałania w znacznym stopniu zależą od rozwiązań przyjętych lokalnie przez twórców schematów metadanych i ich pragmatyki, możliwości tworzenia indywidualnych rozwiązań są więc znaczne. Inaczej wygląda to na poziomie syntaktyki, gdzie istnieje do wyboru kilka standardów globalnych, a tworzenie własnych, lokalnych rozwiązań należy uznać za błąd. W tej sytuacji na poziomie syntaktycznym stosowanych jest znacznie mniej technik zapewnienia współdziałania.

Część narzędzi służących współdziałaniu stosowanych jest na wszystkich poziomach metadanych, na przykład zestawy pochodne (ang. *derivation*) stosowane na poziomie schematów metadanych, chociaż na każdym z nich oznaczają inną funkcjonalność. Format MODS (Metadata Object Description Schema) jest schematem pochodnym MARC 21, w którym m. in. numeryczne etykiety pól zastąpiono etykietami słownymi. Format MARCXML posługuje się kodowaniem w języku XML pól z formatu MARC 21 wraz z ich numerycznymi etykietami. Na poziomie pragmatycznym natomiast tworzone są zestawy pochodne Dublin Core w różnych językach (tłumaczone są etykiety elementów, które oryginalnie powstały w jęz. angielskim), gdyż samo tłumaczenie nazwy elementu nie zmienia jego semantyki ani zasad kodowania.

Z drugiej strony istnieją takie metody zapewnienia współdziałania metadanych jak np. tworzenie profili aplikacyjnych, które zapewniają współdziałanie

tylko na jednym poziomie (w tym przypadku semantyki). W wyniku utworzenia profilu aplikacyjnego powstaje nowy schemat metadanych, a więc nowy zbiór rozwiązań semantycznych, nie powoduje to jednak potrzeby modyfikacji sposobu kodowania (syntaktyka) ani zasad tworzenia wartości poszczególnych elementów metadanych (pragmatyka).

Wnioski

W globalnym, ale heterogenicznym środowisku współczesnego Internetu prace służące zapewnieniu współdziałania elementów tego środowiska są niezbędne, chociaż należy zdawać sobie sprawę z niemożności osiągnięcia pełnego, 100-procentowego sukcesu. Bardzo ważną rolę odgrywa standary-

3. Bojar B.: *O metainformacji i metajęzyku*. „Zagadnienia Informacji Naukowej” 1976, nr 2 (29), s. 43-60.
4. Brachfogel A.: *Terminy metadanych DCMI i możliwości ich wykorzystania w opisie rzeczowym*. „Zagadnienia Informacji Naukowej” 2010, nr 2(96), s. 69-76.
5. Candela L. [et al.]: *The DELOS digital library reference model. Foundations for digital libraries*. Ver. 0.98 [online]. [dostęp: 15.01.2013]. Dostępny w World Wide Web: <<http://www.delos.info/ReferenceModel>>.
6. Candela L. [et al.]: *Setting the foundations of digital libraries. The DELOS Manifesto*. „D-Lib Magazine” 2007, vol. 13, no 3/4. [online]. [dostęp: 15.01.2013]. Dostępny w World Wide Web: <<http://www.dlib.org/dlib/march07/castelli/03castelli.html>>.
7. Carney D., Smith J., Place P.: *Topics in interoperability: infrastructure replacement in a system of systems*. Report nr CMU/SEI2005-TN-031. Pittsburgh 2005.
8. CC:DA (ALCTS/CCS/Committee on Cataloging: Description and Access): *Task Force on Metadata. Final report*. June 16, 2000. [online]. [dostęp: 15.01.2013]. Dostępny w World Wide Web: <<http://www.libraries.psu.edu/tas/jca/ccda/tf-meta6.html>>.
9. Chan L. M., Zeng M.: *Metadata interoperability and standardization – a study of methodology. Part 1. Achieving interoperability at the schema level*. „D-Lib Magazine” 2006, vol. 12, no 6. [online]. [dostęp: 15.01.2013]. Dostępny w World Wide Web: <<http://www.dlib.org/dlib/june06/chan/06chan.html>>.
10. Ding H.: *Challenges in building semantic interoperable digital library system*. [online]. [dostęp: 15.01.2013]. Dostępny w World Wide Web: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.198.8018>>.
11. European Commission: *European interoperability framework for pan-European e-government services*. Luxembourg 2004.
12. Foulonneau M., Riley J.: *Metadata for digital resources: implementation, systems design and interoperability*. Oxford 2008.
13. Gradmann S.: *Interoperability. A key concept for large scale, persistent digital libraries*. [online]. [dostęp: 10.01.2013]. Dostępny w World Wide Web: <<http://www.digitalpreservationeurope.eu/publications/briefs/interoperability.pdf>>.
14. Haslhofer B., Klas W.: *A survey of techniques for achieving metadata interoperability*. „ACM Computing Survey” 2010, vol. 42, no 2, pp. 1-37.
15. IFLA: *Sharing of bibliographic information and resources*. [online]. [dostęp: 15.01.2013]. Dostępny w World Wide Web: <<http://archive.ifla.org/VII/d4/pub/InteroperabilityStandards.pdf>>.
16. Klenczon W.: *Międzynarodowe standardy opisu obiektów muzealnych jako podstawa katalogowania zbiorów*. Warszawa 2011, s. 19-54.
17. Manso M., Wachowicz M., Bernabé M.: *Towards an integrated model of interoperability for spatial data infrastructures*. „Transactions in GIS” 2009, vol. 13, no 1, pp. 43-67.
18. Manso-Callejo M., Wachowicz M., Bernabé-Poveda M.: *Automatic metadata creation for supporting interoperability levels of spatial data infrastructure*. [online]. [dostęp: 15.01.2013]. Dostępny w World Wide Web: <<http://www.gsdi.org/gsdiconf/gsdi11/papers/pdf/194.pdf>>.
19. Mielnicki M. [et al.]: *Agregacja metadanych w skali kraju – kierunki rozwoju Federacji Bibliotek Cyfrowych*. Warszawa 2012, s. 69-83.
20. Miller P.: *Interoperability: what is it and why should I want it?* „Ariadne” 2000, no 24. [online]. [dostęp: 1.02.2013]. Dostępny w World Wide Web: <<http://www.ariadne.ac.uk/issue24/interoperability/>>.
21. Miller S.: *Metadata for digital collections*. New York, London 2011.
22. Nahotko M.: *Struktury danych. W: Biblioteki cyfrowe*. Pod red. M. Janiak, M. Krakowskiej, M. Próchnickiej Warszawa 2012, s. 362-371.
23. Nilsson M.: *From interoperability to harmonization in metadata standardization*. Stockholm 2010.
24. Nilsson M., Baker T., Johnston P.: *Interoperability levels for Dublin Core metadata*. [online]. [dostęp: 15.01.2013]. Dostępny w World Wide Web: <<http://dublincore.org/documents/2009/05/01/interoperability-levels/>>.
25. NISO: *Understanding metadata*. Bethesda 2004.

