

Marek Nahotko

Instytut Studiów Informacyjnych
Uniwersytet Jagielloński
e-mail: marek.nahotko@uj.edu.pl

[Czy (centralne) katalogi biblioteczne są jeszcze potrzebne? OPAC w infotopii

DOI: <http://dx.doi.org/10.18778/0860-7435.29.03>

Abstrakt: W artykule przedstawiono podstawową funkcję katalogów centralnych: zapewnienie współdziałania (ang. *interoperability*) metadanych bibliotecznych. Zdefiniowano problem współdziałania metadanych jako element współdziałania systemów informacyjnych, szczególnie funkcjonujących w sieci. Przedstawiono sposoby stosowane dla zapewnienia współdziałania. Opisano trzy technologie, które w przyszłości powinny ułatwić osiągnięcie współdziałania metadanych: bazy danych NoSQL, aplikacje chmurowe oraz Linked Open Data.

Słowa kluczowe: współdziałanie metadanych, katalogi centralne, OPAC, big data, Linked Open Data

[Wprowadzenie

W bibliotekach od wielu lat główne zasoby metadanych funkcjonują w wersji elektronicznej, co pozwala na ich masowe udostępnianie w Internecie. Katalogi kartkowe już dawno zostały zamknięte, albo po pełnej rekatalogizacji zbiorów, albo po przygotowaniu różnego rodzaju „protezy”, takich jak na przykład zestaw skanów kart katalogowych z możliwością wyszukiwania obrazków według hasła. Doświadczenie uczy, że dokumenty, których opisy nie zostały umieszczone w OPAC, stają się „martwe”, gdyż nie są zamawiane przez użytkowników (Su, 1994, s. 136). Bibliotekarze zauważyli potrzebę

współpracy w zakresie tworzenia i udostępniania metadanych na długo przed zastosowaniem narzędzi elektronicznych. Początkiem tej tendencji były prace prowadzone w drugiej połowie XIX w. przez Panizziego, Cuttera, Dziatzko i innych bibliotekarzy. Komputeryzacja bibliotek, a w szczególności przeniesienie dużej części działalności informacyjnej do Internetu, znacznie pogłębiły potrzebę współpracy do tego stopnia, że biblioteki funkcjonujące indywidualnie, niewspółpracujące, tracą sens istnienia (dotyczy to w szczególności bibliotek naukowych).

Współpraca ta ma dwa uzupełniające się wymiary. Najbardziej oczywista jest współpraca pomiędzy skomputeryzowanymi bibliotekami. Najszybciej i najlepiej zrealizowanym pomysłem jest wymiana metadanych bibliograficznych (opisowych) pomiędzy bibliotekami w różnych momentach tworzenia opisów i na różnych zasadach. Gorzej jest z metadanymi innego rodzaju, np. administracyjnymi, gdyż w bibliotekach nie ma tradycji wymieniania się nimi, a mogłyby one stać się podstawą interesujących, nowych funkcjonalności współpracujących systemów informacyjnych. Dotyczy to na przykład danych o zamawianych nowościach, czytelnikach lub statystykach wykorzystania zasobów informacyjnych.

Drugim wymiarem jest potrzeba współpracy bibliotek nie tylko między sobą, ale także z innymi instytucjami globalnego uniwersum informacyjnego, tworzonoego przez Internet. Chodzi tu o współpracę ze wszystkimi instytucjami i serwisami, które obsługują informacyjnie naszych, rzeczywistych i potencjalnych, użytkowników. W swoim czasie w środowisku bibliotekarzy żywe były dyskusje o zagrożeniach, które może przynieść funkcjonowanie serwisu Google Books. Przeważały głosy krytyczne w stosunku do tego przedsięwzięcia. Tymczasem bibliotekarze powinni traktować tego typu inicjatywy jak sprzymierzeńców i szansę na współpracę, w odróżnieniu od wydawców, dla których mogą one być konkurencją. Google Books jest oczywiście tylko przykładem; nasi użytkownicy korzystają także z Facebooka, YouTube, Flickr, Spotify (obiekty cyfrowe zawierają nie tylko tekst!) i serwisów typu LibraryThing, korzyści przynosi także współpraca z wydawcami i twórcami baz danych, dostawcami usług analizy i wizualizacji danych i wieloma innymi.

Współdziałanie (interoperacyjność)

Współdziałanie, zwane niezbyt fortunnie interoperacyjnością lub interoperatywnością, z ang. *interoperability* (Kienzler, 2003, s. 126), jest bardzo istotnym zagadnieniem związanym z funkcjonowaniem systemów informacyjnych w heterogenicznym środowisku Internetu (Nahotko, 2013, s. 63). Dotyczy ono możliwości współpracy różnych systemów (oraz szerzej – organizacji użytkujących te systemy) na wszystkich płaszczynach. Dążenie do współdziałania wymaga łączenia we wspólnym działaniu różnych standar-

dów, dzięki czemu możliwa jest agregacja wielu zasobów danych i metadanych, pozwalająca na tworzenie nowych, lepszych produktów i usług. Może ono być rozumiane bardzo różnie w zależności od poziomu, na którym rozpatrywana jest współpraca.

Najważniejszą cechą współdziałania jest istnienie jakiegoś rodzaju relacji pomiędzy systemami informacyjnymi, która jest traktowana jako forma komunikacji, wymiany i rozpowszechniania informacji, umożliwiających współpracę ludzi: twórców i użytkowników tych systemów. Rezultatem jest system współdziałających systemów informacyjnych (Carney, Smith & Place, 2005, s. 2). Zazwyczaj różnice w definicjach współdziałania wynikają z odmiennych opisów tych relacji i elementów systemów. Autorzy części definicji skupiają się na problemach systemowych i sprzętowych, inni na usługach dostarczania i rozpowszechniania informacji, a jeszcze inni na możliwościach wykorzystania wymienianej informacji bez konieczności wykonywania dodatkowych prac. Organizacje zarządzające polityką informacyjną zwracają uwagę na aspekty procesów niezbędnych dla wymiany i wielokrotnego wykorzystania informacji (Manso-Callejo, Wachowicz & Bernabé-Poveda, 2009).

