



**INSTYTUT BADAŃ SYSTEMOWYCH
POLSKIEJ AKADEMII NAUK**

**ANALIZA SYSTEMOWA W FINANSACH
I ZARZĄDZANIU**

Wybrane problemy
Tom 4

Pod redakcją
Jerzego HOŁUBCA

Warszawa 2002



**INSTYTUT BADAŃ SYSTEMOWYCH
POLSKIEJ AKADEMII NAUK**

**ANALIZA SYSTEMOWA W FINANSACH
I ZARZĄDZANIU**

**Wybrane problemy
Tom 4**

**Pod redakcją
Jerzego HOŁUBCA**

Warszawa 2002

Wykaz opiniodawców artykułów zamieszczonych w tomie:

doc. dr hab. Mieczysław KŁOPOTEK

prof. dr hab. Stanisław PIASECKI

prof. dr Elżbieta RAKUS-ANDERSON

prof. dr hab. Andrzej STRASZAK

doc. dr hab. Sławomir WIERZCHOŃ

dr Sławomir ZADROŻNY

Publikacja dofinansowana przez
Agencję Wydawniczo-Poligraficzną "ARGRAF", Warszawa

© Instytut Badań Systemowych PAN, Warszawa 2002

ISBN 83-85847-74-X

Wydawca: INSTYTUT BADAŃ SYSTEMOWYCH PAN
ul. Nowelska 6 01-447 Warszawa

Redakcja: Dział Informacji Naukowej i Wydawnictw

Barbara Katuszewska, Joanna Runowska, tel. 837-68-22

Druk: Agencja Wydawniczo-Poligraficzna "ARGRAF", Warszawa

Nakład 200 egz., 15 ark.wyd.; 12,8 .ark. druk.

SIEĆ SEMANTYCZNA - NOWA GENERACJA GLOBALNEJ SIECI INFORMACJI

Klaudia A. Ławcewicz

Zaoczne Studia Doktoranckie IBS PAN

Współczesna sieć WWW stanowi ważne źródło informacji dla wielu ludzi. Nie jest to narzędzie doskonałe – dane są rozproszone, nieuporządkowane i trudne w ocenie ich wiarygodności. Rozwiązaniem może stać się sieć semantyczna – nowa generacja sieci umożliwiająca automatyczne rozpoznawanie, przeszukiwanie i analizowanie informacji. Niniejszy artykuł ma na celu przejrzeć współczesne osiągnięcia w konstruowaniu sieci semantycznej i stosowane w tym zakresie technologie.

Słowa kluczowe: sieć semantyczna, Internet, ontologia, formalizm opisu zasobów (RDF), rozszerzony język znaczników (XML).

Abstract

Current World Wide Web is an important source of information for many people. It is not perfect – data are dispersed, unordered and difficult to assess its reliability. Semantic web, new generation of network, which allows automatic recognition, search and analysis of information, may be a possible solution of this problem. The aim of this article is to review current achievements and employed technologies in this area.

Keywords: semantic web, Internet, ontology, Resource Description Framework (RDF), eXtensible Markup Language (XML).

1. Wprowadzenie

Sens istnienia ogromnej, powszechnie dostępnej bazy wiedzy, jaką jest Internet, we współczesnym świecie jest bezsporny. Równie oczywista jest mnogość wad i ułomności, z jakimi boryka się współczesna sieć. Jednym z podstawowych jej mankamentów jest chaotyczność, rozumiana tu jako brak jakiegokolwiek uporządkowania tematycznego poszczególnych dokumentów, ich nierównomierność jakościową, brak kontroli nad rzetelnością i aktualnością przekazywanych treści. Dzieje się tak, ponieważ współczesna sieć WWW to zbiór informacji tworzonych przez niezliczone organizacje, społeczności i indywidualnych użytkowników. W związku z

tym bardzo trudno tak sformułować warunki przeszukiwania sieci, aby obok kilku interesujących nas informacji nie otrzymać równocześnie wielkiej ilości stron zupełnie tematycznie z naszymi poszukiwaniami niezwiązanych. Jednocześnie ilość informacji zawartych w sieci uniemożliwia takie przeszukanie jej zasobów, aby mieć przynajmniej wrażenie ogarnięcia całokształtu zamieszczonej wiedzy na dany temat.

Wydaje się, więc, że nie ma innego sposobu na poradzenie sobie ze zgromadzoną w Internecie wiedzą, niż oddanie zadania skatalogowania jej „w ręce” inteligentnych agentów. Aby jednak było to możliwe, zawartość sieci musi stać się dla nich zrozumiała. Zagadnieniem tym zajmuje się twórca sieci pierwszej generacji – *World Wide Web (WWW)* – Tim Berners-Lee wraz z zespołem naukowców związanych z kierowanym przez niego *World Wide Web Consortium (W3C)*. Prace trwają od 1994 roku. W 1998 roku przedstawili oni koncepcję sieci nowej generacji – *Sieci Semantycznej*¹, która byłaby w stanie reprezentować znaczenie przechowywanych w niej treści oraz logicznych relacji pomiędzy nimi². Miałaby ona być oparta na standardach i technologiach ułatwiających inteligentnym agentom analizę semantyki i jakości danych dostępnych w sieci. Przykładowo, oprogramowanie przeszukujące sieć w celu znalezienia autorki tego artykułu, automatycznie odrzuci wszystkie dokumenty zawierające słowo „Kludia”, które w danej treści wskazują na nazwę firmy lub magazynu kobiecego. Umożliwi to sprecyzowanie znaczeniowe wyszukiwanego słowa kluczowego: szukamy człowieka, nie organizacji, doprecyzowując szukamy kogoś, kto spełnia kombinacje wymagań, takich jak: <people>, <author>, <first_name>.

