

mgr inż. Adam Piasecki
mgr inż. Michał Socha
Instytut Technik Innowacyjnych EMAG

MOŻLIWOŚCI PRAKTYCZNEGO ZASTOSOWANIA SIECI SEMANTYCZNYCH W PRZEMYŚLE

Sieci semantyczne skupiają coraz większą uwagę zarówno środowisk naukowych, jak i producentów komercyjnych rozwiązań IT. Związane z nimi prace badawcze koncentrują się obecnie na poszukiwaniu właściwych rozwiązań, standaryzacji i konstruowaniu metodologii ich wykorzystania. Jednocześnie w nowym podejściu do zagadnień zarządzania przedsiębiorstwem zwraca się szczególną uwagę na dynamicznie zachodzące zmiany w otoczeniu działania przedsiębiorstwa i konieczną modyfikację stosowanych do tej pory schematów działania. Zapewnienie wydajnej współpracy komórek organizacyjnych przedsiębiorstw wymaga, między innymi wykorzystania nowoczesnych technologii nadążających za dynamiką tych zmian. W artykule została zaprezentowana idea wykorzystania sieci semantycznych jako uzupełnienia dotychczas istniejących rozwiązań IT, głównie w zakresie integracji systemów informatycznych. Wskazane zostały również przedsięwzięcia mające na celu praktyczne zastosowanie sieci semantycznych w przemyśle.

POSSIBLE PRACTICAL APPLICATIONS OF SEMANTIC WEB IN THE INDUSTRY

Semantic web is point of interest both scientific researchers and producers of IT solutions. Research area is focused on finding appropriate solution, standardization and methodology of semantics use. Meanwhile, a new approach to business management issues, requires the use of modern technology that will follow the dynamic changes of the enterprise environment. This article is presenting the idea of using semantic web technology as a support already existing IT solutions, mainly in the area of the IT systems integration, as well as some applications of semantic web in the industry.

1. WPROWADZENIE

Obecnie działalność wszelkiego rodzaju przedsiębiorstw bazuje na silnym wykorzystaniu narzędzi informatycznych, a zakres ich wykorzystania nieustannie się poszerza. Zaobserwować można, że obowiązującym trendem jest wdrażanie systemów informatycznych wyższego rzędu, których działanie opiera się na wykorzystaniu danych wytworzonych przez podstawowe systemy już funkcjonujące w przedsiębiorstwie. Systemy informatyczne coraz częściej występują, jako wsparcie procesów realizowanych w przedsiębiorstwie. W przypadku wspierania działalności człowieka, zwłaszcza w podejmowaniu decyzji, jest to znacznie trudniejsze zadanie niż kontrolowanie przebiegu procesu technologicznego.

W nowym podejściu do zagadnień zarządzania przedsiębiorstwem zwraca się szczególną uwagę na zmiany zachodzące w działaniu i w otoczeniu przedsiębiorstwa. Zmienność warunków wymusza zmianę schematów działania przedsiębiorstw, a przed wykorzystywanymi narzędziami informatycznymi wspierającymi ich działanie stawia trudne

wyzwanie sprostania dynamice tych zmian. Systemy informatyczne wspierające kadre zarządzającą powinny charakteryzować się wieloma funkcjami, których połączenie w ramach jednego systemu jest zadaniem skomplikowanym. Jednym z najtrudniejszych elementów jest integracja heterogenicznego środowiska informatycznego.

Z punktu widzenia zarządzania, integrowanie systemów informatycznych polega na dążeniu do kompleksowego i całościowego ujęcia procesów gospodarczych i biznesowych przedsiębiorstwa oraz jego otoczenia [11]. Systemy informatyczne są zintegrowane, jeżeli możliwy jest między nimi odpowiedni przepływ danych i sygnałów sterujących. Zostanie to dokonane, jeśli spełnione zostaną warunki spójności syntaktycznej, semantycznej i technicznej [14].

Przyczyny, dla których w przedsiębiorstwach integrowane są systemy informatyczne, są dwójakiego rodzaju. Pierwsza przyczyna to potrzeba wymiany komunikatów na poziomie systemów podstawowych. Druga przyczyna to chęć osiągnięcia, dzięki integracji, systemu informacyjnego dostarczającego wiedzy o tej części przedsiębiorstwa, która jest objęta zintegrowanymi systemami podstawowymi. Integracja, której celem jest zapewnienie łączności systemów podstawowych, jest niczym innym jak zarządzaniem procesami technologicznymi. Natomiast stworzenie systemu informacyjnego to podstawa dla systemu zarządzania, który może powstać dopiero na podbudowie w postaci zintegrowanych systemów podstawowych.

