

dr Dariusz Put

Katedra Systemów Obliczeniowych,
Wydział Zarządzania, Uniwersytet Ekonomiczny w Krakowie

Systemy gromadzenia danych dedykowane współczesnym organizacjom sieciowym

WPROWADZENIE

Obserwowany w ostatnich latach rozwój technologii teleinformatycznych był jednym z czynników, które doprowadziły do powstania różnorodnych sieciowych struktur organizacyjnych. Ich głównymi cechami są czasowość istnienia, heterogeniczność systemów wdrożonych w wchodzących w ich skład podmiotach oraz niejednokrotnie duża dynamika zmian liczby tworzących je jednostek. Ze względu na swe własności tradycyjne systemy bazodanowe nie mogą być wykorzystane do współdzielenia informacji przez jednostki tworzące struktury sieciowe. A przecież możliwość wymiany zasobów informacyjnych jest niezwykle istotna, szczególnie z punktu widzenia efektywności działania tego typu organizacji. Celem artykułu jest omówienie problemu współdzielenia zasobów informacyjnych w dynamicznym środowisku, w jakim działają organizacje sieciowe oraz istniejących rozwiązań w tym zakresie.

W opracowaniu zdefiniowano istotę organizacji sieciowych, omówiono ich rodzaje oraz porównano ich własności z cechami organizacji tradycyjnych. Dokonano identyfikacji najistotniejszych własności systemów integracyjnych oraz porównano wybrane rozwiązania dedykowane integracji niejednorodnych zasobów informacyjnych.

WŁASNOŚCI I RODZAJE ORGANIZACJI SIECIOWYCH

Rozwój technologii teleinformatycznych umożliwił tworzenie nowych form organizacyjnych, głównie o strukturze sieciowej, opartych na wzajemnych relacjach przedsiębiorstw niepowiązanych kapitałowo. Sieć łączy jednostki autonomiczne lub powiązane w niewielkim stopniu, przy czym podmioty tworzące sieć współpracują wykorzystując technologie informatyczne, stanowiące wręcz warunek konieczny do zaistnienia tej formy organizacji¹.

¹ B. Barczak, K. Bartusik, A. Kozina, *Modele strukturalne organizacji uczącej się* [w:] *Doskonalenie struktur organizacyjnych przedsiębiorstw w gospodarce opartej na wiedzy*, red. A. Stabryła, Wydawnictwo C.H. Beck, Warszawa 2009, s. 53–116.

Tego typu organizacje łącząc się mogą tworzyć dwie struktury²:

- gwiazdy, w których istnieje firma wiodąca, współpracująca z wieloma kontrahentami,
- połączeń węzłowych, w których wszystkie podmioty są równoprawne.

Połączenia sieciowe mogą być – i często są – tymczasowe, gdy intensyfikacja kontaktów jest zmienna w czasie i uzależniona od bieżących potrzeb, lub sformalizowane, gdy organizacje łączą się oficjalnie, tworząc korporacje działające na danym terenie lub o zasięgu globalnym.

Głównymi czynnikami powstania organizacji sieciowych były przede wszystkim trzy megatrendy³:

- integracja rozwiązań informatycznych i telekomunikacyjnych wpływająca na wzrost wydajności, skalowalności i mobilności systemów,
- globalizacja rynku produktów i zasobów,
- wzrost znaczenia informacji i wiedzy oraz zapotrzebowania na nie.

Celem tworzenia sieci jest poprawa warunków działania, obniżenie kosztów, wzrost efektywności. Do innych korzyści, jakie mogą uzyskać przedsiębiorstwa z tytułu łączenia się w sieci, należy zaliczyć⁴:

- swobodniejszy przepływ informacji,
- łatwiejszy dostęp do informacji,
- wykorzystywanie zasobów należących do partnerów,
- wzrost ilości zasobów informacyjnych, ułatwienia w dostępie do zasobów innych organizacji oraz większy dostęp do zasobów i umiejętności nie będących w posiadaniu danej jednostki.

Wśród słabych stron organizacji sieciowych wymienia się problemy związane z zarządzaniem wzajemnymi związkami, wzrost kosztów transakcji rosnący wraz ze wzrostem liczby uczestników, możliwość pojawienia się zachowań oportunistycznych i egoistycznych, możliwy wyzysk słabszych przez partnerów posiadających kontrolę nad najważniejszymi informacjami, problemy związane z zarządzaniem informacjami i niebezpieczeństwo wycieku informacji⁵. Istotą sieci jest to, że niezależne podmioty, wykorzystując technologie teleinformatyczne, tworzą dobrowolne, luźne, często incydentalne konfiguracje⁶. Organizacje sieciowe mogą przybierać różne formy, m.in.:

² S. Korenik, *Dysproporcje w rozwoju rejonów Polski – wybrane aspekty*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław 2003.

³ A. Jaruga, *Struktury sieciowe przyczynkiem kreowania wirtualnej organizacji przedsiębiorstwa* [w:] *Zarządzanie organizacjami gospodarczymi. Koncepcje i metody*, red. J. Lewandowski, t. 1, Wydawnictwo Politechniki Łódzkiej, Łódź 2005, s. 52–60.

⁴ B. Barczak, K. Bartusik, A. Kozina, *Modele...; Zarządzanie wiedzą w przedsiębiorstwie*, red. K. Perechuda, WN PWN, Warszawa 2005.