2. Architektura sieci semantycznej

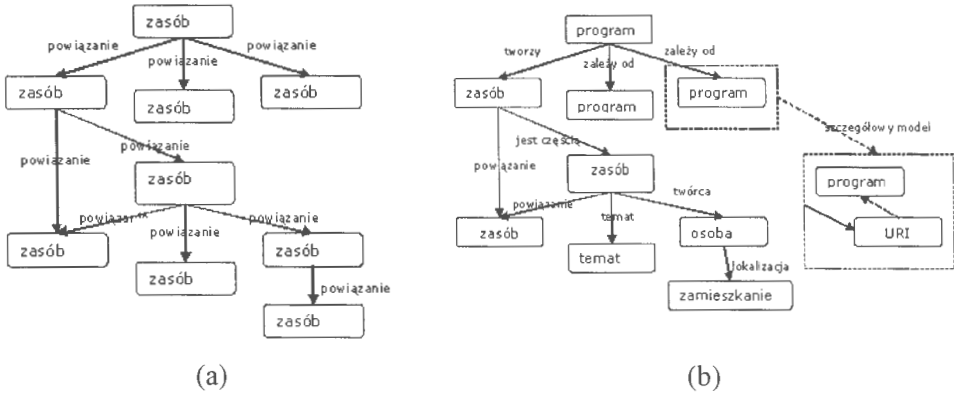
Najważniejszymi funkcjami sieci semantycznej, według autorów koncepcji, mają być: integracja bogatych znaczeniowo danych z różnych źródeł, porównywanie ich, zapewnienie możliwości inteligentnego poruszania się pomiędzy nimi oraz automatyzacja tych czynności na poziomie mechanizmu przeszukującego. Zadania, jakie stoją przed siecią semantyczną daje się podzielić na dwie podstawowe grupy:

¹ Termin “sieć semantyczna” jest tu tłumaczeniem angielskiego określenia “Semantic Web”. Pokrywa się on z wcześniej stosowanym w literaturze polskiej tłumaczeniem ogólniejszego pojęcia „semantic network” określającego metodę reprezentacji wiedzy.

² Berners-Lee T., Semantic Web Road map, <http://www.w3.org/DesignIssues/Semantic.html>

✓ stworzenie takiego modelu adresowania stron, który pozwoli na zamieszczanie w nim kluczowych informacji dotyczących zawartości dokumentu - w sieci semantycznej adresy URI nie ograniczałyby się do wskazywania na konkretne dokumenty, ale na ludzi, idee i powiązania pomiędzy nimi,

✓ zaprojektowanie takiej struktury źródeł i odnośników, która pozwoli na umieszczanie ich w szkieletcie wcześniej zdefiniowanych typów („autor”, „organizacja”, „adres”) powiązanych odpowiednimi relacjami („zależy od”, „ma tytuł”, „jest autorem”).



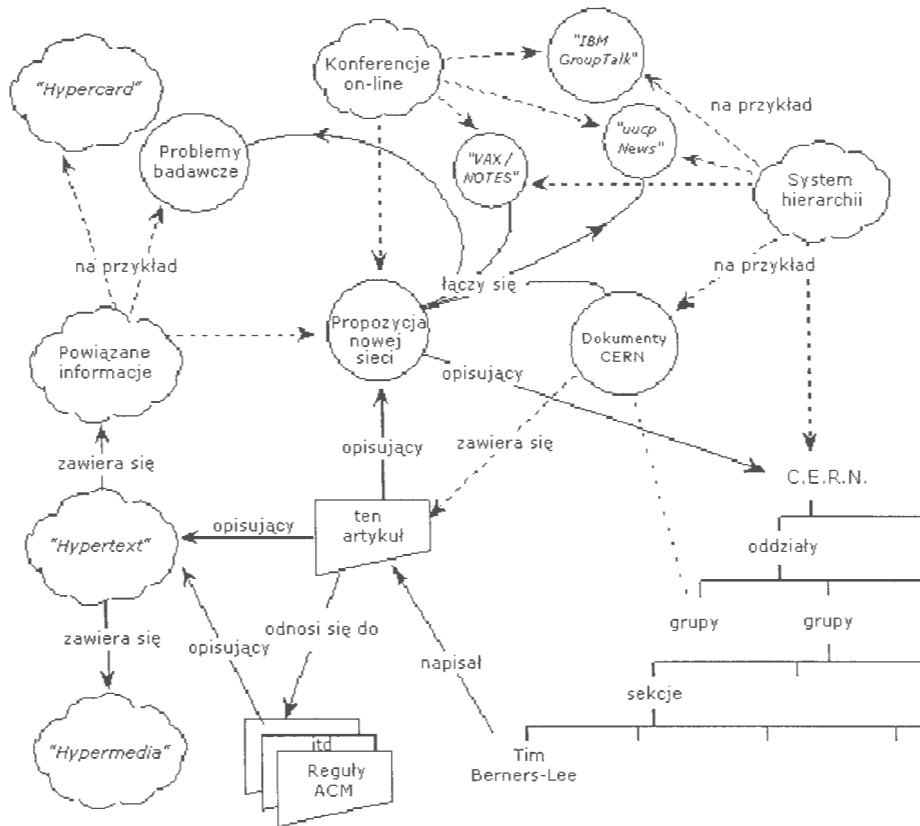
Rys. 1. Internet: sieć współczesna (a), sieć semantyczna (b)³