Rozwój systemów informatycznych, rozpatrując zaawansowanie technologiczne i jakość wykonania, pozwala budować działające zintegrowane systemy informatyczne wraz ze wsparciem zarządzania, oparte na wspólnej dla wszystkich systemów przestrzeni wymiany komunikatów. Rosnąca liczba praktycznych i działających zastosowań pozwala myśleć o kolejnym etapie rozwoju integracji. Zwiększenie elastyczności i otwartości na kolejne systemy dzięki opisowi semantycznemu komunikatów i informacji, które przepływają przez wspólną przestrzeń danych, to cel, który powinien przyświecać rozwojowi integracji.

Dotychczasowe rozwiązania, osobno wspomagają zarządzanie finansami czy mieniem, osobno – zarządzanie poszczególnymi systemami ewidencyjnymi, nawet w ramach jednej organizacji. Nowoczesne rozwiązania informatyczne na poziomie organizacji ewoluują, przechodząc od klasy ewidencji zdarzeń gospodarczych w kierunku controllingu i rachunkowości, umożliwiającej strategiczne planowanie, kontrole finansów firmy, zarządzanie dystrybucją i personelem, planowanie produkcji, prowadzenie sprawozdawczości dla zarządów i instytucji zewnętrznych [14]. Dzięki użytkowaniu zintegrowanych systemów zarządzania uzyskuje się znaczące korzyści:

- wyższe jakościowo wspomaganie procesów decyzyjnych,
- wieloprzekrojowe i na różnych poziomach zarządzania informowanie kierownictwa przedsiębiorstwa,
- objęcie kanałami informacyjnymi wszystkich kluczowych agend przedsiębiorstwa m.in. finanse, logistyka, produkcja itd.,
- podniesienie wiarygodności informacji dzięki wprowadzeniu ich do systemu w miejscu powstawania,
- kontrolowanie kosztów w oparciu o ich pomiar w miejscu powstawania,
- symulowanie budżetowe i analizy finansowe, prognozy zarówno krótko jak i długoterminowe,
- zarządzanie strumieniem materiałów, surowców, półproduktów i usług w ramach wykonywanych prac przedsiębiorstwa,

- mechanizmy zapewniające bezpieczeństwo zasobów danych [14].

Bardzo ważnym elementem w poprawnym działaniu systemu jest efektywne wykorzystanie wszystkich przydatnych m. od. na potrzeby organizacji. Trzeba pamiętać, że moduły każdego systemu są ściśle ze sobą powiązane. Wprowadzenie niepoprawnej informacji w jednym module spowoduje jej błędne przetworzenie w drugim i zniekształcenie informacji końcowej [14].

Praktyka wykazała, że fizyczna integracja systemów informatycznych, polegająca na komunikacji między systemami informatycznymi, wymianie danych za pomocą interfejsów, sieci komputerowych oraz protokołów komunikacyjnych, pozwoliła na osiągnięcie tylko ograniczonego poziomu scalenia. W wyższy poziom zapewnia integracja aplikacji, która dotyczy wspólnego użytkowania danych przez odmienne aplikacje [11].

W działaniach podejmowanych przez autorów niniejszego opracowania szczególne uznanie znalazła definicja integracji określana jako działanie, którego istotą jest utworzenie nowej całości, której elementy połączone są określonymi relacjami i powiązane odpowiednim stopniem zależności od całości [10]. Celem takim rozumianej integracji może być również połączenie w procesie działania, różnych elementów w większą całość, aby w ten sposób uczynić to działanie racjonalnym i optymalnym.

2. SIECI SEMANTYCZNE

Upowszechnienie Internetu, jako medium umożliwiającego kontakt, wymianę oraz wyszukiwanie informacji, jest obecnie jednym z głównych czynników mających wpływ na szybki rozwój gospodarczy, doskonalenie usług, rozwój nowych technologii i wymianę doświadczeń w środowiskach naukowych i biznesowych. W dobie zalewu informacji, mających swoje źródło nie tylko w samej organizacji, ale również poza nią, szczególne zainteresowanie budzi aspekt integracji systemów informatycznych [11] obejmujący jednolitą definicję danych, będących przedmiotem wymiany w ramach organizacji. Rozwiązaniem powiązanych z nim problemów może być zastosowanie sieci semantycznych, których koncepcja ma swoje źródło w rozwiązaniach odnoszących się do globalnej sieci. Jak pokazują dotychczasowe doświadczenia, przejście od koncepcji globalnej do poziomu lokalnego (tu: organizacji) nie nastręcza większych trudności, jednocześnie zastosowanie tej samej koncepcji lokalnie pozwoli na prostą integrację ze strukturami globalnymi.