Współdziałanie systemów informacyjnych może być rozumiane w różny sposób. Ma to swoje odbicie w sformułowanych definicjach, których autorzy współdziałania rozumieją jako:

- dziedzinę wiedzy i praktyki zajmującą się współpracą niezależnie zbudowanych (heterogenicznych) systemów, szczególnie w sieciach komputerowych (Subieta, 1999, s. 201);
- możliwość wymiany danych z minimalną utratą treści i funkcjonalności pomiędzy wieloma systemami, które różnią się stosowanymi: sprzętem i oprogramowaniem, strukturami danych i interfejsami (NISO, 2004, s. 15);
- zdolność dwóch lub więcej systemów lub ich części do wymiany i wykorzystania wymienionej informacji bez potrzeby wykonywania dodatkowych prac w tych systemach (Nilsson, 2010, s. 12);
- kompatybilność dwóch lub więcej systemów pozwalająca wymieniać między nimi informację i dane oraz stosować je bez wykonywania dodatkowych manipulacji (Taylor & Jourdey, 2008, s. 269).

Z definicji tych wynika, że współdziałanie jest cechą wszystkich systemów komputerowych, w tym informacyjnych, gdyż każdy z nich składa się z przynajmniej dwóch elementów, sprzętu i oprogramowania, które muszą ściśle ze sobą współdziałać. Problem ten nie dotyczy jednak wyłącznie zastosowań techniki komputerowej. Przedkomputerowe systemy informacyjne, na przykład systemy biblioteczne, również często współdziałały ze sobą, co ułatwiały wspólnie stosowane standardy, takie jak przepisy katalogowania. Systemy nie służące głównie przetwarzaniu informacji, takie jak system ener-

getyczny lub transportowy (np. kolejowy), także muszą składać się ze współdziałających elementów. Jednym z elementów systemów informacyjnych, istotnym dla zapewnienia współdziałania, są metadane. W związku z tym znajdują one istotne miejsce w każdym modelu reprezentującym zagadnienia współdziałania systemów przetwarzania informacji (Nahotko, 2013, s. 62).

Współdziałanie rozumiane tak, jak w przedstawionych definicjach, oznacza takie „używanie” lub „wykorzystywanie” informacji bądź danych (w tym metadanych), w systemach importujących informację/dane, które zgodne jest z celami systemu, w którym informacja/dane powstały. W przypadku metadanych, będących reprezentacją cech formalnych i treściowych obiektu bibliograficznego oznacza to, że ich interpretacje powinny być spójne. Rozumiemy przez to, że metadane, tworzone przez indeksującego (człowieka) w jednym systemie informacyjnym i następnie eksportowane do innego systemu, będą przetwarzane w systemie importującym w sposób zbliżony z celami i intencjami twórcy metadanych.

W definicji Arlene Taylor i Daniela Joudrey, przedstawionej powyżej, porównano współdziałanie z kompatybilnością systemów. W literaturze zdarza się także odróżnianie współdziałania od kompatybilności systemów (Pacek, 2010, s. 84). Kompatybilność w informatyce rozumiana jest jako zdolność sprzętu i oprogramowania komputerowego do funkcjonowania jako spójna całość. W takim sensie kompatybilność jest uzyskiwana dzięki współdziałaniu wszystkich elementów systemów, współdziałanie jest więc sposobem na osiągnięcie kompatybilności. Innym zjawiskiem związanym ze współdziałaniem jest harmonizacja metadanych (Nilsson, 2010, s. 14). Proces ten pozwala na współdziałanie systemów informacyjnych w sytuacji stosowania w nich wielu schematów metadanych. W takim przypadku w każdym współdziałającym systemie informatycznym przetwarzane są metadane tworzone w różnych systemach i standardach. Przetworzone metadane zapewniają poprawne działanie systemów w sensie osiągania ich planowanej funkcjonalności. Taka harmonizacja może być traktowana jako rodzaj współdziałania na poziomie metadanych, o którym mowa będzie w dalszej części artykułu.

Zastosowanie definicji współdziałania systemów informacyjnych, zaprezentowanych 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 obiektach (bibliograficznych i innych) oraz do interpretowania tych danych opisowych, które podlegały wymianie w sposób spójny z interpretacją twórcy danych (Nilsson, 2010, s. 13).

Najważniejszym elementem przedstawionej definicji jest wymiana informacji lub danych, przy czym stosowanie wymienianych danych powinno być

zgodne z intencjami twórców systemu, w którym dane powstały. W przypadku metadanych istnieje wymóg spójności interpretacji danych jako opisu obiektów różnego rodzaju. Oznacza to, że metadane utworzone przez bibliotekarza w jednym systemie, a następnie przeniesione do innego, będą przetwarzane w tym drugim systemie w sposób spójny z intencjami twórcy metadanych.

Zdolność metadanych do współdziałania traktowana jest jako jedna z najważniejszych cech umożliwiających ich stosowanie, obok prostoty, modularności, wielokrotnego użycia (ang. *reusability*) i rozszerzalności (Zeng & Qin, 2008, s. 268). Cechy te powinny stanowić podstawowe wyznaczniki projektowania i wdrażania systemów i projektów metadanych. W publikacji IFLA (2005) 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.

Współdziałanie metadanych może być zapewnione w bardzo różny sposób, na wszystkich etapach ich tworzenia i stosowania oraz z użyciem wielu narzędzi. W tabeli 1. przedstawione zostały narzędzia służące uzyskiwaniu współdziałania metadanych przed utworzeniem (Chan & Zeng, 2006) i po utworzeniu rekordu metadanych (rekordu katalogowego) (Zeng & Chan, 2006). Zaznaczono narzędzia Linked Data, które zostaną omówione w dalszej części artykułu. Jak widać na rysunku, w obu etapach stosowania metadanych możliwe jest wykorzystanie odpowiednich metod uzyskiwania współdziałania. Metody te funkcjonują także od wielu lat w polskim bibliotekarstwie. Dla zapewnienia współdziałania jeszcze przed utworzeniem rekordu odbywa się współkatalogowanie (w ramach Nukat), co oznacza stosowanie wspólnych, zawczasu przygotowanych standardów, zasad i przepisów katalogowania, wykorzystywanych przez wszystkie biblioteki współpracujące. Nie bez znaczenia jest też wspólny system kartotek wzorcowych (KHW), w tym wspólny system opracowania rzeczowego.