Powyższy diagram przedstawia podstawowe różnice pomiędzy reprezentacją danych w sieci współczesnej i przyszłej sieci semantycznej. Na diagramie 1.b widoczne są powiązania pomiędzy poszczególnymi obiektami, takie relacje dotyczyć mają wnętrza danego dokumentu: „<autor>” □ „,jest twórcą” □ „, <artykuł>” □ „,zatytułowany” □ „, <tytuł artykułu>”, jak i powiązań pomiędzy różnymi dokumentami całej sieci: „, <artykuł>” □ „, należy do” □ „, <zbiór dokumentów PAN>” □ „, zawiera się w ” □ „, <zbiór artykułów sieci semantycznej>”.

Na rysunku 2 przedstawiony jest pierwszy model sieci semantycznej autorstwa Tima Berners-Lee. Obrazuje on powiązania pomiędzy poszczególnymi obiektami sieci - zasobami danych (bazami wiedzy różnych instytucji, pojedynczymi dokumentami), ludźmi uczestniczącymi w różnorodnych formach interaktywnego uczestnictwa w sieci (grupy dyskusyjne, fora, konferencje on-line) oraz relacjami pomiędzy nimi zbudowanymi przy dwóch założeniach:

³ źródło własne

- relacje powinny być, jak bardzo tylko jest to możliwe, zbliżone do relacji, które występują w świecie rzeczywistym,
- relacje powinny być w jak największym stopniu dostosowane do schematów standardów zawartych w słownikach sieci, tak aby można było dowolnie łączyć najróżniejsze zasoby.



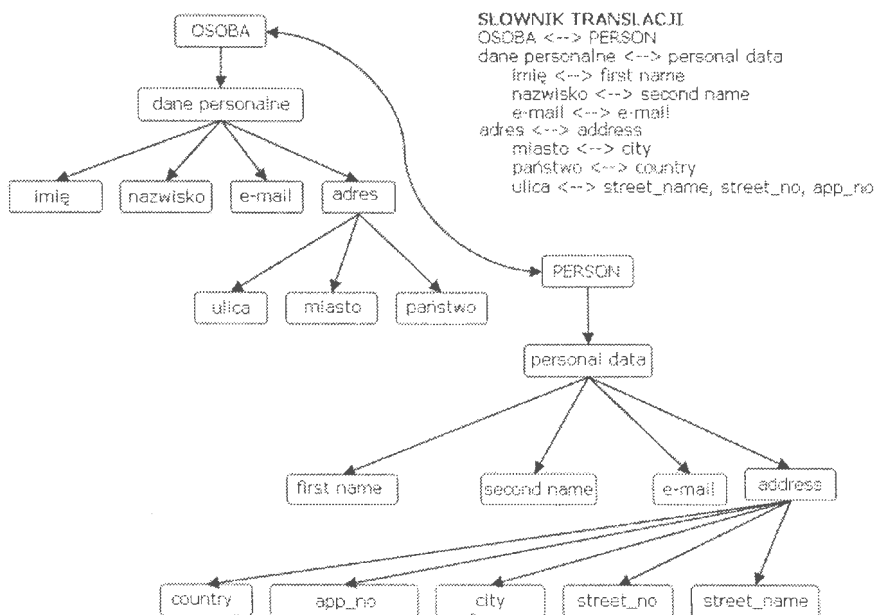
Rys.2. Wizja sieci przyszłości przedstawiona przez Tima Bernersa-Lee w 1998 roku.⁴

Dzięki takiej architekturze sieć semantyczna umożliwi kumulowanie informacji z różnych źródeł bez długotrwałej pracy nad ich ujednoczeniem, nawet, jeżeli łączone informacje zakodowane są przy pomocy innego słownika i schematu. Najtrudniejszym zadaniem będzie skonstruowanie definicji translacyjnych pomiędzy schematami i uruchomienie odpowiednich