Standardy wykorzystywane przez sieć globalną w początkowym okresie jej rozwoju zakładały przede wszystkim możliwość udostępniania dokumentów i plików danych. W takim podejściu o znaczeniu przesyłanych danych decydują ich nadawca (intencja) i odbiorca (interpretacja). Niestety przy s. albo rozwiniętych kanałach komunikacyjnych łączących nadawcę z odbiorcą trudno jest przekazać znaczenie udostępnionych danych. Problem ten urasta do rangi bariery, jeśli udostępnianie danych i dokumentów rozważać będziemy w kontekście anonimowego odbiorcy, który do danych będzie dołączał własną interpretacją oderwaną od intencji nadawcy. (nadawca nie komunikuje się z odbiorcą). Nie dziwi, więc fakt, że w krótkim czasie pojawiły się nowe potrzeby i wymagania, którym dotychczasowe standardy i rozwiązania nie mogły sprostać. Zrodziła się potrzeba takiego sposobu publikowania informacji, tak by mogły być z łatwością wykorzystane przez odbiorcę, w tym, co bardzo istotne, odbiorcę automatycznie przetwarzającego dane. Oczywiście staje się więc konieczność opisywania znaczenia danych w miejscu, w którym się znajdują. Opis znaczenia danych – semantyka, nie powinien być oddzielany od samych danych. Odpowiedni sposób publikowania informacji pozwoli programom – agentom, na automatyczne ich wykorzystanie – odnalezienie odpowiedniej informacji, zinterpretowanie jej lub porównanie z inną i

wykonanie jakiejś czynności na podstawie tej informacji. Agent programowy mógłby np. wyszukać pokoje hotelowe zgodnie z zadaniem i kryteriami i zamówić nocleg na odpowiedni dzień. Bazowanie na danych i ich znaczeniu, składowanych się łącznie na informację, może zapewnić wiele korzyści.

- Możliwość dostosowania treści do preferencji użytkownika. Obecnie serwisy internetowe udostępniają usługi bez względu na preferencje użytkownika korzystającego z nich lub wykorzystują własne rozwiązania do tego celu. Brak możliwości standardowego opisu preferencji użytkownika nie pozwala budować standardowych rozwiązań, które mogłyby być wykorzystywane w różnych serwisach internetowych.
- Ułatwienie integracji na poziomie wymiany danych pomiędzy systemami. Obecnie przeprowadzając integrację systemów informatycznych konieczne jest wypracowanie odpowiednich standardów znaczenia wymienianych informacji – wypracowanie wspólnych pojęć. Taka standaryzacja dotyczy konkretnych dwóch integrowanych systemów – najczęściej nie jest wykorzystywana przy kolejnych integracjach, brakuje wzorca określającego pewne ramy takiej standaryzacji.
- Ułatwienie identyfikacji i wyszukiwania usług sieciowych. Dynamiczny rozwój architektury zorientowanej na usługi [9] (SOA, ang. Service Oriented Architecture) wpływa na coraz większą popularność usług sieciowych. Wiele z udostępnianych usług realizuje podobne czynności lub korzysta z podobnych danych. Konieczne jest więc rozwiązanie pozwalające na łatwe wyszukiwanie usług sieciowych w zależności od ich możliwości. Równie ważne jest wprowadzenie mechanizmów umożliwiających łatwe dostosowanie typów wymienianych danych.

