

LINKED DATA – MODEL DANYCH POWIĄZANYCH W SEMANTIC WEB

Marcin Roszkowski
Instytut Informacji Naukowej
i Studiów Bibliologicznych
Uniwersytetu Warszawskiego

Semantic Web, Linked Data, reprezentacja wiedzy

Tim Berners-Lee w 2007 r., wypowiadając się na temat przyszłości World Wide Web [5], stwierdził, że sukces tego otwartego środowiska informacyjnego jest wynikiem trzech czynników:

- nieograniczonych możliwości łączenia informacji, dokumentów za pomocą np. hiperłączy;

- zastosowania otwartych standardów jako podstawy rozwoju i innowacji aplikacji (m.in. działalność W3C i standardy opisu informacji, np. HTML, XML, RDF, Dublin Core, czy przesyłania danych, np. HTTP, FTP);

- rozdzielenia warstw sieciowych odpowiedzialnych za komunikację, przesyłanie i przetwarzanie informacji i jej wizualizację.

Te trzy cechy charakteryzują dzisiejszą sieć, która jest siecią:

- stworzoną przez ludzi i dla ludzi,

- zbudowaną w większości z dokumentów, które zawierają dane zamknięte w ich strukturze,

- zamkniętych i niepowiązanych ze sobą baz danych, które często porównywane są z silosami danych [5].

Dzisiejsze World Wide Web to sieć dokumentów – statycznych lub generowanych przez aplikacje bazodanowe, ale to wciąż sieć niezrozumiała dla aplikacji je przetwarzających, które „widzą” w nich ciągi znaków oraz zestawy komend związanych ze sposobem ich wyświetlania. Hipertekst, dzięki węzłom i hiperłączom, dał możliwość wizualizacji i eksploracji przestrzeni informacyjnej. Powiązania hipertekstowe, zarówno inter- jak i intratekstualne, odsyłają jednak do dokumentów lub konkretnych miejsc w ich strukturze (np. nagłówek, paragraf). Tym samym centralnym punktem w takiej koncepcji sieci jest dokument, jego metadane oraz powiązania z innymi dokumentami w sieci. Dane zawarte w dokumencie, jego struktura oraz sposoby wyświetlania stanowią jeden ciąg kodu, co utrudnia automatyczną ekstrakcję informacji treściowej i związków znaczeniowych z innymi dokumentami.

W ciągu ostatnich dwudziestu lat dokonała się ewolucja zarówno koncepcji WWW, jak i technologii w niej wykorzystywanych. Obecnie jesteśmy świadkami funkcjonowania trzech generacji sieci. Web 1.0, czyli pierwsza generacja sieci obecna od początku lat 90. ubiegłego stulecia, charakteryzowała się opisem

struktury dokumentu za pomocą języków znacznikowych, przede wszystkim HTML. Mieliśmy do czynienia w większości z dokumentami statycznymi, gdzie metadane nadawane był praktycznie wyłącznie przez twórców dokumentów, projektantów serwisów, rzadko automatycznie przez oprogramowanie i przez użytkowników. Druga generacja sieci – Web 2.0, to sieć partycypacji, interakcji i integracji. Jest to sieć współuczestnictwa użytkowników w tworzeniu jej zawartości, ale także masowe tworzenie metadanych przez użytkowników (np. folksonomie i tagowanie). Integracja w wymiarze społecznym polega na tworzeniu grup użytkowników wokół serwisów społecznościowych, a w wymiarze technologicznym na tworzeniu systemów pobierających informacje z innych serwisów, na przykład na zagnieżdżaniu aplikacji Google, Youtube. Trzecia generacja sieci, czyli Web 3.0 czy też Semantic Web to sieć danych, które dzięki wspólnemu językowi opisu oraz tworzonym przez użytkowników powiązaniom mogą być automatycznie przetwarzane przez programy komputerowe. Semantic Web to sieć, w której najważniejsze są metadane i struktura [22].

Proces tworzenia i rozpowszechniania metadanych w przypadku sieci Web 3.0 zostaje wsparty przez mechanizmy automatycznie je przetwarzające, które generują rezultaty wyszukiwania. Społecznościowy charakter metadanych jest tutaj zachowany, nie jest to sieć maszyn. Zmienia się tylko sposób opisu rzeczywistości poprzez jego sformalizowanie. Mamy więc do czynienia z przejściem od sieci dokumentów zawierających dane dostępne za pomocą wyszukiwarek i przeglądarek, przez sieć aplikacji prezentujących w jednym miejscu zawartość wielu serwisów i aplikacji, do sieci danych, które wewnątrz są uporządkowane i powiązane dzięki wspólnemu modelowi reprezentacji wiedzy.

Semantic Web to wizja sieci, która dąży do utrzymania postulatów „każdy może wypowiedzieć się na dowolny temat” (tzw. postulat 3xA – Anyone can say Anything about Any topic) [1], która ma się „przyczynić do utworzenia i rozpowszechnienia standardów opisywania treści w Internecie w sposób, który umożliwi maszynom i programom (tzw. agentom) przetwarzanie informacji w sposób odpowiedni do ich znaczenia.” [34] Jest to wizja Tima Berners'a-Lee, którego zdaniem zrozumienie przez aplikację danych polegałoby na zastosowaniu takiego modelu ich opisu, który zakładałby możliwość powiązania ich znaczenia w ramach wspólnego kontekstu.

Linked Data – założenia

Semantic Web nie polega tylko na umieszczaniu danych w sieci, lecz także na tworzeniu powiązań między nimi. Ma to umożliwić manualną lub automatyczną eksplorację zbiorów i odkrywanie nowych danych. Za pomocą danych powiązanych ze sobą tego typu przeszukiwanie staje się możliwe [6]. Koncepcja danych powiązanych Tima Berners'a-Lee, znana jako *Linked Data*, polega na wykorzystaniu World Wide Web i jego technologii do tworzenia formalnych połączeń między danymi pochodzącymi z różnych zbiorów. Umożliwia tworzenie relacji pomiędzy zbiorami danych operującymi różnymi formatami opisu, które różnią się zarówno pod względem treściowym, jak i formalnym. Jest to możliwe dzięki zastosowaniu wspólnego modelu reprezentacji wiedzy oraz rozwiązań

technologicznych uznawanych obecnie za standardy w WWW. Linked Data odwołuje się do metod ekspresji, reprezentacji, łączenia i współdzielenia danych w Semantic Web, które wykorzystują istniejące standardy i narzędzia sieciowe [36]. Nie jest to nowy schemat metadanych, system reprezentacji wiedzy, lecz zestaw wytycznych dotyczących publikowania i udostępniania danych w sieci semantycznej.

Koncepcja Linked Data jest oparta na języku RDF (Resource Description Format), formalnym modelu reprezentacji wiedzy, rekomendowanym przez Konsorcjum WWW. Polega na tworzeniu deklaracji opisujących otaczającą nas rzeczywistość i ustanawianiu związków znaczeniowych pomiędzy jej elementami. Linked Data wychodzi więc poza organizację rzeczywistości wirtualnej i jest gotowe do opisu świata realnego.