26. Pacek J.: *Bibliografia w zmieniającym się środowisku informacyjnym*. Warszawa 2010.
27. Potęga J., Wróbel A.: *The Dublin Core Metadata Element Set, Ver. 1.1 a potrzeby i oczekiwania bibliotekarzy cyfrowych – analiza przypadków*. W: *Polskie biblioteki cyfrowe 2009. Materiały z konferencji*. Pod red. C. Mazurek, M. Stroiński, J. Węglarz. Poznań 2010, s. 71-78.
28. Richmond G.: *Interoperability as desideratum, problem, and process*. W: *Proceedings of the 1st ICSS Conceptual Structures Tool Interoperability Workshop*. Ed. A. de Moore, S. Polovina, H. Delugach. Aalborg 2006, pp. 1-17.
29. *Słownik encyklopedyczny informacji, języków i systemów informacyjno-wyszukiwawczych*. Oprac. B. Bojar: Warszawa 2002.
30. Subieta K.: *Słownik terminów z zakresu obiektowości*. Warszawa 1999.
31. Taylor A., Joudrey D.: *The Organization of information*. Westport 2008.
32. Tolkoff I.: *The path toward global interoperability in cataloging*. „Information Technology and Libraries” 2010, vol. 29, no 1, pp. 30-33, p. 39.
33. Veltman K.: *Syntactic and semantic interoperability: new approaches to knowledge and the semantic web*. „New Review of Information Networking” 2001, vol. 7, no 1, pp. 159-183.
34. Weibel S.: *Metadata: Semantics; structure; syntax*. [online]. [dostęp: 15.01.2013]. Dostępny w World Wide Web: <<http://weibel-lines.typepad.com/weibelines/2008/02/metadata-semantic.html>>.
35. Zeng M., Chan L. M.: *Metadata interoperability and standardization – a study of methodology. Part 2. Achieving interoperability at the record and repository levels*. „D-Lib Magazine” 2006, vol. 12, nr 6. [online]. [dostęp: 15.01.2013]. Dostępny w World Wide Web: <<http://www.dlib.org/dlib/june06/zeng/06zeng.html>>.
36. Zeng M., Qin J.: *Metadata*. New York 2008.

Summary

Interoperability is one of the most important features of the modern information systems, concerning all aspects of their functioning, and particularly important in global networks. The article presents problems of metadata interoperability as an element of information systems' interoperability, in particular digital libraries. Both types of interoperability are defined. The most important concepts of placing metadata interoperability among different types of information systems' interoperability are discussed. The author lists the most popular methods of metadata interoperability on: schema, record, and repository levels. Finally, he develops a model of metadata interoperability referring to its' semantics, syntactic, and pragmatics, and offers a set of tools used on the mentioned levels.