W przypadku, gdy potrzeba współdziałania metadanych pojawia się już po ich utworzeniu, stosowane są inne sposoby jej zapewnienia, takie jak konwersja z jednego formatu do innego lub stosowanie standardowych protokołów, takich jak Z39.50 oraz OAI-PMH (ang. *Open Archives Initiative Protocol for Metadata Harvesting*). Ta ostatnia metoda także jest stosowana w polskich systemach informacyjnych; Z39.50 jest podstawą funkcjonowania katalogu KARO, natomiast OAI-PMH prawdopodobnie stosowany będzie w systemie OMNIS.

Wspomaganie współdziałania zarówno przed, jak i po utworzeniu metadanych, ma swoje zalety w określonych sytuacjach środowiska informacyjnego, nie można więc wyróżnić jednej metody jako najlepszej. Żaden sposób nie zapewnia też stuprocentowego współdziałania. Zwykle mamy do czynienia z mniejszymi lub większymi stratami w semantyce współdziałających metadanych.

Tab. 1. Narzędzia do uzyskiwania współdziałania metadanych

Współdziałanie metadanych						
Przed utworzeniem rekordu			Po utworzeniu rekordu			
Ujednolicenie	Współdziałanie schematów		Współdziałanie rekordów		Współdziałanie repozytoriów	
Wspólne standardy: Np. MARC, Dublin Core, MODS	Pochodne zestawy: MODS oparty na MARC, Rozszerzenie DC	Profile aplikacyjne: DC-Library AP, DC-Government AP	Konwersja: Np. z MARC do DC, z DC do MODS. Podczas konwersji z bogatszego schematu do uboższego następuje utrata danych.	Linked Data: Pozwala na wielokrotne stosowanie danych i zdecentralizowaną integrację. Np. METS, RDF	Protokoły wymiany: Np. Z39.50, OAI-PMH	Agregacja: Łączenie wielu zasobów dla wzbogacenia metadanych.
	Tablice przejścia: Mapowanie elementów jednego schematu na elementy innego schematu.	Schematy przełącznikowe: Mapowanie pomiędzy wieloma schematami jednocześnie.				
	Struktura nadrzędna: Np. ADN, OAIS	Rejestry metadanych: Np. UKOLN, OMR, LOV				
Niepraktyczne w środowisku heterogenicznym	Konwersja stratna sprowadzająca metadane do najniższego wspólnego mianownika: utrata semantyki danych zamiast jej wzbogacenia			Modułowe rozwiązania integrujące różne standardy w jedną strukturę: metadane w różnych standardach dla różnych zasobów mogą być ze sobą łączone		

Na podstawie: <http://commons.wikimedia.org/wiki/File:Metadata-interoperability.png>

Nowe metody zapewnienia współdziałania metadanych

Jak pisze Karen Coyle, biblioteki, wraz z innymi podobnymi organizacjami, zawsze były instytucjami innowacyjnymi w zakresie organizacji informacji. Jednak w XX i na początku XXI w. technologie biblioteczne stanowią wtórną implementację technologii wcześniej wykorzystywanych w innych miejscach, służących niebibliotecznej organizacji informacji. Od około lat 80. XX w. katalogi biblioteczne wyszły poza mury tych instytucji: jako OPAC udostępniane były początkowo na terminalach w sieciach lokalnych, a później w sieciach globalnych. Od tego czasu OPAC zyskał wiele nowych cech, takich jak wyszukiwanie zintegrowane (katalogi centralne), charakterystyki związane z Web 2.0 (OPAC 2.0) oraz uwzględnianie metadanych gatunków informacji wcześniej nieobecnych w katalogach, np. artykułów z czasopism. Pomimo tych zmian stosowana technologia informacyjna pozostawała bez zmian, systemy biblioteczne oparte były na stosowaniu relacyjnych baz danych o strukturach zaprojektowanych w latach 60. XX wieku (Coyle, 2017).

Od czasu pojawienia się pierwszych OPAC technologie informacyjne uległy olbrzymiemu rozwojowi. Nastąpiły tam zmiany, które tylko powierzchownie były uwzględniane w bibliecznych systemach informacyjnych i w ich OPAC. Po pojawieniu się XML (ang. *Extensible Markup Language*) przygotowano wersję formatu MARC w tym języku, jednak była to jedynie zmiana w syntaktyce metadanych, sam format (semantyka) oraz model danych pozostał bez zmian. Spowodowało to nawet pojawienie się twierdzeń o końcu („śmierci”) formatu MARC (Tennant, 2002, s. 26), które były jedną z przyczyn rozwoju prac nad nowymi formatami, takimi jak Dublin Core. Twierdzenia te były przedwczesne, nie uwzględniały bowiem tego, że metadane są największym bogactwem bibliotecznego systemu informacyjnego, jego najbardziej cenną częścią. Nie da się więc z dnia na dzień porzucić milionów rekordów MARC utworzonych na świecie. Technologie informacyjne się zmieniają, a metadane pozostają. Nie zmienia to jednak faktu, że one także powinny być dostosowywane do możliwości nowych technologii, stąd problem współdziałania jest tak istotny.