Osiągnięcie powyższych korzyści w dużej mierze wymaga wcześniejszego przygotowania. Należy zdefiniować pojęcia¹, hierarchię i relacje między tymi pojęciami. Istotne jest by całość (pojęcia, hierarchia i relacje) była powszechnie znana i wykorzystywana. W oparciu o tak zdefiniowaną strukturę pojęć możliwe jest odpowiednie klasyfikowanie danych, a także łatwe ich wyszukiwanie. Sieć pojęć określa również w jakimś stopniu i kontekście rzeczywistość – pojęcia i relacje między nimi, związane z codziennym życiem. Staje się więc opisem rzeczywistości – ontologią². W ujęciu filozoficznym ontologia pozwala na wyłumaczenie zależności pomiędzy bytami i w właściwości tych bytów, tak by opisać otaczającą nas rzeczywistość, w znaczeniu informatycznym ontologia oznacza ustandaryzowany sposób zapisu znaczenia danych. Dzięki standaryzacji zapisu ontologii, komputery i oprogramowanie posługując się pojęciami z ontologii, mogą w pewnym stopniu „rozumieć” rzeczywistość i tym samym lepiej korzystać z danych opisanych ontologią.

W 1997 roku organizacja World Wide Web Consortium (W3C) [16] rozpoczęła proces standaryzacji sposobu zapisu ontologii, publikując standard Resource Description Framework (RDF) [19], który w 1999 roku stał się oficjalnym standardem W3C. RDF pozwala na prosty zapis trójek pojęć. Każda trójka składa się z tematu (*subject*), właściwości (*predicate*) i obiektu (*object*). Taki zapis pojęć podstawowy element sieci definicji (każdy obiekt może być również tematem w innej trójce).

¹ abstrakcyjny, myślowy odpowiednik obiektu, zjawiska, itp.

² Termin „ontologia” zaczerpnięty jest z filozofii: „jeden z podstawowych działów filozofii, nauka o bycie (...) sprowadza się do definiowania pojęć” [2].

Standard RDF został uzupełniony w 2004 roku o specyfikację RDF Schema (RDF-S) [20]. RDF-S wprowadził możliwość budowania meta-pojęć – klas, podklas, właściwości – oraz standardowy sposób definiowania nazwy pojęcia (*label*) i jego opisu (*comment*).

Kolejnym etapem rozbudowy standardów sieci semantycznych było zwiększenie ekspresywności języków przeznaczonych do zapisu ontologii. Organizacja W3C opublikowała standard OWL (Web Ontology Language) [18]. Pozwala on m.in. na wyrażenie liczności zbiorów pojęć, zawierania się i różnicę pojęć, warunków koniecznych i wystarczających dla danego pojęcia. Większa ekspresywność języka pozwala na weryfikację pojęć dodawanych do ontologii, a także znajdowanie pewnych faktów i właściwości nie-wprost. Dodatkowo język ten zapewnia możliwość integracji dwóch ontologii kojarząc identyczne pojęcia ze sobą. Standardy opisu ontologii pozwalają więc na opisanie pojęć oraz sieci powiązań pomiędzy pojęciami. Podstawowym mechanizmem wykorzystywanym przez RDF do identyfikacji podmiotu, predykatu i obiektu jest URI [8].

3. SIECI SEMANTYCZNE W INTEGRACJI SYSTEMÓW

Wiele cennych danych przechowywanych jest obecnie w relacyjnych bazach danych. Wiele z tych danych opisuje podobne elementy (produkty, osoby itp.) z różnych punktów widzenia. Przydatne jest także zintegrowanie tych danych, które pozwoliłoby na wykorzystanie w pełni wiedzy w nich zgromadzonych. Dobrym rozwiązaniem wydaje się być zastosowanie sieci semantycznych i przeprowadzenie semantycznej integracji systemów informatycznych. W głównej mierze sprowadza się ona do wypracowania wspólnej terminologii w nazywaniu i określaniu danych, z których korzystać mogą różne aplikacje. Pod nazwanymi i pojęciami mają być dostępne dane z poszczególnych systemów informatycznych. Dla pojedynczego elementu mogą być dostępne dane z różnych systemów informatycznych, w ten sposób uzupełniając informacje o pojedynczym elemencie.