Linked Data jako model publikowania, łączenia i udostępniania danych w Semantic Web zakłada cztery ogólne reguły postępowania [6]:

- używaj URI (Uniform Resource Identifier) jako nazw i sposobu odróżniania obiektów;
- stosuj protokół HTTP, aby uzyskać informacje o opisywanych obiektach;
- udostępniaj użyteczne informacje o obiekcie identyfikowanym przez jego URI za pomocą standardów RDF/XML;
- wykorzystuj powiązania z innymi obiektami za pomocą URI, aby zapewnić możliwość eksploracji i odkrywania informacji o innych obiektach.

Pierwsza zasada odnosi się do tworzenia unikalnych nazw dla obiektów i zasobów za pomocą standardu URI. Druga i trzecia zasada odwołują się do wykorzystania sieciowego modelu organizacji World Wide Web do pozyskiwania informacji o zidentyfikowanych obiektach. Czwarta kładzie nacisk na tworzenie powiązań między obiektami i zasobami [11].

Technologie Linked Data

Koncepcja Linked Data jest oparta na dwóch, fundamentalnych dla WWW, technologiach: identyfikatorach URI [4] (Uniform Resource Identifier) oraz protokole przesyłania danych HTTP. Identyfikatory URI nie muszą w swojej budowie wskazywać na umiejscowienie obiektu, który oznaczają (tak jak w przypadku URL). Za ich pomocą można zidentyfikować dowolną jednostkę zarówno w świecie wirtualnym, jak i rzeczywistym, konkretną i abstrakcyjną. Celem standardu URI jest identyfikacja zasobów i obiektów za pomocą unikalnych ciągów znaków [32].

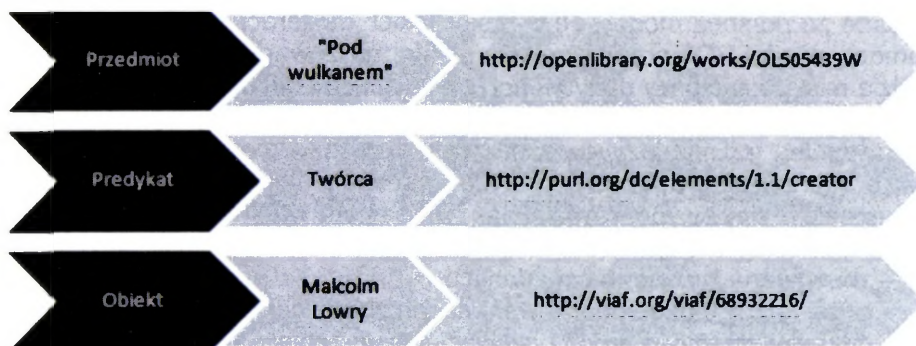
Protokół komunikacyjny HTTP (Hypertext Transfer Protocol) zapewnia znormalizowany sposób komunikacji między komputerami. Umożliwia przesłanie informacji o obiekcie lub zasobie zidentyfikowanym za pomocą URI do maszyny, w której wywołano identyfikator za pomocą poleceń protokołu HTTP. Jest on uniwersalnym mechanizmem pozyskiwania zasobów lub informacji o obiektach nieposiadających postaci cyfrowych, które nie mogą zostać przesłane bezpośrednio za jego pomocą.

O ile identyfikatory URI i protokół komunikacyjny HTTP stanowią podstawę technologii Linked Data, to ideą tej koncepcji jest zastosowanie jednego, współ-

nego dla wszystkich źródeł danych modelu reprezentacji wiedzy o obiektach i zasobach. Tę funkcję pełni RDF jako język reprezentacji wiedzy stosowany do wyrażania metainformacji o obiektach i relacjach, w jakie wchodzi z innymi. „Jest to specyfikacja modelu metadanych, ustanowiona przez World Wide Web Consortium (W3C) na podstawie reguł języka XML. Podstawowym założeniem wyjściowym jest tutaj rozdzielenie warstwy wizualizacji informacji od warstwy ich przetwarzania, czyli oddzielenie struktury dokumentu od jego treści. RDF udostępnia przede wszystkim standaryzowaną syntaktykę wykorzystywaną do reprezentacji metainformacji dla obiektu informacyjnego. Oznacza to, że określa reguły gramatyczne tworzenia interpretowalnych przez programy „wypowiedzi” o obiektach informacyjnych i ich właściwościach. RDF zakłada trójelementową budowę takich wyrażen, które w jej dokumentacji noszą nazwę *trójek RDF* (ang. *RDF triples*). Składają się one z:

- podmiotu, który jest opisywanym zasobem,
- predykatu określającego opisywaną własność,
- obiektu, który określa wartość tej własności” [31].

Przedmiot, który jest identyfikowany za pomocą jego unikalnego URI w danym zbiorze, jest charakteryzowany poprzez wskazanie na jego własność (predykat) oraz wartość tej cechy (obiekt). Najbardziej pożądaną sytuacją jest taka, w której wszystkie elementy takiej deklaracji są przywoływane za pomocą ich unikalnych URI (rys. 1).



Rys. 1. Budowa deklaracji RDF

Odwzorowanie informacji o przedmiocie odbywa się poprzez odwołania bezpośrednio do identyfikatorów wskazujących na własność oraz jej wartość, a pośrednio do sformalizowanych przestrzeni nazw. Są to zbiory słownictwa, na przykład w postaci schematów metadanych, słowników systemów organizacji wiedzy, ontologii internetowych, które opisuje się za pomocą standaryzowanych języków reprezentacji. Należą do nich *RDF Schema (RDFS)* czy *Web Ontology Language (OWL)*. W przytoczonym przykładzie deklaracja reprezentująca stwierdzenie, że autorem dzieła „Pod wulkanem” jest Malcolm Lowry, stosuje odwołania do:

- URI dzieła „Pod wulkanem” z bibliograficznego zbioru danych projektu Open Library (<http://openlibrary.org/>),

- URI atrybutu dc:Creator ze schematu metadanych Dublin Core,
- URI rekordu kartoteki wzorcowej nazw osobowych VIAF (The Virtual International Authority File, <http://www.viaf.org>).