Według Heather Moulaison Sandy środowisko pracy katalogujących bibliotekarzy od roku 2000 uległo rozszerzeniu i skomplikowaniu przez wzrost liczby standardów służących pracy z obiektami elektronicznymi, przy czym same standardy coraz częściej dostępne są w wersji elektronicznej (2019, s. 384). Wskazać można całe zestawy standardów i narzędzi, służących wyspecjalizowanym działaniom, które musi znać i stosować współczesny bibliotekarz. Wiele z nich utworzonych i zastosowanych zostało w skali międzynarodowej. Technologie baz danych NoSQL, chmury obliczeniowej i Linked Open Data, opisane pokrótce dalej, stanowią przykłady standardów powstałych poza bibliotekami, które wdzierają się do zastosowań bibliecznych. Każdy z nich

wymaga stosowania wielu nieznanych wcześniej bibliotekarzom narzędzi (np. edytorów, narzędzi programistycznych) i standardów, co wymaga zdobywania nowych umiejętności.

Jednym z elementów nowego środowiska informatycznego są nierelacyjne bazy danych tzw. „NoSQL”. Skrót ten stosowany jest dla wskazania na odmiennosć nowych technologii w stosunku do języka zapytań SQL (ang. *Structured Query Language*), traktowanego jako symbol szczytowych osiągnięć w zakresie relacyjnych baz danych. Tłumaczony bywa także jako „not only SQL”, co z kolei wskazuje na symbiozę relacyjnych i innych baz danych w obszarze wyszukiwania informacji. Bazy danych NoSQL kojarzone są ze zjawiskiem tzw. big data, chociaż zbiory danych organizowane z zastosowaniem relacji typowych dla SQL też tam funkcjonują. Możemy także mówić o „big metadata” – metadane to też dane, które jednak różnią się wieloma cechami od metadanych gromadzonych w bibliotekach. Metadane big data gromadzone bywają planowo, tak jak w typowej bibliotece, mogą jednak być także gromadzone pasywnie, „na wszelki wypadek”. Często alternatywne możliwości ich późniejszego użycia nie są wówczas jeszcze znane, lecz niski koszt pozyskania i gromadzenia danych daje ekonomiczne uzasadnienie takiego postępowania; rozsądne jest założenie, że w przyszłości dane znajdą jakieś (obecnie nieznane) zastosowanie. Zjawisko zbierania danych o wszystkim „na wszelki wypadek” określane jest niekiedy jako „danetyzacja” (ang. *datafication*). Do niedawna przydatność danych kończyła się z chwilą osiągnięcia celu, w którym były zbierane. Obecnie zgromadzone dane mogą być nadal surowcem dla biznesu (w tym produkcji), administracji publicznej i nauki.

Big data to bardzo duża liczba danych, których obróbka wymaga zastosowania nowych technologii i architektur, tak by była możliwa ekstrakcja wartości płynących z tych danych poprzez procesy gromadzenia danych i ich analizy (Katal, Wazid & Goudar, 2013, s. 405). Ze względu na masowość danych niezbędne jest stosowanie nowych technik analitycznych. Big Data to określenie stosowane dla takich zbiorów danych, które jednocześnie charakteryzują się dużą objętością, różnorodnością, strumieniowym napływem w czasie rzeczywistym, zmiennością, złożonością, jak również wymagają zastosowania innowacyjnych technologii, narzędzi i metod informatycznych w celu wydobycia z nich nowej i użytecznej wiedzy. Analiza zbiorów big data różni się od zwykłej analizy statystycznej, gdyż następuje odejście od analizy przyczynowo-skutkowej na korzyść badania korelacji pomiędzy danymi. Analizy dokonywane są na całym zbiorze danych, zamiast precyzyjnego wyznaczania i analizy reprezentatywnej próby. Analizy te służą głównie do poszukiwania w czasie rzeczywistym wzorców danych w słabo uporządkowanym (ustrukturyzowanym) zbiorze, zwanym często „jeziorem danych” (w odróżnieniu od ściśle ustrukturyzowanych relacyjnych baz danych SQL).

Biblioteki są jednym z tych miejsc, w których znajdujemy dane masowe, miliony rekordów dotyczących obiektów informacyjnych, ich użytkowników oraz transakcji, których dokonują. Metadane biblioteczne od typowych danych w big data różnią się swoim ścisłym ustrukturyzowaniem. Kolejną różnica polega na tym, że metody big data dotyczą korelacji w zbiorze danych, a nie pojedynczych obiektów, takich jak książki, zdania i słowa. Korelacje obejmują wyszukane obiekty w ich otoczeniu kontekstualnym, wraz z relacjami, w które wchodzi w tym otoczeniu. Dzięki relacjom zapisanym w OPAC i KHW można twierdzić, że OPAC zawiera fenomenologiczny opis otoczenia opisywanego obiektu utworzony w celu dotarcia do reprezentacji wnętrza (treści) tego obiektu (Krajewski, 2017, s. 227).

OPAC może służyć jako narzędzie big data na dwa sposoby. Po pierwsze, przeciwdziała nadmiarowi informacji dzięki funkcjom strukturyzacji treści oraz lokalizowania i wyszukiwania informacji. Po drugie, zawiera on nową informację, tworzoną przez bibliotekarzy specjalnie w celu organizowania różnych obiektów przy pomocy wspólnych elementów ich treści pobieranych z notatek, abstraktów i innych paratektów. Metadane, traktowane jako big data, służące określeniu korelacji pomiędzy obiektami informacyjnymi, statystykami ich udostępniania, cytowania, (współ)występowania słów, pozwalają na łatwe określenie częstotliwości wyszukiwania każdego obiektu, wskazywania tych części tekstu, które są najczęściej czytane (lub zaznaczane, jak w Amazon Kindle), wyznaczania podobieństwa tekstów. Ta informacja z kolei może prowadzić do stosowania OPAC w nowej roli – konsultanta proponującego nową lekturę na podstawie algorytmu korelacji pomiędzy zachowaniami na kolejnych stronach (ekranach) tekstu, ruchów myszą, czasu spędzonego na czytaniu. To znowu pozwala na tworzenie korelacji między tymi obiektami, które zostały wykorzystane (niekoniecznie przeczytane) a całym zbiorem big data, którym dysponuje OPAC, a właściwie globalna sieć współdziałających OPACów: światowy katalog centralny.