Semantyczna integracja motywowana jest uproszczeniem dostępu do danych, możliwością odczytu znaczenia danych i ukrytej semantyki nie zapisanej bezpośrednio w relacjach w bazie danych. Przykładowe zalety wykorzystania sieci semantycznych w dostępie do baz danych są następujące [6]:

- Bazując na języku RDF i OWL stosuje się URI do identyfikacji zasobów sieciowych. Z jednej strony umożliwia to odnoszenie się do zewnętrznych lub specyficznych dla danej dziedziny ontologii; z drugiej strony umożliwia to synchronizację danych z różnymi systemami np. gromadzącymi informacje o klientach, produktach itp.
- OWL pozwala na definiowanie „bogatszych” w właściwości i relacji. Właściwości obiektowe mogą zostać zdefiniowane, jako symetryczne, funkcjonalne, odwrotne, przechodnie. Właściwości obiektowe są odpowiednie do opisanie złożonych relacji między produktami oraz pomiędzy produktami i innymi encjami składającymi się na informację o produkcie.
- Ekspresywność OWL pozwala na określenie logicznych klas (iloczyn, złączenie, dopełnienie zbiorów), co umożliwia automatyczną klasyfikację pozycji produktu. Na przykład kategorie nowego produktu można określić, jako iloczyn dwóch innych: przykładem mogą być smartfony, które posiadają cechy zarówno telefonów, jak i palmtopów. Każdy produkt będący jednocześnie telefonem i palmtopem jest zatem smartfonem.
- Za pomocą ograniczeń w języku OWL można definiować dynamiczne kategorie, które nie występowały we wcześniej opracowanej hierarchii kategorii i są określone przez użytkowników w czasie wprowadzania zapytania. Restrykcje mogą przedstawiać

złożone i potencjalnie zmieniające się kategorie. Na przykład, stosując ograniczenie o minimalnej liczności, możliwe jest zdefiniowanie kategorii "produkty przestarzałe", która zawiera wszystkie produkty zastąpione przynajmniej przez jeden inny produkt. Pozycje należące do dynamicznych kategorii mogą być pozyskiwane za pomocą wnioskowania ontologii OWL.

Istotnym zagadnieniem jest umożliwienie mapowania relacyjnych baz danych w przestrzeni sieci semantycznych. Obecnie zagadnieniem tym zajmuje się grupa robocza RdfRDB [12] pod patronatem organizacji W3C. Artykuły prezentowane w ramach prac tej grupy wymieniają różne doświadczenia związane z wdrażaniem tego typu rozwiązań. Wskazuje się między innymi na potencjalne problemy, z jakimi można się spotkać w przypadku mapowania bazy danych na RDF [7].

Zagadnienie semantycznej integracji baz danych można podzielić na dwa etapy:

- udostępnianie danych pojedynczej bazy danych za pomocą zdefiniowanej ontologii,
- wykorzystanie jednej lub wielu źródeł tak udostępnionych danych w konkretnym zastosowaniu biznesowym.

Udostępnienie danych semantycznych polega na dodaniu komponentu – narzędzia, które umożliwi odpytywanie danych zapisanych w bazie danych komunikując się z wykorzystaniem protokołu HTTP i posługując się zapytaniami uwzględniającymi nazewnictwo danych zapisane w ontologii.

W dziedzinie udostępniania danych semantycznych wypracowano dotąd dwie standardowe metody. Pierwszą z nich jest udostępnianie danych poprzez dokumenty RDF, dostępne pod adresami URL, odnoszącymi się do opisywanych elementów. Drugim sposobem jest udostępnianie interfejsów SPARQL (tzw. SPARQL endpoint) [13], które pozwalają na wysyłanie zapytań w języku SPARQL (za pomocą protokołu HTTP) i otrzymywanie odpowiedzi zawężonej do danych, których dotyczyło zapytanie.

4. SIECI SEMANTYCZNE W PRZEMYŚLE

Projekty związane z zastosowaniem sieci semantycznych w przemyśle [1] mają na celu przede wszystkim ujednoczenie terminologii używanej na różnych etapach produkcji, monitorowania parametrów itp. Taka systematyzacja pozwala na właściwe zrozumienie odpowiednich pojęć i umożliwia łatwą wymianę danych pomiędzy różnymi przedsiębiorstwami uczestniczącymi w różnych etapach produkcji. Struktury danych opisujące poszczególne parametry porządkują zależności między nimi, są łatwo rozszerzalne i łatwo zarządzane, a standaryzacja zapisu pojęć pozwala na łatwą integrację różnych systemów informatycznych zarówno pomiędzy firmami jak i wewnątrz firmy.

4.1. The Norwegian Daily Production Report (DPR)

Celem projektu była standaryzacja danych związanych z procesem wydobywania ropy naftowej tak, aby możliwe było śledzenie ich przez partnerów i nadzorców przedsiębiorstwa. Projekt był testowany na polach naftowych Hydro's Åsgård, a udziałowcami projektu były firmy Shell, BP, Chevron, Exxon, Statoil, ConocoPhillips, i Total. DPR³ jest częścią wielopoziomowego systemu raportowania obejmującego zarówno wydobywanie jak i przygotowanie złóż do eksploatacji.