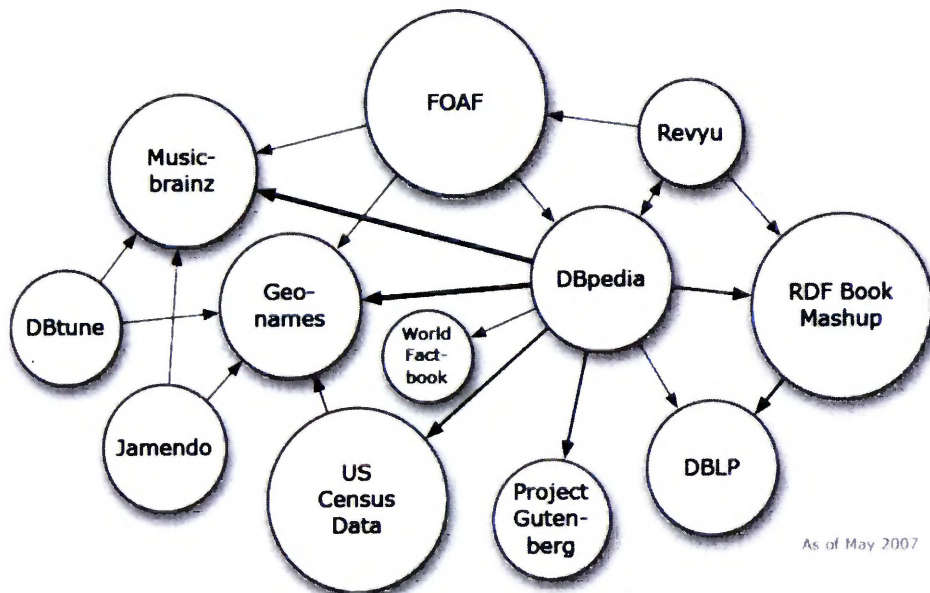
Taki sposób ekspresji powoduje, że aplikacja, która ją odczytuje, za każdym razem odwołuje się do zadeklarowanych przestrzeni nazw, w których zidentyfikowano obiekty, ich własności czy relacje, jakie mogą tworzyć z innymi. Tym samym mamy do czynienia nie z ciągami znaków „niezrozumiałymi” dla programów je przetwarzających, lecz elementami, których kontekst znaczeniowy został zdefiniowany w przestrzeniach nazw. Stąd obecność przymiotnika „semantyczna” w odniesieniu do Semantic Web. U jej podstaw leży stosowanie sformalizowanego systemu odniesienia do opisu jej zasobów. Pozwala to na przetwarzanie zasobów przez aplikacje, które każdorazowo, aby „zrozumieć” znaczenie terminów w deklaracjach, odwołują się do przestrzeni nazw, z których one pochodzą. Termin ma znaczenie dla aplikacji tylko w takim stopniu, w jakim zostało ono wcześniej zdefiniowane w deklaracjach oraz wywołujących przestrzeniach nazw. Koncepcja Linked Data zakłada jedynie łączenie i ekspresję danych wykorzystujących istniejące standardy. W tym przypadku w jednej deklaracji wykorzystano elementy pochodzące z trzech różnych zbiorów słownictwa: dwóch kartotek wzorcowych oraz schematu metadanych. Pomimo heterogenicznego charakteru tych zasobów istnieje możliwość odwoływania się do nich w jednej deklaracji, zachowując jednocześnie warunek przetwarzalności przez aplikacje. Zarówno deklaracja, jak i struktura tych zbiorów są bowiem opisane za pomocą modelu RDF. Zbiór takich deklaracji, skonstruowany za pomocą składni RDF oraz odwołania do sformalizowanych przestrzeni nazw, będzie miał postać bazy danych reprezentującej model Linked Data.

Tak skonstruowana sieć zbiorów danych i powiązań między nimi, operująca wspólną rodziną języków reprezentacji wiedzy, może być postrzegana jako odrębna warstwa World Wide Web. Jest ona umiejscowiona pod warstwą dokumentów sieciowych, rozdzielając tym samym warstwę opisu danych od warstwy ich prezentacji. Jest to sieć:

- zbudowana z różnego typu danych,
- gdzie każdy może opublikować własne zbiory danych,
- gdzie twórcy mogą wykorzystywać dowolne zbiory słownictwa do reprezentacji danych,
- której elementy są połączone ze sobą za pomocą powiązań w składni RDF i tworzą globalną przestrzeń danych obejmującą także całe zbiory danych i słownictwa wykorzystywanego do ich opisu.

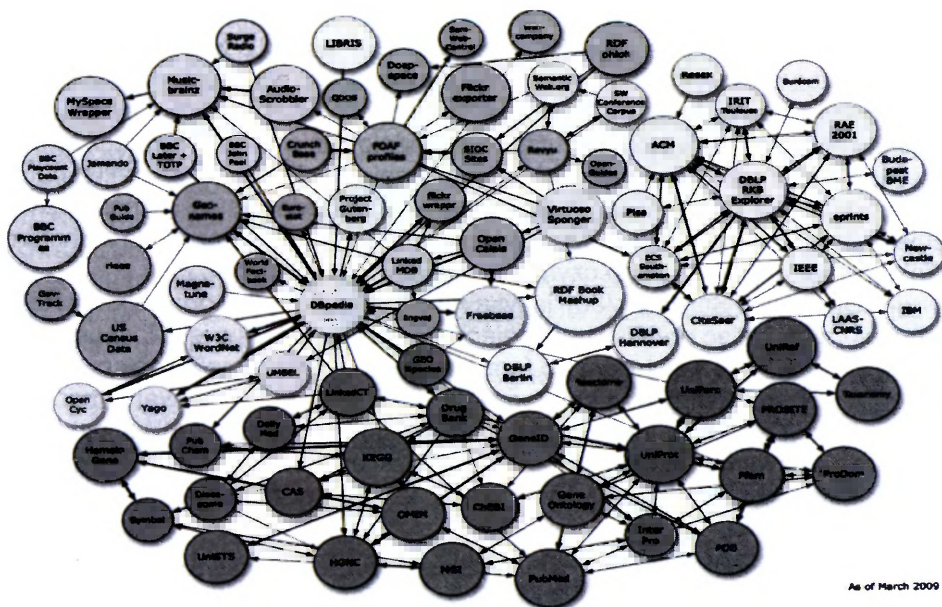
Z punktu widzenia przetwarzania informacji, sieć danych cechuje:

- oddzielenie danych od warstw formatowania i prezentowania;
- samoopisujący się charakter danych (*self-describing*). Deklaracje RDF zawierają odwołania do identyfikatorów własności oraz ich wartości, które są zdefiniowane w przestrzeniach nazw. Aplikacje przetwarzające takie deklaracje są przekierowywane do takich zbiorów w celu identyfikacji kontekstu znaczeniowego dla zastosowanego terminu;
- dostęp do zbiorów danych poprzez interfejsy programowania aplikacji (API wykorzystujące standaryzowany język zapytań SPARQL) oraz protokół przesyłania danych HTTP;



As of May 2007

Rys. 2. Chmura Linked Data w maju 2007 r. [35]



As of March 2009

Rys. 3. Chmura Linked Data w marcu 2009 r. [10]

– otwarty charakter (mechanizmy pozyskiwania danych nie są programowane dla konkretnego zbioru danych, lecz dzięki wspólnemu modelowi reprezentacji oraz powiązaniom RDF mogą eksplorować inne zasoby) [10].

Chmura Linked Data

Zbiory danych opublikowane zgodnie z modelem Linked Data tworzą tzw. chmurę danych powiązanych (*linked data Cloud*), czyli globalną sieć zbiorów danych stosujących wzajemne odwołania. W maju 2007 r. liczyła ona zaledwie kilka zbiorów danych realizujących ten model. Obecnie jest ich ponad 200 i szacuje się, że zawarta w nich łączna liczba deklaracji przekracza 25 miliardów trójek RDF [16].

Diagram chmury danych powiązanych pokazuje wielkość zbiorów danych i występowanie relacji pomiędzy nimi. Zmiany w rozmiarze chmury oraz zakres i kierunki odwołań można prześledzić na stronie internetowej Richarda Cygniaka [15], który od 2007 r. opracowuje tego typu diagramy.