Kolejną technologią informacyjną, stanowiącą przyszłość systemów informacyjnych, są aplikacje chmurowe. Rozwój sieci komputerowych, wzrost ich szybkości, przepustowości i niezawodności, stał się przesłanką do tego, by lokalne – rozproszone zbiory danych mogły lepiej ze sobą współdziałać. Spowodował on, że rozsądne stało się przekazanie lokalnych serwerów bibliotecznych wraz z oprogramowaniem (w tym z OPAC) wyspecjalizowanym firmom, które dbają o ich bezawaryjne funkcjonowanie, oczywiście za odpowiednią odpłatnością. Bibliotekarze, podobnie jak inni użytkownicy, dostęp do systemu bibliotecznego w chmurze mają poprzez Internet. Jak pisze Aleksander Radwański (2015, s. 33), umieszczenie systemów informacyjnych w chmurze nie tylko pozwala na komunikowanie się z nimi najszerzej rozumianych użytkowników, ale także samych systemów (ich algorytmów) między sobą, co stwarza nową jakość w funkcjonowaniu tych systemów.

Powszechne zastosowanie tej technologii pozwala na uzyskanie przez OPAC biblioteczny nowych funkcjonalności. Użytkownik powinien mieć możliwość otrzymania potrzebnego mu materiału znając dowolne jego cechy (metadane), w tym dysponując także obrazem okładki, np. sfotografowanej. Po odnalezieniu potrzebnego obiektu informacyjnego powinien być kierowany do najbliższego miejsca, w którym obiekt ten jest dostępny (wersja drukowana) lub bezpośrednio do wersji elektronicznej (jeśli istnieje). Również otwarcie konta czytelniczego (ważnego we wszystkich bibliotekach w chmurze) nie powinno być trudniejsze niż w jakiegokolwiek aplikacji internetowej, do której dostęp uzyskać można poprzez konto w Google lub Facebooku. Dzięki tym funkcjom użytkownik w możliwie najszybszy sposób ma możliwość dotarcia do interesującego go obiektu w najbardziej mu przydatnej, dostępnej formie. Warto także zaznaczyć, że tak działająca chmura bibliotek byłaby źródłem danych (big data) służących do wielu różnych analiz i budowy korelacji.

Tak zaprojektowana powszechna usługa chmurowa uwidacznia potrzebę współdziałania systemów bibliotecznych nie tylko ze sobą nawzajem, ale także z innymi systemami informacyjnymi, tworzonymi w innych środowiskach. Dzięki takiemu współdziałaniu można by na przykład połączyć dane dotyczące geolokalizacji z danymi o dostępności egzemplarza (gdzie znajduje się biblioteka, w której dostępny jest egzemplarz poszukiwanej książki?). Dla osiągnięcia takiego efektu niezbędna jest możliwość uzyskania dostępu do wszystkich danych bibliotecznych (rekordu bibliograficznego, egzemplarza, KHW) bez potrzeby korzystania z pośrednictwa interfejsu systemu informacyjnego. Dostępne powinny być też informacje o wszystkich relacjach funkcjonujących w zbiorze metadanych, np. powiązań rekordu KHW z rekordami bibliograficznymi lub rekordów egzemplarza z rekordami użytkowników. Dzięki temu możliwe jest wyszukiwanie według wszystkich elementów opisu i innych metadanych oraz każdej ich kombinacji, nie tylko tych, które zaprojektowane zostały przez twórcę interfejsu systemu bibliotecznego dla standardowego użytkownika lokalnego.

Dla osiągnięcia opisanych celów systemy biblioteczne muszą stać się systemami wzajemnie otwartymi. Wówczas można oczekiwać, że z funkcjonalnego punktu widzenia działać będą jak jeden system. Osiągnąć to można dwoma sposobami:

- zastosowanie jednego, wspólnego systemu komercyjnego, w ramach którego poszczególne biblioteki działać mogą jako zintegrowane części większej całości. W ten sposób działają systemy nowego rodzaju, tzw. Library Services Platform, takie jak system WorldShare Management Services OCLC, pozwalający na ujednoczenie procedur zarządzania wszystkimi obiektami informacyjnymi (drukowanymi i elektronicznymi). W tej aplikacji chmurowej zasób każdej biblioteki lokalnej

i rekordy zasobu łączone są z rekordem światowego katalogu centralnego WorldCat. Podobnie działa Alma, ale tylko w obrębie bibliotek posiadających ten system (katalog centralny tworzony jest dla bibliotek użytkujących Almę),

- zastosowanie współdziałających systemów otwartych (Open Source), dla których nie ma problemów natury licencyjnej, utrudniających pełne współdziałanie (pozostają problemy natury technicznej i mentalnej).

Chmurowe aplikacje systemów komercyjnych mają tendencję zamykania metadanych i relacji między nimi w lokalnej chmurze danych, czym przypominają nieco nowe osiedla budowane w naszych miastach, dokładnie odgrozione od podobnych, sąsiednich osiedli. Takie postępowanie daje złudne poczucie bezpieczeństwa, jednocześnie nie poprawiając współdziałania systemów i ich metadanych, podobnie jak płoty nie ułatwiają integracji sąsiedzkiej.

Technologią zasadniczo zmieniającą tę sytuację jest Linked Open Data (LOD), związana z pracami nad Semantycznym Web (Nahotko, 2014, s. 5). W szczególności chodzi o zastosowanie RDF (ang. *Resource Description Framework*) i wyników prac dotyczących modelowania metadanych z jego wykorzystaniem (Dunsire i in., 2012, s. 4). Linked Data ma służyć tworzeniu danych dotyczących wszelkich obiektów, nie tylko bibliograficznych, ale także na przykład ludzi, organizacji, procesów i pojęć. Są to dane utworzone i opublikowane w taki sposób, aby umożliwić tworzenie połączeń pomiędzy zestawami danych oraz słownikami. Te ostatnie rozumiane są szerzej niż to było dotąd przyjęte w bibliotekarstwie, gdyż nazwą tą obejmuje się zarówno słowniki stosowane na poziomie pragmatyki, zawierające wartości metadanych, zwane także słownikami kontrolowanymi (KHW, słowniki JIW, wykazy kodów nazw języków, nazw geograficznych itp.), jak również zestawy pojęć stosowane na poziomie semantycznym, dotąd nazywane formatami lub schematami (wykazy elementów metadanych, pól i podpól formatów metadanych wraz z ich definicjami). W ten sposób do chmury danych trafiają słowniki zarówno poziomu semantycznego, jak i pragmatycznego metadanych, kodowane przy użyciu standardowych języków kodowania, a więc dostępne do bezpośredniego przetwarzania komputerowego. Każdy element słownika jest jednoznacznie identyfikowany przy pomocy identyfikatora; na ogół stosowany jest URI (ang. *Uniform Resource Identifier*) ze względu na jego otwartość.