³ Szczegóły projektu można znaleźć na stronie <http://production.posccaesar.org/browser/projects> [2009-11-14]

4.2. The Active Knowledge Systems for Integrated Operations (AKSIO)

W ramach projektu zbudowany został zintegrowany system do zarządzania wiedzą wspierający operacje odwiertów na platformach wiertniczych (wydobycie ropy naftowej i gazu). Takie przedsięwzięcie wymagało integracji danych pochodzących z baz danych, aplikacji i baz wiedzy oraz integracji bieżących informacji otrzymywanych z platform wiertniczych. Większość funkcjonalności systemu jest wspomagana przez technologię sieci semantycznych z uwzględnieniem adnotacji treści pojęciami z ontologii i inteligentnego wyszukiwania treści.

4.3. The Integrated Information Platform (IIP)

Projekt⁴ był wspierany przez Norwegian Research Council (NRC) i miał na celu utworzenie platformy integrującej dotychczasowe dane i standardy technologiczne związane z przemysłem naftowym oraz realizację nowych ontologii (ontologie ropy i gazu) [3].

W ramach projektu za pomocą ontologii i taksonomii zostały zintegrowane dane niezbędne dla podwodnych urządzeń sejsmicznych, odwiertów, produkcji, zarządzania i nadzoru całego procesu wydobycia gazu i ropy oraz systemy ekspertowe. Zastosowanie takiego rozwiązania ułatwia wymianę danych i pozwala na optymalizację poszczególnych etapów procesu wydobycia. Dzięki wprowadzeniu ontologii możliwe jest usystematyzowanie informacji o kolejnych etapach procesu technologicznego, co ważne jest w procesie nadzoru nad wydobyciem. Wykorzystano również wyniki innego projektu badawczego – Petrotechnical Open Software Corporation's POSC Caesar. W ramach tego projektu wykorzystując standard OWL utworzono dotychczas ontologię zawierającą około 60000 klas, które opisują wyposażenie pól naftowych (odpowiednik standardu ISO 15926-7). Reguły w OWL opisują specyficzne właściwości klas i czynności jakie należy podjąć, jeżeli któraś z reguł zostanie przekroczona. Projekt został zakończony w 2007 roku.

4.4. InfoWeb

W projekcie InfoWeb[17] wykorzystano sieci semantyczne do utworzenia bazy wiedzy zgodnej ze standardem ISO 15926. Baza ta ma być źródłem informacji do tworzenia przewodników, procedur, wymiany danych, oprogramowania pomiędzy różnymi przedsiębiorstwami produkcyjnymi. Celem takiej standaryzacji słownictwa i pojęć ma być m.in.: ułatwienie zarządzania łańcuchem dostaw (zamówienia części, komponentów do dalszej produkcji itp.), możliwość porównywania ofert, właściwości produktów, ułatwiona ochrona, zabezpieczanie i przeglądanie danych, ułatwienie współpracy pomiędzy projektantami a wytwórcami. Standard obejmuje m.in. przemysł związany z rafineriami, instalacjami chemicznymi i energetycznym, budownictwo itp. Pojęcia w bazie wiedzy zostały zapisane za pomocą standardu RDF i mogą stanowić podstawę wymiany danych i integracji złożonych operacji pomiędzy przedsiębiorstwami przemysłowymi.

4.5. The Geosciences Network

System GEON's CHRONOS [15] jest narodowym portalem stratygraficznym w USA. Bazuje on na systemach ArcIMS (system dostarczający dynamiczne mapy typu GIS poprzez Internet) oraz the National Carbon Sequestration (system integrujący różne źródła danych geologicznych w USA). System GEON zapewnia dostęp do rozproszonych baz danych w całym kraju. W wykorzystuje on ontologie zapisane w OWL do reprezentacji wiedzy, słownictwa, hierarchii pojęć i bardziej zaawansowanych relacji pomiędzy terminami naukowymi. Dzięki temu ułatwione zostało wyszukiwanie i wymiana danych, usług sieciowych itp. związanych z sejsmologią.