⁴ Berners-Lee T., Semantic Web Road map, <http://www.w3.org/DesignIssues/Semantic.html>, (tłumaczenie własne)

mechanizmów, które przeprowadzą cały proces automatycznie. Łączenie danych można będzie przeprowadzić na dwa sposoby:

- przy łączeniu trwałym (np. przy pochłonięciu jednej organizacji przez drugą) dane zostają przetłumaczone i zakodowane według wspólnego słownika,
- przy chwilowym złączeniu danych (np. dla przeprowadzenia analizy rynku) skonstruowane zostają odpowiednie mechanizmy „przepytujące” inteligentnie oba zasoby również przy pomocy odpowiedniego słownika definiującego pojęcia z różnych źródeł oraz relacje pomiędzy nimi. Obecnie konstruuje się języki zbliżone do SQL’a, które umożliwiają przepytwanie jednocześnie kilku źródeł zakodowanych przy pomocy różnych słowników, mając do dyspozycji słownik translacyjny.



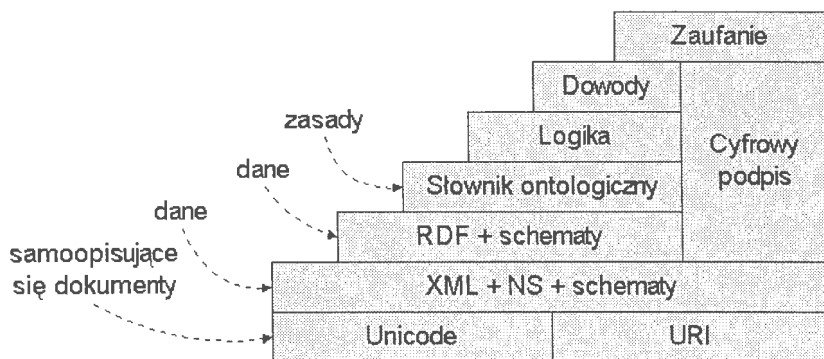
Rys.3. Przykładowy model łączenia dwóch schematów danych przez tworzenie reguł słownika translacyjnego⁵.

3. Model warstwowy sieci semantycznej

Na stronach W3C udostępniony jest model warstwowy sieci semantycznej. Przed nami jednak jeszcze zdefiniowanie i zaimplementowanie standardów przyszłej sieci oraz dokładne zdefiniowanie

⁵ źródło własne

funkcjonalności warstw, z jakich będzie się ona składać. Proponowane warstwy zaprezentowano na poniższym rysunku:



Rys.4. Warstwy sieci semantycznej⁶

Znaczenie poszczególnych warstw sieci semantycznej:

1. warstwy „Unicode” i „URI” zapewniają międzynarodowy charakter obiektów znajdujących się w sieci semantycznej przez użycie URI (*Uniform Resource Identifier*), który wskazuje zasób, a nie lokalizację, dzięki czemu m.in. można znaleźć najbliższą kopię lub przeniesiony dokument oraz poprzez wprowadzenie kodowania Unicode zawierającego litery i symbole z alfabetu łacińskiego i innych alfabetów środkowo- i zachodnioeuropejskich, cyrylicy, pisma greckiego, hebrajskiego, arabskiego, dewanagari (sanskryt), ideogramów Han, pisma chińskiego, japońskiego, sylab Hangul, koreańskiego i wielu, wielu innych,
2. warstwa „XML + NS (przestrzeń nazw) + definicje schematów” porządkuje dane poprzez usystematyzowanie ich i pogrupowanie według określonych schematów przy użyciu rozszerzalnego języka znakowania informacji XML (*eXtensible Markup Language*), warstwa ta nie określa właściwego znaczenia informacji, wskazuje natomiast, gdzie można znaleźć jej znaczenie, czyli opisujące ją hasło RDF,
3. warstwa „RDF + schematy” jest to warstwa słownikowa i synonimiczna, umożliwia nadawanie treści znaczeniowej danym zawartym w warstwie drugiej, tworzenie i grupowanie typów danych, pozwalających na określanie relacji pomiędzy nimi oraz konstruowanie wyrażeń

⁶ Koivunen M.R., Miller E., W3C Semantic Web activity, materiały konferencji: „Semantic Web Kick-off Seminar in Finland”, 2.11.2001

przeszukujących sieć przy pomocy inteligentnych agentów kumulujących dane z różnych obszarów sieci,

4. **warstwa „Słownik ontologiczny”** nie jest jeszcze doprecyzowana, w założeniach ma to być rodzaj encyklopedii wskazującej na szczegółowe relacje pomiędzy poszczególnymi obiektami sieci (szczególnie relacje pomiędzy znacznikami XML i definicjami RDF),

5. **warstwa „Logika”** pozwala na stworzenie reguł pomiędzy poszczególnymi metadanymi, dzięki którym inteligentni agenci będą mogli przeprowadzać czynności dedukcyjne podczas obsługi zapytań dotyczących zasobów bądź uprawnień, warstwa ta również nie jest jeszcze ostatecznie doprecyzowana, konieczne jest ustalenie zbioru dopuszczalnych predykatów logicznych (negacji, alternatywy, implikacji, ...) oraz kwantyfikatorów (dla wszystkich ..., istnieje taki ...), tu Tim Berners-Lee szczególnie trudność widzi w uwzględnieniu operatora negacji