Relacje pomiędzy identyfikowanymi elementami (dla danych bibliograficznych: dzieła, materializacje, egzemplarze, osoby, instytucje, miejsca, elementy treści itp.) opisywane są przy pomocy języków, takich jak RDF i OWL (ang. *Web Ontology Language*), których wyrażenia mogą być przetwarzane bezpośrednio przez komputery. Relacje te pozwalają na nawigację pomiędzy danymi zawartymi w źródłach publikowanych w Web, co powoduje ich integrację ułatwiającą współdziałanie techniczne na poziomie syntaktyki. Współ-

działanie organizacyjne (prawne) uzyskiwane jest dzięki otwarciu danych. Otwartość danych nie jest niezbędna dla osiągnięcia współdziałania technicznego, ale umożliwia pełną efektywność stosowanych technologii.

Z punktu widzenia współdziałania metadanych tworzonych w tym środowisku ważne jest oddzielenie semantyki od syntaktyki metadanych. W Linked (meta)Data znaczenie metadanych (semantyka) jest oddzielone od ich syntaktyki, dzięki czemu zmiana struktur nie powoduje modyfikacji znaczenia. Tylko pozornie różnica polega na zamianie jednych standardów innymi. W rzeczywistości nowe standardy mają jedną, ale bardzo ważną nową cechę – zarówno semantyka, jak i pragmatyka kodowane są przy pomocy tej samej syntaktyki (głównie XML/RDF, choć istnieją inne serializacje), a co więcej, jest to syntaktyka, która umożliwia automatyczne przetwarzanie danych (dzięki jej „rozumieniu” przez komputery), bez udziału człowieka. Obecnie wszystkie duże, uniwersalne klasyfikacje dostępne są w wersji dla Linked Data z użyciem SKOS (ang. *Simple Knowledge Organization System*). Dawne schematy metadanych nadal są używane (łącznie ze strukturą pól i podpól MARC 21), ale struktury te są kodowane w RDFS (ang. *RDF Schema*) lub OWL (zmianie uległa syntaktyka), przez co mogą stać się ontologiami umieszczanymi w chmurze danych. Do każdego symbolu klasyfikacji i do każdej relacji wyrażonej strukturą MARC (a więc do etykiety każdego pola tego formatu) oraz między tymi symbolami i etykietami można poprowadzić kwalifikowany odnośnik, czyli taki, który identyfikuje i nazywa relację występującą pomiędzy połączonymi elementami.

Tak zorganizowane środowisko metadanych umożliwia dwojakiego rodzaju współdziałanie metadanych. Pierwszy sposób na zapewnienie współdziałania związany jest z istnieniem słowników dwóch opisanych rodzajów, przydatnych szczególnie, gdy chcemy łączyć dane z różnych obszarów (dziedzin wiedzy, zastosowań praktycznych). Użytkownik Web może przeglądać jego zasoby bez potrzeby znajomości stale zmieniających się technologii i wynikających z nich struktur, stanowiących podstawę działania Sieci. Podczas przeglądania swobodnie przekraczane są granice między zasobami, bez względu na ich fizyczne oddalenie. Na tej samej zasadzie można w Linked Data przeglądać zbiory danych podążając za odnośnikami od jednego zasobu do innego, nawet wtedy, gdy są fizycznie umieszczone w różnych miejscach i zapisane w różnych formatach (Bermes, 2011).

Zasobami linkowanych danych są także zawartości słowników obu wcześniej wymienionych rodzajów. Umożliwiają one współdziałanie dzięki temu, że działają jak centrum przełącznikowe, umieszczone w chmurze danych, łączące dane wyrażone zgodnie z różnymi semantykami danych. Takie centrum umożliwia w Linked Data nawigację od jednego zbioru metadanych do innego, przez podążanie za odnośnikami, czyli URI, nawet jeżeli dane podłączone do centrum są heterogeniczne. Wyobraźmy sobie na przykład, że ist-

nieją dwa zasoby metadanych, w których strukturach znajdują się m.in. symbole UKD. Bez względu na to, w jakich miejscach różnych struktur metadanych (polach, elementach) umieszczone zostały te symbole, odsyłają one przy pomocy URI do ontologii UKD, wykonanej w SKOS i dostępnej w Internecie. Dzięki temu, po pierwsze następuje ujednoczenie znaczenia (semantyki) danych, po drugie możliwe jest wyszukiwanie podobnych danych poprzez proste podążanie za odnośnikami URI. Sposób ten pozwala na unikanie niejednorodności semantyki dzięki porozumieniu ontologicznemu (Heath & Bizer, 2011, s. 24).