⁴ Szczegóły dotyczące projektu można znaleźć na stronie <http://research.idi.ntnu.no/IIP> [2009-11-14]

5. DOŚWIADCZENIA EMAG W ZAKRESIE SIECI SEMANTYCZNYCH

5.1. Uniwersalna warstwa danych dla spółek węglowych.

W spółkach węglowych działają specjalne komórki organizacyjne odpowiedzialne za koordynowanie poziomu zużycia poszczególnych mediów, takich jak energia elektryczna, ciepła, woda oraz sprężone powietrze. Koordynacja ta jest wymuszona poprzez konieczność deklarowania u dostawców mediów poziomu ich wykorzystania. Na bieżąco kontrolowane jest zużycie poszczególnych mediów tak, by nie dopuścić do przekroczenia zadeklarowanych poziomów mocy. Istnieją systemy, które na bieżąco pobierają dane z kopalń i prezentują bieżące wartości. Zadaniem dyspozytora jest planowanie wykorzystania mediów oraz w razie awarii skoordynowanie odpowiednich działań mających na celu optymalizację wykorzystania mediów ramach spółki.

Zadaniem dyspozytorów jest również tworzenie odpowiednich raportów i przedstawianie ich zarządom spółki. Raporty te zazwyczaj tworzone są w oparciu o dane związane ze zużyciem energii, jednak okresowo zachodzi również konieczność tworzenia raportów w oparciu o dane pochodzące z innych źródeł.

W wyniku prowadzonych w EMAG prac badawczych opracowano koncepcję realizacji uniwersalnej warstwy danych. Zakłada ona wprowadzenie warstwy wspólnej dla wszystkich źródeł danych. Warstwa ta oparta na danych relacyjnych zmapowanych w oparciu o ontologię zapewni wspólny schemat danych. Rozwiązanie takie zapewnia przezroczystość dostępu do danych pochodzących z różnych źródeł oraz możliwość definiowania dodatkowych relacji, hierarchii opisów danych, nieistniejących w bazach relacyjnych. W ramach systemu zapewniono możliwość tworzenia dedykowanych aplikacji korzystających z takiej warstwy danych: inteligentnych formularzy, modułu raportów (na wzór arkusza kalkulacyjnego), modułu doraźnych raportów. Aplikacje te są klientami wspólnej warstwy danych. W warstwie danych uzupełniono o narzędzie pozwalające na zarządzanie danymi poprzez eksport i import danych do popularnych formatów (pliki Microsoft Office, pliki CSV, pliki XML), przeglądanie danych itp.

5.2. Wirtualny Konsultant Usług Publicznych

Prace związane z zagadnieniami sieci semantycznych prowadzone były w ramach projektu Wirtualnego Konsultanta Usług Publicznych (WKUP), którego celem była budowa spersonalizowanego systemu informacyjnego realizującego usługi administracji publicznej na poziomie regionalnym. Zadaniem WKUP jest zidentyfikowanie problemu obywatela lub przedsiębiorcy, odnalezienie na tej podstawie odpowiadającej temu problemowi usługi publicznej (lub kilku usług), a następnie przeprowadzenie przez proces kompletowania informacji niezbędnych do jej realizacji. System poprzez interfejs użytkownika zapewnia obsługę elektronicznych formularzy zastępujących dotychczasowe formularze papierowe. W systemie zastosowano narzędzia przetwarzania języka naturalnego oraz powiązano usługi sieciowe z opracowanymi w ramach projektu ontologiami dla środowiska świadczenia usług publicznych. Realizacja usług oparta została na wyszukiwaniu opartym na ontologii SKOS [4] i wnioskowaniu semantycznym.

5.3. Metodyka ATOM

Na podstawie badań i prac, jakie były prowadzone w EMAG, a które dotyczyły zastosowania technik semantycznych, powstała metodyka ATOM. Określa ona etapy realizacji integracji semantycznej oraz opisuje sposób typowego postępowania w przypadku integracji danych

z pomocą sieci semantycznych. Metodyka składa się z czterech podstawowych elementów: analizy, wyboru narzędzi do integracji, budowy ontologii dziedzinowej oraz mapowania danych. Szczegółowy opis poszczególnych elementów metodyki jest tematem odrębnej publikacji [4].