Zasoby chmury Linked Data

Zasoby Linked Data można interpretować w kategoriach gotowych danych opublikowanych za pomocą technologii Semantic Web oraz założeń tego modelu publikowania lub w postaci sformalizowanych przestrzeni nazw (schematów metadanych, ontologii), czyli zbiorów słownictwa stosowanych do ich opisu.

Złożoność otaczającego nas świata oraz jego projekcji zawartej w zasobach sieciowych wymusza stosowanie wielu narzędzi do jego efektywnego opisu i uchwycenia istotnych cech oraz relacji, jakie w nim zachodzą. W kontekście Semantic Web mamy do czynienia z dużą liczbą narzędzi wykorzystywanych do tego celu, zaczynając od schematów metadanych, przez zbiory słownictwa stosowane do wyrażania wartości określonych własności obiektów, po skomplikowane ontologie internetowe służące do modelowania fragmentu rzeczywistości, który za ich pomocą ma zostać formalnie opisany. Formalny charakter opisu sporządzany za ich pomocą, czyli zgodny ze standardami sieci semantycznej, powoduje, że są one gotowymi narzędziami opisu danych włączanych w chmurę Linked Data. Mamy więc do czynienia z wykorzystaniem np.:

- RDFS (RDF Schema) [30], języka reprezentacji wiedzy opartego na RDF, który wprowadzając podstawowe pojęcia (np. pojęcie klasy, jej własności, zakres), służy przede wszystkim do formalnego uporządkowania zapisów;

- OWL (Web Ontology Language) [26], narzędzia do budowania bardziej szczegółowych opisów fragmentów rzeczywistości niż RDFS. OWL jest standardem sieciowym przyjętym przez W3C. Rozszerza on w dużym stopniu możliwości modelowania rzeczywistości RDFS i jest narzędziem o wyjątkowo dużej sile semantycznej;

- Dublin Core Metadata Element Set [20], najczęściej stosowanego schematu metadanych dla obiektów cyfrowych, którego zapisy są wyrażane za pomocą RDF;

- FOAF (Friend of a Friend) [23], schematu formatowania metadanych służących do budowania wizytówek sieciowych i tworzenia sieci społecznych;

- SKOS (Simple Knowledge Organization System) [38], który „jest rozszerzeniem RDF wykorzystywanym do reprezentacji wiedzy w sieciowych systemach organizacji wiedzy. Jest to język opisu prostych systemów organizacji

wiedzy dla zasobów sieciowych” [31]. SKOS jest formalnym językiem opisu słowników systemów leksykalnych zarówno tradycyjnych, jak i sieciowych;

– SIOC (Semantically-Interlinked Online Communities) [37], ontologii stworzonej na bazie RDFS i OWL służącej do modelowania i opisu obiegu informacji w społecznościach sieciowych. Jest to standard służący do opisu treści generowanych przez użytkowników takich społeczności (np. fora dyskusyjne, blogi, galerie zdjęć), który wraz z FOAF jest wykorzystywany do tworzenia zbiorów danych przetwarzalnych przez aplikacje sieciowe;

– DOAP (Description of a Project) [21], ontologii wykorzystującej RDFS oraz m.in. zestaw elementów Dublin Core i FOAF do tworzenia opisów projektów oprogramowania, w szczególności Open Source;

– GeoNames Ontology [42], ontologii dedykowanej, opracowanej na potrzeby projektu GeoNames (<http://www.geonames.org/>), służącej do formalnej ekspresji metadanych geoprzestrzennych;

– Music Ontology [27], ontologii identyfikującej obiekty, własności oraz relacje dotyczące dokumentów dźwiękowych (np. wykonawca, album, utwór, wykonanie, aranżacja, itd.).

W efekcie prac prowadzonych nad technologiami dla Semantic Web opracowano wiele narzędzi tego typu w postaci schematów metadanych, zbiorów słownictwa czy gotowych ontologii internetowych je wykorzystujących.

Zagadnienia związane z Linked Data można rozpatrywać z punktu widzenia zbiorów danych opublikowanych dla sieci semantycznej. Rozmiar chmury Linked Data pokazuje, że mamy do czynienia z ciągle rozrastającą się globalną siecią zbiorów powiązanych danych. Warto wymienić kilka z nich. Są to:

– DBpedia (<http://dbpedia.org/>), czyli „semantyczna” wersja Wikipedii. Jest to projekt Freie Universität w Berlinie oraz Uniwersytetu w Lipsku. W 2007 r. opublikowano pierwszą wersję Wikipedii w formie danych powiązanych. Najnowsza wersja pochodzi z kwietnia 2010 r. i zawiera opisy 3,4 miliona obiektów [9]. Konwersja danych z Wikipedii do jej semantycznej wersji polega na z kodowaniu abstraktu artykułu oraz metadanych zawartych w bocznej części opisu do postaci deklaracji RDF za pomocą ontologii stosowanej w tym projekcie, np. „semantyczna” forma artykułu w DBpedii dla hasła Malcolm Lowry – http://dbpedia.org/page/Michael_Lowry;

– The DBLP Computer Science Bibliography [40], bibliograficzna baza danych z zakresu informatyki. Zawiera opisy ponad 800 tys. jednostek bibliograficznych. Dostępna jest w postaci hipertekstowej oraz „surowej”, czyli deklaracji RDF;

– GeoNames (<http://www.geonames.org/>), baza danych zawierająca dane geograficzne o ponad 2,6 miliona jednostek geograficznych na całym świecie. Została opublikowana w formie graficznej oraz w modelu Linked Data;

– BBC Music (<http://www.bbc.co.uk/music/artists>), baza danych o artystach prezentowanych w rozgłośni BBC. Również dostępna w formie graficznej oraz w modelu Linked Data, np.:

- informacje o grupie Nirvana w formie graficznej: <http://www.bbc.co.uk/music/artists/5b11f4ce-a62d-471e-81fc-a69a8278c7da>,

- informacje o grupie Nirvana w formie deklaracji RDF: <http://www.bbc.co.uk/music/artists/5b11f4ce-a62d-471e-81fc-a69a8278c7da.rdf>;

– CIA World Factbook (<http://www4.wiwiss.fu-berlin.de/factbook/>), publikacją Centralnej Agencji Wywiadowczej USA, która zawiera podstawowe dane na temat wszystkich krajów świata. Wersję bazy danych dla przetwarzania przez aplikacje przygotowano w Freie Universitat w Berlinie;

– Dane rządowe, raporty, analizy, zestawienia statystyczne:

- rządu USA (<http://www.data.gov/catalog/raw>),
- rządu Wielkiej Brytanii (<http://data.gov.uk/>).

Przytoczone przykłady to tylko wybrane inicjatywy związane z publikowaniem danych w chmurze Linked Data. Obszerny zbiór baz danych dla Semantic Web dostępny jest w serwisie Comprehensive Knowledge Archive Network (CKAN) [16], projektu realizowanego przez Open Knowledge Foundation.