Linked Data stanowi system rozproszony. Nie istnieje jedno, główne centrum danych, ale dowolnie wiele takich centrów, połączonych ze sobą odnośnikami. Jako centrum przełącznikowe w Linked Data może funkcjonować dowolny zasób udostępniony w tej technologii. Odnajdywanie informacji może polegać na podążaniu od jednego centrum danych do innego centrum za odnośnikami łączącymi dane, które tam się znajdują. Zaletą takiego postępowania jest to, że proces ten odbywa się w sposób intuicyjny. Jest to druga metoda zapewnienia współdziałania metadanych, nazwana „podążaniem za własnym nosem”. Wszystkie zasoby łączone odnośnikami, budowane z wykorzystaniem standardów RDF i URI stanowią globalny graf informacyjny, dostępny dla niczym nieograniczonego manualnego przeglądania przez użytkowników (ludzi) i automatycznego skanowania przez aplikacje podążające za odnośnikami URI. Odnośniki te opisują relacje charakteryzujące każdy element danych, w które wchodzi on z innymi elementami. Element danych identyfikujący osobę może na przykład być charakteryzowany przez przyłączone do niego obiekty, miejsca, inne osoby itp. przy pomocy odnośników oznaczających relacje bycia autorem, pracownikiem, miejscem urodzenia lub pobytu, ojcem/matką, mieszkańcem, fanem i dowolnej innej. Taka forma aktywności bywa obrazowo nazywana „toURIsm'em” (od *tourism* – ang. turystyka).

Linked Data wciąż jest zamierzeniem pozostającym w sferze raczej idei niż działających aplikacji, choć prace służące tworzeniu nowych standardów dają pierwsze pozytywne rezultaty. Opisane rozwiązania mogą w niedalekiej przyszłości stać się podstawą tworzenia przez dostawców oprogramowania bibliotecznego zintegrowanych systemów bibliecznych nowej generacji (Wilson, 2012, s. 110). Integracja jest tu rozumiana inaczej lub nie tylko tak, jak dotychczas w zintegrowanych systemach bibliecznych, jako integracja modułów oprogramowania i procesów realizowanych w ramach systemu bibliotecznego. W nowych systemach integracja polegać ma na ścisłym współdziałaniu systemu bibliotecznego ze światem zewnętrznym, co być może jest warunkiem przetrwania i dalszego rozwoju tych systemów. Według Marshalla Breedinga (2012, s. 14) wkrótce systemy tego typu pozwolą na tworzenie większości metadanych jako zasobu LOD, co spowoduje efekt śnieżnej kuli.

Każdy sukces osiągnięty w praktycznym zastosowaniu technologii Linked Data będzie powodował dalszy postęp w rozwoju dostępu bibliotek i użytkowników do zlinkowanych metadanych i zasobów informacji. Otwartość technologii i zasobów może spowodować redukcję uzależnienia od rozwiązań komercyjnych, co w pewnym stopniu obserwujemy już obecnie (Nahotko, 2014, s. 18).

Dzięki technologii LOD biblioteki zyskują takie możliwości, jak:

- udostępnienie własnych danych jako LOD dla innych systemów informacyjnych, nie tylko, a nawet nie przede wszystkim bibliotecznych,
- pobieranie i wykorzystanie LOD z innych, zewnętrznych zasobów wszelkiego rodzaju dla wzbogacenia własnych danych,
- stosowanie LOD w celu stworzenia całkowicie nowej, niezależnej od dotychczasowych dostawców, Infrastruktury Web jako podstawy tworzenia metadanych, co jest celem projektu BIBFRAME (ang. *Bibliographic Framework Initiative*).

Cele te wprost i bezpośrednio prowadzą do stworzenia środowiska, w którym zapewnione zostanie współdziałanie na wszystkich poziomach systemów informacyjnych.

Zakończenie

W artykule przedstawione zostały technologiczne aspekty rozwoju OPAC, w tym katalogów centralnych. Jest to bardzo ważne zagadnienie, gdyż katalogi biblioteczne zawsze funkcjonowały dzięki stosowaniu aktualnie innowacyjnych technologii, począwszy od tabliczek glinianych i zwojów papierowych po najnowsze technologie elektroniczne. Istnieje jednak dodatkowy aspekt projektowania systemów informacyjnych, który nie jest natury technicznej. Są nimi użytkownicy systemów informacyjnych – ludzie (w odróżnieniu od użytkowników – programów komputerowych). Potrzeby użytkowników powinny być głównym czynnikiem implementacji technologii informacyjnych i wyznacznikiem budowy standardów, takich jak model FRBR (ang. *Functional Requirements for Bibliographic Records*) i zasady katalogowania RDA (ang. *Resource Description and Access*), zbliżających biblioteki w kierunku LOD. Pomimo tego, że o potrzebach i zadaniach użytkownika mówi się od czasów Panizziego i Cuttera, poprzez Zasady Paryskie aż do FRBR, problem ten nie został ostatecznie rozwiązany. Jeżeli bowiem zintegrowane systemy informacyjne działałyby tak, jak to przedstawiono w artykule, to ich funkcjonalność nie może kończyć się na dostarczeniu żądanych obiektów informacyjnych. Wręcz przeciwnie, dopiero na tej podstawie powinny być budowane kolejne usługi eksploracji zasobów informacyjnych.

Zastosowanie nowych technologii powoduje, że OPAC, w tym katalogi centralne, stają się częścią uniwersalnego środowiska informacyjnego komputerowych sieci globalnych. Z tego natomiast wynika, że zamiast (a może obok) tradycyjnej dbałości o jakość metadanych, będących efektem pracy intelektualnej, bibliotekarze muszą dbać także o ich współdziałanie ze strukturami stosowanymi w innych środowiskach. Przejawem tej dbałości są na przykład starania o zapewnienie współdziałania standardów takich jak ONIX (przemysł księgarski) i MARC (Mitchell, 2013, s. 299). Jest to tym bardziej pożądane, że nasi użytkownicy, o potrzebach których była przed chwilą mowa w artykule, funkcjonują właśnie w tych innych środowiskach (Facebook, YouTube, Instagram, Amazon...), spędzając tam znacznie więcej czasu, niż w bibliotece. Stosując często przywoływaną analogię, powinniśmy zburzyć mur postawiony wokół wypielegnowanych rabatki naszych OPAC otwierając je na działanie ludzi znajdujących się poza tym murem, nawet, jeśli narazimy rabatki na podeptanie podczas świadczenia znacznie obszerniejszych usług informacyjnych, niż dotychczas.