6. **warstwa „Dowody”** w przyszłości – po dokładnym określeniu jej funkcjonalności – będzie przeprowadzała operacje na złączeniu wyżej opisanych warstw, przy czym warstwa ta będzie musiała spełniać dwa następujące wymagania:

6.1. agent przystosowany do analizy n dokumentów określonego typu w wersji n powinien móc „nauczyć się” analizować dokumenty w wersji n+1,

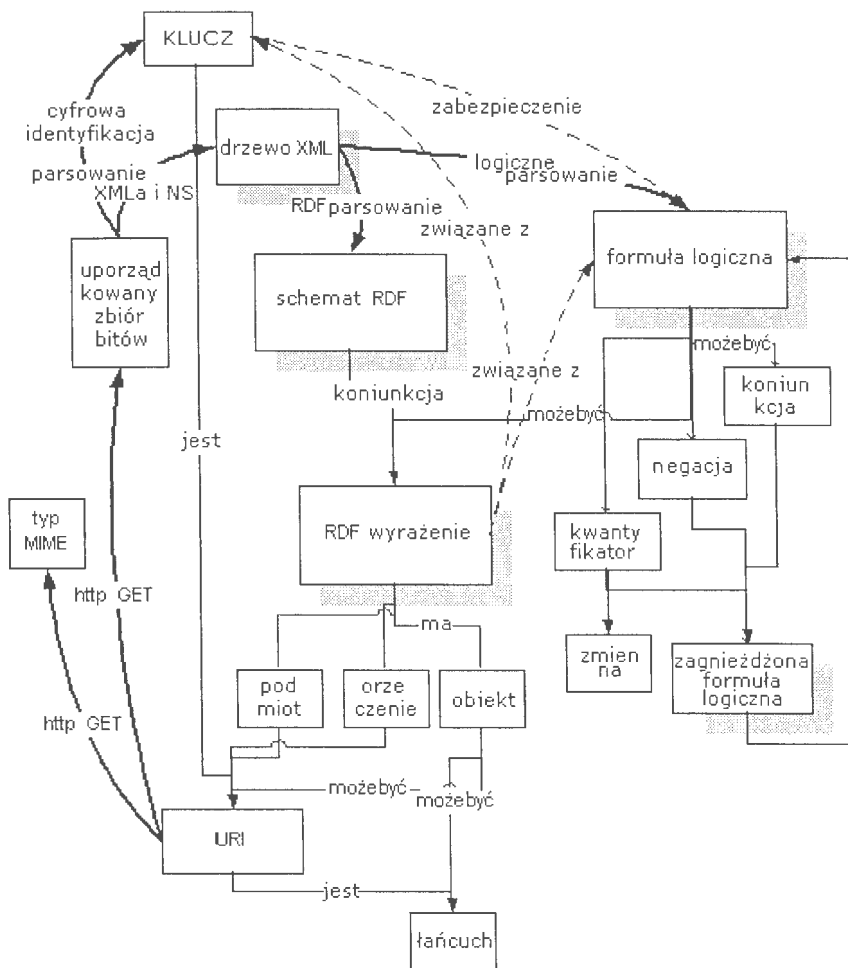
6.2. różne aplikacje o podobnej funkcjonalności powinny móc „nauczyć się” jak analizować właściwe dla każdej z nich dane,

Po spełnieniu powyższych warunków proces dowodzenia w sieci semantycznej będzie polegał na samym zaistnieniu akcji (jeśli aplikacja klienckiej udało się przesłać żądanie □ miała takie uprawnienia, jeżeli uzyskała odpowiedź □system był do tego uprawniony); ta warstwa również wymaga dopracowania,

7. **warstwa „Cyfrowy podpis”** pozwala na analizowanie (parsowanie) dokumentów nie tylko pod względem powiązań i zawartości, ale również drzew uprawnień dostępu do danych wiadomości – hierarchii cyfrowych podpisów, dzięki czemu informacje docierałyby wyłącznie do faktycznych adresatów,

8. **warstwa „Zaufanie”** umożliwi użytkownikom weryfikację dostarczanych treści; również nie jest to warstwa dopracowana.

Taki model zapewnia wielopoziomowe uporządkowanie i powiązanie zasobów wiedzy znajdującej się w Internecie, jednocześnie nie zakłócając naturalnego rozrostu, rozproszenia i powszechności. Diagram na rys. 5 przedstawia proces działania i wzajemnych przepływów w sieci semantycznej.



Rys.5. Model przepływu danych i akcji w sieci semantycznej.⁷

W kolejnych punktach omówimy niektóre aspekty wybranych warstw modelu sieci semantycznej.