Zasada tożsamości

Kluczem do eksploracji chmury danych powiązanych i pozyskiwania danych z wielu źródeł jest możliwość sformalizowania zasady tożsamości. Polega to na zastosowaniu relacji wykładnika języka OWL – `owl:sameAs` jako związku tożsamości znaczeniowej ustanawianego pomiędzy identyfikatorami obiektów z różnych zbiorów. W warstwie technicznej pozwala to na agregację rozproszonych metadanych odnoszących się do wielu obiektów, dla których ustanowiono ten typ zależności na przykład we wspomnianej DBpedii, URI obiektu odnoszącego się do osoby Woodiego Allena powiązано wykładnikiem `owl:sameAs` z URI obiektów z baz New York Times, Open Cyc i Zitgist. Aplikacja napotykalająca na taki zapis będzie mogła uzupełnić dane o Woodym Allenie z Dbpedii o zapisy pozostałych baz dzięki formalnemu wykładnikowi tożsamości. Formalizacja prawa identyczności ma istotne znaczenie dla pozyskiwania danych w chmurze Linked Data. Ważne są także własności tego typu relacji ustalone w ramach języka OWL, np. symetryczność (zob. [18]).

Dostęp do danych i publikowanie Linked Data

Dostęp do zbiorów danych polega na tzw. negocjacji zawartości pomiędzy serwerem udostępniającym dane a aplikacją, która wymusza ich pobieranie. Odbywa się to za pomocą standaryzowanego języka SPARQL (Simple Protocol And RDF Query Language), który jest językiem zapytań i protokołem dla plików RDF. Forma i zakres zapytań oraz zasady przetwarzania pobranych danych są opracowywane dla aplikacji je przetwarzających i tworzą ich interfejsy programowania (API – Application Programming Interface). Inną formę dostępu realizują narzędzia wyszukiwania informacji, które wykorzystują interfejsy graficzne do prezentacji rezultatów wyszukiwania i prezentowania powiązań między danymi. Umożliwiają one również eksplorację tych danych dzięki wizualizacji powiązań za pomocą hipertekstu. Są to różnego typu przeglądarki (np. Marbles – <http://www5.wiwiss.fu-berlin.de/marbles>, Tabulator – <http://www.w3.org/2005/ajar/tab>) oraz wyszukiwarki (np. Falcons – <http://iws.seu.edu.cn/services/falcons/>).

Problematyka publikowania danych w modelu Linked Data została przedstawiona w licznych opracowaniach i poradnikach (np. [8], [19]). Powstało także wiele aplikacji, w większości darmowych, które umożliwiają konwersję danych w istniejących formatach do postaci RDF (tzw. RDFizer) oraz oprogramowania służącego do generowania relacyjnych baz danych na podstawie zbiorów deklaracji (np. D2R Server). [10]

Linked Data a dane bibliograficzne

W styczniu 2010 r. ukazał się raport przygotowany przez American Library Association pt. „Understanding the Semantic Web: Bibliographic Data and Metadata” [40]. W całości poświęcono go zagadnieniom uczestnictwa bibliotek w Semantic Web. Stwierdzono tam, że zmiany, które muszą nastąpić w bibliotekach, mają polegać na przekształceniu katalogu bibliotecznego z formy zamkniętej do postaci zbioru danych powiązanych, gdzie odwołuje się do innych zasobów sieciowych i do którego prowadzą powiązania z innych źródeł danych [13]. Uczestniczenie bibliotek w tworzeniu zasobów Semantic Web oznacza konieczność stosowania modelu Linked Data. Dane bibliograficzne są sporządzane w bibliotekach zgodnie z przyjętymi standardami opisu (m.in. normy krajowe, ISBD, AACR2) oraz na podstawie przyjętych formatów zapisu (np. MARC 21). Zapowiadana w tym dokumencie zmiana paradygmatu ma wymusić formalizację sposobu zapisu metadanych bibliograficznych z wykorzystaniem technologii Semantic Web. Celem zmiany ma być współtworzenie zasobów sieci trzeciej generacji, a nie wyłącznie obecność katalogów w środowisku sieciowym. Wymaga to takiego zapisu i ekspresji danych, który będzie „czytelny” dla aplikacji Semantic Web.

21 maja 2010 r. w ramach konsorcjum World Wide Web powstał projekt o nazwie Library Linked Data Incubator Group [42]. Celem projektu jest pomoc w zwiększeniu interoperacyjności danych bibliotecznych w sieci poprzez stworzenie platformy wymiany poglądów dla osób zainteresowanych tematyką wdrażania koncepcji Linked Data w środowisku bibliotekarskim. Grupę tworzy 43 członków, 20 organizacji współpracujących oraz 10 zaproszonych ekspertów. Do zadań grupy należy:

- zbieranie informacji o projektach wdrażania technologii Semantic Web w bibliotekach,
- wspieranie współpracy pomiędzy podmiotami (biblioteki, muzea, archiwa, wydawcy) zainteresowanymi wprowadzeniem danych o dobrach kultury do przestrzeni informacyjnej Semantic Web,
- identyfikacja modeli danych, zbiorów słownictwa, ontologii oraz zasad i tworzenia lub poprawy interoperacyjności pomiędzy nimi,
- wskazywanie na zapotrzebowanie na nowe standardy, wytyczne i metody działań [2].

Efektom prac prowadzonych w ramach projektu są tzw. wzorce rozwiązań (ang. *case templates*), które zawierają rozwiązania i algorytmy postępowania dla konkretnych projektów dotyczących uczestnictwa bibliotek w Semantic Web. Aktualnie opracowano 13 takich miniprojektów, które dotyczą m.in. adaptacji

istniejących modeli konceptualnych i systemów organizacji wiedzy, wykorzystania technologii Semantic Web do przetwarzania danych bibliograficznych oraz zarządzania i dystrybucji metadanych (http://www.w3.org/2005/Incubator/llid/wiki/Use_Cases).

Problematyka danych bibliograficznych jako Linked Data jest obszerna i obejmuje zagadnienia związane zarówno z konwersją danych, jak i metodami i licencjami na jakich są udostępniane. Pierwszym krokiem jest jednak przygotowanie nowych narzędzi i adaptacja już istniejących do procesu konwersji danych. Wymaga to formalizacji opisu zarówno standardów opisu bibliograficznego, jak i kartotek wzorcowych oraz słowników języków informacyjno-wyszukiwawczych stosowanych do opisu zawartości dokumentów.

Functional Requirements for Bibliographic Records (FRBR) [28] jest modelem konceptualnym, który na najwyższym poziomie abstrakcji porządkuje uniwersum bibliograficzne. Wyznacza on grupy jednostek oraz relacje, jakie mogą między nimi zachodzić w szerokim kontekście opisu bibliograficznego. Konsekwencją takiej perspektywy interpretacji uniwersum bibliograficznego są dwa kolejne modele opracowane przez IFLA: Functional Requirements for Authority Data (FRAD) [24] i Functional Requirements for Subject Authority Data (FRSAD) [25], wykorzystywane do porządkowania struktury kartotek wzorcowych. Pomimo trudności z ich zastosowaniem w prezentacji informacji w katalogach i kartotekach bibliotecznych (np. [3], [12]) modele te pełnią ważną funkcję w aspekcie konceptualnym i terminologicznym. Dlatego też zostały wykorzystane do modelowania uniwersum bibliograficznego w postaci ontologii internetowych z przeznaczeniem do wykorzystania w opisie danych dla Semantic Web. Obecnie trwają prace nad przełożeniem zapisu struktury tych modeli z wykorzystaniem składni RDF i opracowaniem gotowych ontologii internetowych, do których będzie można się odwoływać w deklaracjach [17]. W serwisie Open Metadata Registry (<http://metadatarregistry.org>) można na bieżąco śledzić postępy tych prac. Zamieszczono tam wyniki prac nad „semantyzacją” FRBR, ISBD, FRSAD, a także Resource Description and Access (RDA) [29], czyli nowych anglo-amerykańskich zasad katalogowania, które w założeniu mają być wykorzystywane do opisu zasobów sieciowych.