Oznacza to, że cechy OPAC powinny predestynować go do funkcjonowania mniej jako inwentarza, stanowiącego spis majątku instytucji, a bardziej jako systemu informacyjnego – bibliografii, a może nawet wyszukiwarki internetowej, oferującej użytkownikom wszystko, co zaspokaja ich potrzeby informacyjne, bez względu na to gdzie znajdują się żądane materiały, w jakiej formie (drukowana, elektroniczna) oraz w jakich sposobach dostępu. Okazuje się więc, że nasze OPAC czekają dalsze zmiany, a jedyne, co można w tym zakresie uważać za niezmiennie, to stała potrzeba tworzenia nowych rozwiązań.

Bibliografia

- Bermes, Emmanuelle (2011). Convergence and interoperability: a Linked Data perspective. In: *World Library and Information Congress: 77th IFLA General Conference and Assembly*. Puerto Rico, 13–18 August 2011. Pobrane 28 maja 2019, z: <https://www.ifla.org/past-wlic/2011/149-bermes-en.pdf>
- Breeding, Marshall (2012). Current and future trends in information technologies for information units. *El profesional de la información*, 21, 9–15.
- Carney, David J. & Smith, James & Place, Patrick R. (2005). *Topics in interoperability: infrastructure replacement in a system of systems*. Report nr CMU/SEI2005-TN-031. Pittsburgh: Carnegie Mellon Univ.
- Chan, Lois M. & Zeng, Marcia L. (2006). Metadata interoperability and standardization – a study of methodology part I. Achieving interoperability at the schema level. *D-Lib Magazine*, 12 (6). Pobrane 25 maja 2019, z: <http://www.dlib.org/dlib/june06/chan/06chan.html>
- Coyle, Karen (2017). *Creating the catalog, before and after FRBR*. Pobrane 15 czerwca 2019, z: <http://kcoyle.net/mexico.html>
- Dunsire, Gordon i in. (2012). Linked Data vocabulary management: infrastructure support, data integration, and interoperability. *Information Standards Quarterly*, 24, 4–13.

- Heath, Tom & Bizer, Christian (2011). *Linked Data. Evolving the Web into a Global Data Space*. San Rafael, CA: Morgan & Claypool Publ.
- IFLA (2005). Sharing of bibliographic information and resources. Pobrane 12 czerwca 2019, z: <http://archive.ifla.org/VII/d4/pub/InteroperabilityStandards.pdf>
- Katal, Avita & Wazid, Mohammad & Goudar, Rayan H. (2013). Big data: issues, challenges and good practices. In: M. Parashar (ed.), *Sixth International Conference on Contemporary Computing (IC3)*, Noida, India 8–10 Aug. 2013 (p. 404–409). Piscataway, NJ: IEEE.
- Kienzler, Iwona (2003). *Słownik terminologii komputerowej angielsko-polski i polsko-angielski*. Gdynia: IVAX.
- Krajewski, Markusz (2017). Tell data from meta: tracing the origins of big data, bibliometrics, and the OPAC. *OSIRIS*, 32 (1), 224–240.
- Manso-Callejo, Miguel & Wachowicz, Monica & Bernabé-Poveda, Miguel-Angel (2009). Automatic metadata creation for supporting interoperability levels of spatial data infrastructure. Pobrane 12 maja 2019, z: <http://www.gsdi.org/gsdiconf/gsdi11/papers/pdf/194.pdf>
- Mitchell, Erik (2013). Programmatic tools and the implications of automation in the next generation of metadata. *Technical Services Quarterly*, 30, 296–310.
- Nahotko, Marek (2013). Współdziałanie metadanych w systemach informacyjnych. *Zagadnienia Informacji Naukowej*, 1, 61–83.
- Nahotko, Marek (2014). Współdziałanie metadanych w chmurze. *Przegląd Biblioteczny*, 1, 3–24.
- Nilsson, Mikael (2010). *From interoperability to harmonization in metadata standardization*. Stockholm: KTH School of Computer Science and Communication.
- NISO (2004). *Understanding metadata*. Bethesda: NISO Press.
- Pacek, Jarosław (2010). *Bibliografia w zmieniającym się środowisku informacyjnym*. Warszawa: Wydawnictwo SBP.
- Radwański, Aleksander (2015). System biblioteczny jako powszechna usługa sieciowa i baza danych. W: M. Odlaniecka-Poczobutt (red.), *Systemy biblioteczne nowej generacji. Platformy usług* (s. 31–38). Gliwice: TYPO Usługi Wydawnicze.
- Sandy, Heather M. (2019). Explaining cataloging to a six year old? *Technical Services Quarterly*, 36, 379–390.
- Su, Shiao-Feng (1994). Dialogue with an OPAC: how visionary was Swanson in 1964? *The Library Quarterly*, 64, 130–161.
- Subieta, Kazimierz (1999). *Słownik terminów z zakresu obiektowości*. Warszawa: Akademicka Oficyna Wydawnicza PLJ.
- Taylor, Arlene & Joudrey, Daniel (2008). *The Organization of information*. Westport: Libraries Unlimited.
- Tennant, Roy (2002). MARC must die. *Library Journal*, October 15, 26–28.
- Wilson, Kristen (2012). Introducing the next generation of library management systems. *Serials Review*, 38, 110–123.
- Zeng, Marcia L. & Chan, Lois M. (2006). Metadata interoperability and standardization – a study of methodology part II. Achieving interoperability at the record and repository levels. *D-Lib Magazine*, 12. Pobrane 5 czerwca 2019, z: <http://www.dlib.org/dlib/june06/zeng/06zeng.html>
- Zeng, Marcia L. & Qin, Jian (2008). *Metadata*. New York: Neal-Schuman Publ.

Do (union) library catalogues are still necessary? OPAC in the infotopia

ABSTRACT: The article presents the basic function of union catalogues: ensuring interoperability of library metadata. The problem of metadata interoperability was defined as an element of cooperation of information systems, especially functioning in the network. The methods used to ensure interoperability were presented. Three technologies were described that should facilitate metadata interoperability in the future: NoSQL databases, cloud applications and Linked Open Data.

KEYWORDS: metadata interoperability, union catalogues, OPAC, big data, Linked Open Data