⁷ Berners-Lee T., Semantic Web Road map, <http://www.w3.org/DesignIssues/Semantic.html>, (tłumaczenie własne)

4. Formalizm opisu zasobów RDF

Szerszego opisu wymaga formalizm **opisu zasobów** (*RDF – Resource Description Framework*), który dodaje semantykę do ustrukturyzowanej zawartości. RDF są to zestawy metadanych (danych opisujących dane), czyli określających zestaw zdefiniowanych elementów i dopuszczalną ich zawartość. Ramowe opisy zasobów umożliwiają standaryzację terminów (nazw znaczników). Mogą być przez siebie dziedziczone (kolejne definicje RDF mogą korzystać z poprzednio utworzonych i dowolnie je rozszerzać), co pozwala na elastyczne rozszerzanie słownika pojęć, kiedy tylko zaistnieje taka potrzeba. Fragment takiej hierarchii terminów może mieć następującą postać:

- publikacja
 - artykuł
 - artykuł z podziałem na rozdziały
 - artykuł w konkretnym miesięczniku
 - ...

Pozwala to na dowolne doprecyzowywanie struktury zamieszczanych treści jednocześnie zachowując ich pełną kompatybilność z innymi dokumentami wykorzystującymi wyższe poziomy hierarchii terminów.

Używanie technologii RDF przy opisie dokumentów umożliwia wszelkiego rodzaju agentom internetowym na wyszukiwanie precyzyjnie określonych treści. Jest to jeden z najważniejszych elementów, który nada sieci współczesnej semantyczny charakter. Nie jest to technologia istniejąca wyłącznie na papierze – powszechnie wykorzystuje się ją w wielu dziedzinach już dzisiaj, między innymi w opisach standardów informatycznych, przy tworzeniu materiałów szkoleniowych. Jednym z najbardziej popularnych RDF'ów jest Dublin Core – słownik przeznaczony do opisu danych o stronach sieciowych. Poniżej zamieszczony jest fragment opisu artykułu przy użyciu słownika Dublin Core.

```
<RDF:RDF xmlns:RDF=http://www.w3.org/1999/02/22-rdf-syntax-ns#
          xmlns:DC="http://purl.org/DC/">
<RDF:Description about="http://metalab.unc.edu/xml/">
<DC:Creator>Klaudia A. Ławcewicz</DC:Creator>
<DC:Language>pl</DC:Language>
<DC:Format>HTML</DC:Format>
<DC>Date>2002-05-04</DC>Date>
<DC>Type>artykuł</DC>Type>
<DC>Title>Opowieści_o_sieci_semantycznej </DC>Title>
</RDF:Description>
</RDF:RDF>
```

Jako że RDF powstał jako rezultat współpracy i porozumienia wielu społeczności co do podstawowych zasad reprezentacji i transportu metadanych, czerpie on inspiracje z wielu źródeł. Głównym źródłem tych inspiracji są środowiska zajmujące się standardami WWW, bibliotekoznawstwem, strukturalizacją dokumentów w formie SGML i XML, jak również w reprezentacją wiedzy. Również inne technologie, takie jak programowanie zorientowane obiektowo, języki modelowania oraz bazy danych, miały wpływ na postać RDF. Chociaż RDF czerpie inspirację z metodologii reprezentacji wiedzy nie obejmuje specyfikacji mechanizmów wnioskowania. RDF można więc scharakteryzować jako prosty system ramek. Mechanizm wnioskowania może być dobudowany ponad ramkami RDF.

RDF może znaleźć różnorodne zastosowania

- w odkrywaniu zasobów (co podniesie możliwości maszyn wyszukiwujących (ang. search engines))
- w katalogowaniu (dla opisywania zawartości i powiązań występujących na konkretnej stronie WWW, portalu, lub bibliotece elektronicznej),
- przez inteligentnych agentów wspomagających wymianę i współdzielenie wiedzy,
- w ocenie treści,
- przy opisywaniu kolekcji stron które reprezentują pojedynczy logiczny „dokument”,
- dla opisywania praw własności intelektualnej stron WWW,
- dla określania preferowanego poziomu ochrony prywatności poszczególnych użytkowników lub zasad stosowanych w tym względzie przez witryn WWW,

Połączenie formatu RDF z cyfrowym podpisem może umożliwić skuteczne zbudowanie „Sieci zaufania” dla celów gospodarki elektronicznej, współpracy i innych aplikacji.

5. Język znaczników XML

XML jest metajęzykiem, który doskonale nadaje się do opisu różnego rodzaju złożonych danych. Z punktu widzenia sieci semantycznej jest to doskonałe narzędzie do budowania schematów dla różnych kategorii tekstów, ponieważ podczas dzielenia tekstu bierze on pod uwagę rodzaj wstawianej treści, a nie sposób jej formatowania. Zawiera on w sobie technologię definicji typu dokumentu (DTD *Document Type Definition* lub *XML Schema*), dzięki czemu można tworzyć opis słownika i składni dla

konkretnych zastosowań. Z uwagi na olbrzymią możliwość manipulacji danymi zapisanymi przy użyciu XML, znalazł on już praktyczne zastosowanie w wielu firmach zajmujących się przetwarzaniem i gromadzeniem treści. Większość używanych współcześnie baz danych obsługuje ten język, dzięki czemu uzyskuje się kompatybilność różnych systemów zarządzania bazą danych i plików XML. Powstało wiele standardów powszechnie uznanych za obowiązujące, najczęściej używane to:

– **chemiczny język znaczników CML** (*Chemical Markup Language*) opisany jako słownik i składnia nauk molekularnych: chemii, krytalografii, fizyki ciała stałego, jako przykład można podać opis w tym standardzie cząsteczki wody:

```
<?xml version="1.0"?>
<CML>
  <MOL TITLE="Water">
    <ATOMS>
      <ARRAY BUILTIN="ELSYM">H O H</ARRAY>
    </ATOMS>
    <BONDS>
      <ARRAY BUILTIN="ATID1">1 2</ARRAY>
      <ARRAY BUILTIN="ATID2">2 3</ARRAY>
      <ARRAY BUILTIN="ORDER">1 1</ARRAY>
    </BONDS>
  </MOL>
</CML>
```

- **MusicalML** – obsługa zapisu nutowego,
- **MathML** (*Mathematical Markup Language*) – obsługuje równania matematyczne,
- **SMIL** (*Synchronized Multimedia Integration Language*) – język synchronizacji multimedialnych – do tworzenia prezentacji multimedialnych,
- **OSD** (*Open Software Description*) – umożliwia automatyczną aktualizację oprogramowania,
- **SVG** (*Vector Markup Language*), **VML** (*Scalable Vector Graphics*) – umożliwiają tworzenie skalowanej grafiki wektorowej,
- **VoxML** – służy do obsługi głosu,
- **OFX** – format wymiany danych finansowych

To jedynie drobny wycinek obecnie używanych możliwości technologii XML. Zastępuje ona obowiązujący współcześnie standard kodowania danych HTML bardzo opornie, jeśli pod uwagę bierzemy wyłącznie całkowite zastępowanie. Znacznie częściej wtapia się w jego

formuły prawie niezauważalnie i dane z rozszerzeniem *.html, które obecnie widzimy na stronach przez nas przeglądanych bardzo często są generowane automatycznie przy pomocy technologii opartych na metajęzykach takich jak XML.

6. Dostęp do danych i ich wiarygodność

Współczesna sieć poświęca integralność i wiarygodność powiązań na rzecz ich skalowalności. Bardzo często zamiast akceptacji żądania i przesłania informacji dostajemy jedną z dwóch odpowiedzi zwrotnych:

- brak strony (obecnie znany błąd 404),
- odmowa uzyskania dostępu do strony.

Sieć semantyczna umożliwi tworzenie różnych typów powiązań pomiędzy zasobami. Pozwoli to na uzyskanie częściowych informacji, nawet wtedy, gdy części zasobów brakuje (np. informacje o pracownikach danej firmy, gdy sama firma, jak również jej domena, od dawna już nie istnieją) lub nie do wszystkich mamy wystarczające uprawnienia.

Zadaniem sieci semantycznej jest również rozwiązanie problemu związanego z zaufaniem do otrzymywanych przez nas danych. Organizacja W3C pracuje obecnie nad zdefiniowaniem i zaprojektowaniem standardu „kluczy zaufania”, dzięki którym użytkownik mógłby zadeklarować, jakie części sieci mają być przeszukiwane, jakim autorom, bądź organizacjom pozwala czerpać dane od siebie i skąd mają być pobierane informacje dla niego.

Informacje prezentowane w sieci przyszłości, obecnie najczęściej oderwane od ich autorów i źródeł, będą mogły być przetwarzane przez filtry informacyjne, które pozwolą określać stopień wiarygodności elektronicznych wypowiedzi w zależności od tego:

- kto mówi (czyj autorytet stoi za daną wypowiedzią),
- kiedy (czy jest to wiadomość aktualna),
- dokładnie, co mówi (czy dana informacja nie została w jakiś sposób zniekształcona)

Każdy użytkownik sieci semantycznej będzie mógł dowolnie zwiększać lub zmniejszać poziom wiarygodności dla danego autora, strony, wydawcy, etc. Pojawi się również możliwość dziedziczenia zaufania – powstaną „łańcuchy zaufania” definiujące źródła skąd czerpiemy i gdzie wysyłamy informacje. Konieczność zaistnienia takich ograniczeń jest zauważalna dziś na wszelkiego rodzaju interaktywnych miejscach sieci (listach dyskusyjnych, forach, ...), gdzie umożliwiłyby one ograniczenie

dostępu osobom zachowującym się w sposób niezgodny z ogólnie przyjętą etykietą. Tego typu podejście do praw dostępu uprości również bardzo skomplikowane współcześnie systemy autoryzacji oraz pozwoli uczynić je zależnymi od wielu czynników, co znacznie zwiększy ich wiarygodność.