6. PODSUMOWANIE

Technologie semantyczne posiadają duży potencjał rozwoju dzięki zapewnieniu jednolitego języka opisu danych i metadanych. Ustandaryzowany sposób opisu informacji może ułatwić organizację łańcucha dostaw oraz dobór najbardziej odpowiadających elementów pod względem jakości i ceny. Semantyczna standaryzacja parametrów monitorowanych podczas procesu produkcji może ułatwić ich wykorzystanie na różnych warstwach zarządzania w przedsiębiorstwie. Takie informacje powiązane ze sobą znaczeniowo (semantycznie) mogą ułatwiać przekrojową analizę danych pochodzących z różnych systemów monitorowania produkcji. Może to przyczynić się do podejmowania lepszych decyzji związanych zarówno z zarządzaniem przedsiębiorstwem (fabryką, procesem wydobywania, energetyką), jak i z bezpieczeństwem.

Jak wskazują przytoczone do świadczenia norweskie, zastosowanie technik semantycznych pozwala zrealizować integrację przedsiębiorstw i organizacji skupionych wokół jednej dziedziny. Stanowią one przykład projektów badawczych, które realizowane były z myślą o przemyśle, a ich wyniki zostały z powodzeniem wdrożone.

Sieci semantyczną mogą, więc być przydanym narzędziem dla przedsiębiorstw z różnych branż jednak sposób ich wykorzystania obecnie jest dopiero stopniowo definiowany i precyzowany. Mimo sukcesów w wykorzystaniu sieci semantycznych należy pamiętać, że jest to dziedzina wymagająca dalszych prac badawczych, a w realizacji związanych z nimi projektów niezbędny jest udział jednostek naukowych, posiadających doświadczenie w definiowaniu i stosowaniu ontologii.

Bibliografia

1. Chum F., *Semantic Web Use Cases and Case Studies*, <http://www.w3.org/2001/sw/sweo/public/UseCases/Chevron/> [14.11.2009].
2. Encyklopedia podręczna, Wydawnictwo Zielona Sowa, Kraków 2005.
3. Fjellheim R., Norheim D., *Semantic Technology for Knowledge and Work Process Management in the Oil & Gas Industry*, W3C Workshop on Semantic Web in Energy Industries, 2008.
4. Górka W., Piasecki A., Szymocha A., *Semantyczna integracja systemów informatycznych oraz zarządzanie wiedzą w organizacjach*, Technologie wiedzy w zarządzaniu publicznym '09, Akademia Ekonomiczna, Katowice 2009.
5. Górka W., Socha M., Piasecki A., *Wyszukiwarka informacji oparta na ontologii Simple Knowledge Organisation System (SKOS)*, Konferencja EMTECH-2008, Materiały konferencyjne, Katowice 2008.
6. Ma L., Kulkarni K., Fokoue A., *Semantic Web Technologies and Data Management*, 2007.
7. Malhotra A., *Integrating Data from Relational Databases Using RDF*, 2007.
8. Network Working Group, Uniform Resource Identifier (URI): Generic Syntax, <http://labs.apache.org/webarch/uri/rfc/rfc3986.html> [14.11.2009]
9. Newcomer E., Lomow G., *Understanding SOA with Web Services*, Addison-Wesley Professional, 2005.
10. Ochman J., *Integracja w systemach informatycznych zarządzania*, PWE, Warszawa 1992.

11. Olszak C., Sroka H., Zintegrowane systemy informatyczne w zarządzaniu, Wydawnictwo Uczelniane AE w Katowicach, Katowice 2001.
12. RdfRDB - W3C Workshop on RDF Access to Relational Databases, <http://www.w3.org/2007/03/RdfRDB> [14.11.2009].
13. SPARQL – język zapytań dla sieci semantycznych zapisanych za pomocą RDF (<http://www.w3.org/TR/rdf-sparql-query>).
14. Stempnakowski Z., Efektywność organizacji a jej systemy informatyczne, <http://www.e-marketing.pl/artyk/artyk89.php> [14.11.2009]
15. The Geosciences Network ,<http://www.geongrid.org/>, [14.11.2009]
16. W3C, <http://www.w3c.org>, [14.11.2009]
17. W3C, Extensible Markup Language (XML), <http://www.w3.org/XML/>, [14.11.2009]
18. W3C, OWL Web Ontology Language Overview, <http://www.w3.org/TR/owl-features/>, [14.11.2009]
19. W3C, Resource Description Framework (RDF), <http://www.w3.org/RDF>, [14.11.2009]
20. W3C, RDF Vocabulary Description Language 1.0: RDF Schema, <http://www.w3.org/TR/rdf-schema/>, [14.11.2009]