Obok projektów zmierzających do formalizacji zapisu modeli danych bibliograficznych i standardów katalogowania prowadzi się również badania nad zastosowaniem nowych narzędzi do zapisu metadanych dla jednostek bibliograficznych. Jednym z takich projektów jest ontologia Bibliographic Ontology (BIBO – <http://bibliontology.com>). Jest to prosta ontologia definiująca własności jednostek bibliograficznych, która oprócz własnej specyfikacji wykorzystuje także schemat Dublin Core i wspomniany FOAF. Innym przykładem jest działalność Biblioteki Kongresu na rzecz „semantyzacji” zbiorów słownictwa i formatów opisu. Serwis internetowy (<http://www.loc.gov/standards/marcxml/>) tej biblioteki udostępnia darmowe narzędzia i dokumentację, która pozwala na konwersję formatu MARC 21 do formatu RDF z wykorzystaniem m.in. elementów schematu Dublin Core. Ciekawy przykład inicjatywy spoza środowiska bibliotekarskiego to projekt MarcOnt Bibliographic Ontology [33]. Ontologia MarcOnt (<http://www.marcont.org/>) jest przeznaczona dla bibliotek cyfrowych i wykorzystuje schemat Dublin Core rozszerzony o jednostki, klasy i własności zaprojektowane za

pomocą OWL. MarcOnt ma niewiele wspólnego z formatem MARC pomimo nazwy, która może nasuwać takie skojarzenie.

Podstawę semantycznego charakteru sieci trzeciej generacji stanowią sformalizowane przestrzenie nazw, do których odwołują się deklaracje identyfikujące i opisujące obiekty. Są to zbiory słownictwa wyznaczające kontekst znaczeniowy dla terminów w nich występujących. W licznych pracach analizujących szanse udziału danych bibliograficznych w chmurze Linked Data szczególnie podkreśla się rolę zbiorów słownictwa opracowanego w społeczności bibliotekarskiej. Kartoteki wzorcowe, słowniki języków informacyjno-wyszukiwawczych to gotowe narzędzia leksykalne, na których „semantyzację” czekają twórcy aplikacji dla Semantic Web. Do najważniejszych prac prowadzonych w środowisku bibliotekarskim, które zmierzają do konwersji tego typu źródeł, należą projekt VIAF oraz publikowanie zasobów leksykalnych za pomocą języka SKOS.

VIAF (The Virtual International Authority File, <http://viaf.org/>) to projekt realizowany przez Bibliotekę Kongresu oraz Bibliotekę Narodową Niemiec. Jest to międzynarodowa kartoteka nazw osobowych, której zadaniem jest ułatwienie wyszukiwania informacji poprzez zniesienie barier językowych. W projekcie bierze udział 18 instytucji z całego świata, w tym także z Polski (NUKAT). Obok podstawowego założenia projektu, czyli udostępnienia międzynarodowej kartoteki nazw osobowych, celem ma być publikacja tych danych w chmurze Linked Data (zob. rys 1).

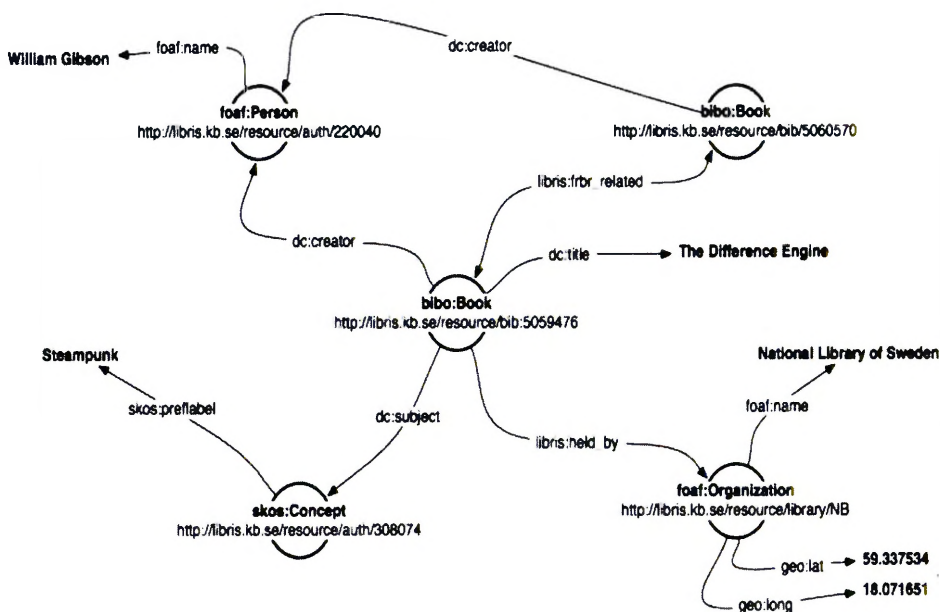
Dzięki opracowaniu formalnego języka opisu słowników prostych systemów organizacji wiedzy, czyli SKOS [33, 38], udostępniono narzędzie służące do przekładu struktury słowników tradycyjnych języków informacyjno-wyszukiwawczych do postaci przetwarzalnej przez aplikacje sieciowe. SKOS oprócz umożliwiania opisu struktury paradygmatycznej systemów organizacji wiedzy działa także jako język przejścia. Pozwala na odwzorowanie zależności semantycznych pomiędzy różnymi systemami organizacji wiedzy. W strukturze SKOS wyodrębniono wykładniki typów zgodności zakresowej pomiędzy wiązanymi w ten sposób jednostkami z różnych źródeł. Cały czas trwają prace nad konwersją zapisu słowników wybranych języków informacyjnych oraz źródeł słownictwa wykorzystywanych do opisu zasobów bibliotek. Do zakończonych projektów należą m.in. :

- AGROVOC Agricultural Thesaurus,
- Australian Public Affairs Information Service (APAIS) Thesaurus,
- DDC Dewey Decimal Classification (projekt eksperymentalny, wybrane klasy najwyższego stopnia podziału),
- GEMET General Multilingual Environmental Thesaurus,
- GeoNames, kartoteka wzorcowa nazw geograficznych,
- HPMulti Thesaurus for Health Promotion,
- ISO 3166-1 Country Codes,
- ISO639 Codes for the representation of names of languages,
- IVOA astronomy vocabularies,
- Kartoteki wzorcowe Biblioteki Narodowej Niemiec,
- Library of Congress Subject Headings,
- Medical Subject Headings,
- NASA taxonomy,

- New York Times subjects,
- RAMEAU,
- STW Thesaurus for Economics,
- The General Multilingual Environmental Thesaurus (GEMET),
- Thesaurus for Graphic Materials,
- Thesaurus for the Social Sciences,
- Thesaurus of territorial units of Spain and France,
- UK Archival Thesaurus (UKAT),
- UK Government Category List (GCL),
- UNESCO Thesaurus [39].