7. Przyszłość koncepcji sieci semantycznej

Wydaje się, że sieć semantyczna, pomimo ewidentnych, rewolucyjnych zalet, nie zdobywa oczekiwanej popularności. Jej twórca i główny promotor Tim Berners-Lee wydaje się być wystarczająco dobrą wizytówką nowej technologii, jednak pomimo jego entuzjastycznych wypowiedzi, nie sposób porównać rozwoju tej nowej technologii z tajfunem, jaki ogarnął świat przy wdrażaniu jego poprzedniego wynalazku – sieci WWW. Powody mało entuzjastycznego przyjęcia sieci nowej generacji są prozaiczne. Technologia ta wydaje się przeciętnym użytkownikom komputerów bardziej skomplikowana. Można jednak wyobrazić sobie najpopularniejsze edytory tekstu, które zamiast przysłowiowej czystej białej kartki pytają o uporządkowane dane: autora tekstu, temat, słowa kluczowe. Następnie akapitami lub rozdziałami umożliwiają wprowadzanie dokumentu do pamięci. Tak zakodowaną treść można byłoby w bardzo prosty sposób przenosić do sieci semantycznej, jednak jest to zupełnie nowe podejście do tworzenia dokumentów tekstowych i dlatego jego zaakceptowanie wymaga po prostu czasu. Mimo to każdy, kto miał z tą technologią styczność uważa, że sieć semantyczna oferuje idealne środowisko do pracy i co ważniejsze – do współpracy.

Zaprojektowanie tak złożonego organizmu wymaga jednak wiele czasu, dlatego też prace nad zdefiniowaniem poszczególnych warstw sieci semantycznej przebiegają bardzo powoli. Obecnie priorytetem organizacji W3C są prace:

- grupy opracowującej technologię RDF, w założeniu ma ona reprezentować metadane jako standard akceptowany przez większość aplikacji, zadaniem grupy jest stworzenie ogólnego, koncepcyjnego modelu schematów RDF uwzględniających składnie XML'a oraz słowników opisujących relację pomiędzy nimi,
- grupy pracującej nad skonstruowaniem ontologicznych zasad funkcjonowania sieci, ma ona być nakładką na warstwę RDF tworzącą klasy i podklasy obiektów, charakteryzująca ich właściwości i hierarchie oraz dziedziczenie pomiędzy nimi.

Trwają również prace nad:

- narzędziami konwertującymi obecną sieć,

- narzędziami opisu źródeł,
- współpracą i zaufaniem w sieci,
- prawami dostępu do źródeł,
- monitorowaniem konstruowania logicznych reguł sieci,
- współdziałaniem inteligentnych agentów i korelacjami pomiędzy nimi,
- integracją struktur XML.

Sieć semantyczna tym różniłaby się od sieci współczesnej, czym zbiór plików tekstowych różni się od bazy danych. Nietrudno jest sobie wyobrazić kontrast pomiędzy możliwościami przeszukiwania i analizowania tak różnych dwóch tworów. Osobiście uważam, że niedocenianym przez wielu ludzi problemem może okazać się konwersja współczesnej sieci. Problem skonstruowania inteligentnych agentów przeznaczonych do automatycznego przekształcania tradycyjnych stron WWW w zbiór powiązanych danych nowej generacji może okazać się wyzwaniem niebanalnym.

Zastosowanie sieci semantycznej zmieni w tym samym stopniu życie wielkich korporacji, co pojedynczych ludzi. Przy jej pomocy automatycznie będziemy znajdować najlepszą z interesujących nas ofert i organizować nasz czas w optymalny sposób – sieć semantyczna umożliwi błyskawiczną ocenę ogromnej liczby kombinacji. Rozwój nowej sieci semantycznej, po przejściu punktu krytycznego w niedalekiej przyszłości, zmieni cały świat, tak jak obecna sieć WWW zmieniła go kilkanaście lat temu.

Literatura

- Anderson Dawid. P., Kubiawicz J., *Ogólnoświatowy komputer*, Świat Nauki, maj 2002
- Berners-Lee T., *Semantic Web Road map*, <http://www.w3.org/DesignIssues/Semantic.html>, wrzesień 1998
- Berners-Lee Jamek., Hendler J., Lassila O., *Sieć semantyczna*, Świat Nauki, lipiec 2001
- Herold E.R., *XML. Księga eksperta*, Gliwice 2000
- Koivunen M.R., Miller E., *W3C Semantic Web activity*, materiały konferencji: „Semantic Web Kick-off Seminar in Finland”, listopad 2001
- Masadyński M., *Światowa Sieć jako Super Mózg*, http://neur.am.put.poznan.pl/cempel/pr_kon/praca.html, Poznań 2000

Port O., *Wielka sieć pomysły za ciebie*, BusinessWeek, kwiecień 2002

W3C, *Requirements for a Web Ontology Language*, <http://www.w3.org/TR/2002/WD-webont-req-20020307/>, marzec 2002

W3C, *RDF Model Theory*, <http://www.w3.org/TR/2002/WD-rdf-mt-20020429/>, kwiecień 2002

W3C, *RDF/XML Syntax Specification (Revised)*, <http://www.w3.org/TR/2002/WD-rdf-syntax-grammar-20020325/>, marzec 2002

ISBN 83-85847-74-X

)