Dotychczas niewiele bibliotek podjęło decyzję o opublikowaniu swoich zasobów w chmurze Linked Data. Są to Niemiecka Biblioteka Narodowa, Biblioteka Narodowa Węgier oraz LIBRIS – katalog zbiorczy szwedzkich bibliotek akademickich. Pierwsza inicjatywa bibliotek w tym kontekście to opublikowanie pod koniec 2008 roku katalogu LIBRIS. Dane z tego zbioru opisano za pomocą wielu narzędzi ekspresji metadanych (rys. 4), m.in.:

- Schematu metadanych Dublin Core,
- Schematu FOAF,
- Ontologii BIBO,
- SKOS jako sposobu ekspresji opisu rzeczowego,
- Opracowanej na potrzeby systemu ontologii LIBRIS.



Rys. 4. Schemat modelowania uniwersum bibliograficznego w projekcie LIBRIS
[\[http://blog.libris.kb.se/semweb/wp-content/uploads/2008/12/linked-libris1.jpg\]](http://blog.libris.kb.se/semweb/wp-content/uploads/2008/12/linked-libris1.jpg)

Rysunek 4 przedstawia w formie schematu zestaw deklaracji dotyczących dokumentu pt. *The Difference Engine*, którego autorem jest William Gibson. Reprezentacja wiedzy dla tego fragmentu uniwersum bibliograficznego została

skonstruowana za pomocą wspomnianych ontologii oraz własności i relacji, które można wyrazić za ich pomocą. W centrum grafu znajduje się URI dla dokumentu, który zidentyfikowano jako *książka* (bibo:Book), której tytuł (dc:title) ma formę *The Difference Engine*, jej autorem (dc:creator) jest obiekt zidentyfikowany jako osoba (foaf:Person), która nazywa się (foaf:name) *William Gibson*. Treść dokumentu jest opisana poprzez wskazanie na jego przedmiot (dc:subject), który jest zidentyfikowany jako *pojęcie* (skos:Concept) pochodzące z określonego systemu organizacji wiedzy i ma formę (skos:preflabel) *Steam-punk*. Książka ta jest w posiadaniu (libris:held_by) organizacji (foaf:Organization) o nazwie (foaf:name) *Biblioteka Narodowa Szwecji (National Library of Sweden)*, która zlokalizowana jest na 59 stopniu szerokości (geo:lat) i 18 stopniu długości geograficznej (geo:long), itd.

Siła ekspresji tego modelu reprezentacji wiedzy (RDF) jest olbrzymia. Wymaga jednak stosowania adekwatnych narzędzi modelowania fragmentu rzeczywistości (ontologii, schematów metadanych, zasobów leksykalnych). W przypadku uniwersum bibliograficznego ten wymóg jest spełniony poprzez funkcjonowanie ugruntowanych standardów i schematów opisu oraz systemów porządkowania pojęć, które pozwalają na identyfikację dzieła na odpowiednim poziomie szczegółowości, wskazanie na istotne cechy jego treści oraz odwzorowanie relacji bibliograficznych, w jakie wchodzi z innymi dziełami. Charakterystyczną cechą modelu reprezentacji wiedzy opartego na RDF jest jego elastyczność. Z punktu widzenia konstruowania metadanych bibliograficznych mamy do czynienia z otwartą formą rekordu bibliograficznego, który tworzą deklaracje składające się na konkretny komunikat o charakterze metainformacyjnym. Wykorzystanie różnych narzędzi opisu jednostek bibliograficznych pozwala na tworzenie kompleksowych narzędzi opisu, dostosowanych do potrzeb organizacji lub opisywanej dziedziny czy domeny. Otwarty model reprezentacji wiedzy oparty na współdzielonej składni (RDF) pozwala na budowanie profili aplikacyjnych dostosowanych do wymogów reprezentacji metadanych w poszczególnych dziedzinach wiedzy, systemach, aplikacjach, usługach (zob. np. zalecenia ustaleń singapurskich dla Dublin Core) [14]. Korzyści płynące dla bibliotek z tego otwartego modelu reprezentacji są dwójakie: z jednej strony mamy możliwość uzupełniania opisów jednostek bibliograficznych poprzez odwoływanie się do źródeł zewnętrznych, z drugiej, to dane bibliograficzne mogą zostać wykorzystane poza kontekstem wyznaczanym przez miejsce i usługi bibliotek w systemach i serwisach informacyjnych.

Problemy

Tim Berners-Lee, mówiąc w 2009 r. o Linked Data [6] na prestiżowej konferencji „TED – Ideas worth spreading”, operował sloganem „raw data now!”. Odniósł się do konieczności przede wszystkim udostępniania zbiorów danych zgodnie z założeniami koncepcji Linked Data. Problem tkwi w tym, że do końca nie jesteśmy w stanie określić wszystkich możliwych sposobów wykorzystania takiego modelu danych w odniesieniu do naszego środowiska. Duże zainteresowanie chmurą danych powiązanych z różnych sfer rejestrowania dorobku

intelektualnego i dziedzictwa kulturowego spowodowało pierwszą falę krytyki oraz wskazanie jej niedostatków. Pierwszy problem to brak skodyfikowanej wiedzy typu know-how, a więc wszelkiego rodzaju poradników, instrukcji i metodyki tworzenia zasobów w ramach koncepcji Linked Data. Cały czas poruszamy się na poziomie przygotowania i publikacji danych, nie odnosząc się do konieczności zaprojektowania i wdrożenia mechanizmów ich przetwarzania i pozyskiwania informacji. Nie ma ostatecznych rozstrzygnięć dotyczących długoterminowej dostępności do zbiorów danych oraz stabilności mechanizmów odsyłania (URI). Równie istotny problem to wiarygodność deklaracji oraz mechanizmów jej weryfikacji, czyli problem fałszywych metadanych i tego konsekwencji. Z punktu widzenia ekspresji i reprezentacji danych pojawił się także problem opisu całych kolekcji danych, czyli metadanych na poziomie kolekcji (ang. *collection level description*) oraz braku mechanizmów weryfikacji nieścisłości w opisie tego samego atrybutu dla dokumentu za pomocą różnych narzędzi (przy jednoczesnym wykorzystaniu zasady tożsamości – owl:sameAs).

Należy pamiętać, że Semantic Web to wizja, której jedną z materializacji jest chmura danych powiązanych. Koncepcja Linked Data dotyczy metod ekspresji, reprezentacji, łączenia i współdzielenia danych w semantycznej sieci, które wykorzystują istniejące standardy i narzędzia sieciowe. Stabilność modeli opisu uniwersum bibliograficznego i narzędzi wykorzystywanych w bibliotekach do tego celu pozwala mieć nadzieję, że będziemy aktywnymi uczestnikami tworzenia zawartości Semantic Web, o ile znajdą się chętni i pieniądze na sfinansowanie tego typu inicjatyw.

Bibliografia

1. Allemang D., Hendler J.: *Semantic web for the working ontologist modeling in RDF, RDFS and OWL*. Burlington 2007, s. 13.
2. Baker T., Bermes E.: *Library Linked Data Incubator Group Charter*. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://www.w3.org/2005/Incubator/lld/charter>>.
3. Bennett R., Lavoie B., O'Neill E.: *The Concept of a Work in WorldCat: An Application of FRBR*. „Library Collections, Acquisitions, and Technical Services” 2003, vol. 27, no. 1, pp. 45-59.
4. Berners-Lee T.: *AI. Uniform Resource Identifier (URI): Generic Syntax*. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://tools.ietf.org/html/rfc3986>>.
5. Berners-Lee T.: *Hearing on the Digital Future of the United States: Part I – The Future of the World Wide Web*. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://dig.csail.mit.edu/2007/03/01-ushouse-future-of-the-web.html>>.
6. Berners-Lee T.: *Linked Data – Design Issues*. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://www.w3.org/DesignIssues/LinkedData.html>>.
7. Berners-Lee T.: *On the next Web*. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <http://www.ted.com/talks/tim_berniers_lee_on_the_next_web.html>.
8. Bizer C., Cyganiak R., Heath T.: *How to Publish Linked Data on the Web*. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://www4.wiwiss.fu-berlin.de/bizer/pub/LinkedDataTutorial/>>.
9. Bizer Ch.: *DBpedia 3.5 released*. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://blog.dbpedia.org/2010/04/12/dbpedia-35-released/>>.
10. Bizer Ch., Heath T., Berners-Lee T.: *Linked Data – The Story So Far*. „International Journal on Semantic Web and Information Systems” 2009, no. 3, pp. 1-22.
11. Campbell L., MacNeill S.: *The Semantic Web, Linked and Open Data*. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <http://wiki.cetis.ac.uk/images/1/1a/The_Semantic_Web.pdf>.

12. Clunker H.: *FRBR: Functional Requirements for Bibliographic Records Application of the Entity-Relationship Model*. „Library Resources & Technical Services” 2002, vol. 46, no. 4. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <http://www.oclc.org/research/publications/archive/2002/oneill_frbr22.pdf>.
13. Coyle K.: *Library Data in a Modern Context*. „Library Technology Reports” 2010, no. 1, pp. 5-13.
14. Coyle K.: *Metadata models for the World Wide Web*. „Library Technology Report” 2010, no. 46, pp. 12-19.
15. Cyganiak R.: *About the Linking Open Data dataset cloud*. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://richard.cyganiak.de/2007/10/lod/>>.
16. *Datasets in the next LOD Cloud*. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://www4.wiwiw.fu-berlin.de/lodcloud/>>.
17. Davis I., Newman R.: *Expression of Core FRBR Concepts in RDF*. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://vocab.org/frbr/core.html>>.
18. Ding L., Finin T., Shinavier J., McGuinness D.: *owl:sameAs and Linked Data: An Empirical Study*. W: *Proceedings of the WebSci10: Extending the Frontiers of Society Online*, Raleigh. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://journal.webscience.org/403/>>.
19. Dodds L., Davis I.: *Linked Data Patterns. A pattern catalogue for modelling, publishing, and consuming Linked Data*. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://patterns.dataincubator.org/book/linked-data-patterns.pdf>>.
20. *Dublin Core Metadata Element Set, Version 1.1*. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://dublincore.org/documents/dces/>>.
21. Dumbill E.: *Description of a Project*. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://trac.usefulinc.com/doap>>.
22. Fay R.: *Linked: Semantic Web & metadata*. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://www.slideshare.net/robinfay/semantic-web-cataloging-metadata>>.
23. *FOAF Vocabulary Specification 0.98*. Red. D. Brickley, L. Miller. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://xmlns.com/foaf/spec/>>.
24. *Functional Requirements for Authority Data*. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://www.ifla.org/publications/functional-requirements-for-authority-data>>.
25. *Functional Requirements for Subject Authority Data (FRSAD) A Conceptual Model*. Red. M. Zeng, M. Žumer, A. Salaba. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://nkos.slis.kent.edu/FRSAR/FRSAD-Report.pdf>>.
26. Herman I.: *Web Ontology Language (OWL)*. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://www.w3.org/2004/OWL/>>.
27. *Music Ontology Specification*. Red. Y. Raimond, F. Giasson. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://musicontology.com/>>.
28. Padziński A.: *Wymagania funkcjonalne dotyczące rekordów bibliograficznych – FRBR: możliwości zastosowania w katalogach bibliotecznych*. „Przegląd Biblioteczny” 2004, z. 3/4, s. 173-194.
29. *RDA: Resource Description and Access*. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://www.rda-jsc.org/rda.html>>.
30. *RDF Vocabulary Description Language 1.0: RDF Schema*. Pod red. D. Brickley, V. N. Guha. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://www.w3.org/TR/rdf-schema/>>.
31. Roszkowski M.: *Simple Knowledge Organization System (SKOS) – reprezentacja wiedzy w sieciowych systemach organizacji wiedzy*. „Zagadnienia Informatyki Naukowej” 2009, nr 1(93), s. 89-102.
32. Segaran T., Evans C., Taylor J.: *Programming the Semantic Web*. Sebastopol 2009.
33. *Semantic Digital Libraries*. Pod red. S. Kruk, B. McDaniel. Berlin, Heidelberg 2009, s. 114-119.
34. Sequeda J.: *Consuming Linked Data*. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://www.slideshare.net/juansequeda/consuming-linked-data-semtech2010>>.

35. Shakya A., Takeda H.: *A Report on Linked Data*. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://sigsw.org/papers/SIG-SWO-A801/SIG-SWO-A801-07.pdf>>.
36. *SIOC Core Ontology Specification*. Pod red. U. Bojārs, J. Breslin. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://rdfs.org/sioc/spec/>>.
37. *SKOS Simple Knowledge Organization System Primer*. Pod red. A. Isaac, E. Summers. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://www.w3.org/TR/2009/NOTE-skos-primer-20090818/>>.
38. *SKOS/Datasets*. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://www.w3.org/2001/sw/wiki/SKOS/Datasets>>.
39. *The DBLP Computer Science Bibliography*. Pod red. M. Ley. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://www.informatik.uni-trier.de/~ley/db/>>.
40. *Understanding the Semantic Web: Bibliographic Data and Metadata*. „Library Technology Reports”. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://www.alatechsource.org/library-technology-reports/understanding-the-semantic-web-bibliographic-data-and-metadata>>.
41. Vatant B.: *GeoNames Ontology*. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://www.geonames.org/ontology/>>.
42. *W3C Launches Library Linked Data Incubator Group*. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <<http://www.w3.org/News/2010#entry-8803>>.
43. Wikipedia. *Semantic Web*. [online]. [dostęp: 05.11.2010]. Dostępny w World Wide Web: <http://pl.wikipedia.org/wiki/Semantic_Web>.

Summary

The paper presents foundations of Linked Data – a recommended best practice for exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web. Practical applications of this model were presented. The paper includes issues on bibliographic data on Semantic Web, major problems and possible advantages for library community.