⁵ B. Miłucha, A. Pietruszka-Ortyl, *Organizacje sieciowe*, Zeszyty Naukowe Akademii Ekonomicznej w Krakowie, Prace z zakresu zachowań organizacyjnych, nr 715, 113–130, Kraków 2006.

⁶ A. Jaruga, *Struktury...*

- klastrow, skupiających konkurujące firmy z tej samej branży, a także inne jednostki działające w danym regionie. Klastry mogą powstawać w efekcie polityki władz lokalnych, regionalnych, na poziomie narodowym lub ponadnarodowym albo oddolnie, z inicjatywy przedsiębiorstw,
- parków technologicznych, czyli kompleksów gospodarczych, których głównym zadaniem jest tworzenie warunków do prowadzenia działalności gospodarczej. W ich ramach realizowana jest polityka w zakresie⁷: wspomagania innowacyjnych przedsiębiorstw nastawionych na rozwój produktów i metod wytwarzania w technologicznie zaawansowanych branżach, optymalizacji warunków transferu technologii i komercjalizacji rezultatów badań z instytucji naukowych do praktyki gospodarczej,
- organizacji procesowych powstających w wyniku wyodrębnienia ciągów czynności, które składają się na procesy przebiegające w przedsiębiorstwie oraz ich identyfikacji i nazwania w celu racjonalnego zarządzania nimi. Najistotniejszym zadaniem podczas tworzenia organizacji procesowej jest rozpoznanie i koordynacja (integracja) procesów zachodzących w organizacji, prowadzących do osiągnięcia celu,
- organizacji wirtualnych, w których łączone są zasoby współpracujących jednostek realizujących wspólne przedsięwzięcia w celu osiągnięcia większych korzyści, niż wtedy, gdyby każda z nich funkcjonowała indywidualnie⁸. Organizacje tego typu optymalizują przede wszystkim wykorzystanie zasobów niematerialnych. Jedną z cech organizacji wirtualnych jest czasowość współpracy firm, często dobieranych dynamicznie, zlokalizowanych w różnych miejscach, które łączy wspólny cel,
- fraktali bazujących na samoorganizujących się autonomicznych zespołach pracowniczych, opierających swoją działalność na pełnej dostępności do informacji i jej drożnym, dynamicznym przepływie⁹. Fraktem może być całe przedsiębiorstwo (makrofraktal), jak i jego jednostki organizacyjne (mikrofraktale). Strukturę organizacyjną tego typu jednostek charakteryzuje elastyczność, dynamika zmian, procesowość oraz sieciowość.

Wspólnymi cechami wymienionych form organizacyjnych są rozproszenie i niejednorodność, aby mogły sprawnie funkcjonować muszą korzystać z technologii teleinformatycznych umożliwiających współdzielenie danych, informacji i wiedzy. Wymaga to zaprojektowania i wdrożenia systemów zarządzania zasobami informacyjnymi, umożliwiających odrębne ich gromadzenie, ale wspólne wykorzystanie. Omówione struktury charakteryzuje duża dynamika zmian,

⁷ K.B. Matusiak (red.), *Ośrodki innowacji w Polsce. Analiza krajowych instytucji wspierających innowacyjność i transfer technologii*, PARP, SOOIPP, Warszawa 2005.

⁸ Dziembek D., *Struktura podmiotowa organizacji wirtualnej* [w:] *Technologie i systemy komunikacji oraz zarządzania informacją i wiedzą*, red. L. Kiełtyka, Warszawa 2008, s. 153–162.

⁹ Binsztok A., *Organizacja fraktalna* [w:] *Zarządzanie przedsiębiorstwem w turbulentnym otoczeniu. Ku superelastycznej organizacji*, red. R. Krupski, PWE, Warszawa 2005, rozdz. 6.

związki między ich składnikami są nieformalne, łączone są odrębne jednostki, organizowane są, zależnie od potrzeb rynkowych, w odpowiedzi na pojawiające się wyzwania i zmieniające się warunki prowadzenia działalności (por. tabela 1). Aby taka dynamiczna funkcjonalność mogła zostać osiągnięta niezbędna jest koordynacja procesów, połączenie aplikacji i systemów oraz odpowiednie zarządzanie zasobami informacyjnymi. Wspomagające je systemy informatyczne muszą charakteryzować się rozszerzalnością, elastycznością, pełną lub możliwie dużą automatyzacją czynności i procesów oraz łatwością obsługi i przyjaznym interfejsem umożliwiającym m.in. nieskomplikowany dostęp do zasobów informacyjnych, niezbędnych podczas realizacji codziennych zadań.

Tabela 1. Wybrane własności organizacji tradycyjnych i rozproszonych form organizacyjnych

Organizacje tradycyjne	Organizacje sieciowe
niewielka dynamika zmian	duża dynamika zmian
oparte na formalnych związkach	związki między partnerami są nieformalne
jednostki współpracują przez dłuższy czas	często zmieniająca się liczba kooperantów
posiadają ustaloną strukturę organizacyjną	brak organów zarządzających
oparte na centralnym projekcie bazy danych	brak wspólnego projektu bazy danych
potrzeby informacyjne zidentyfikowane na etapie projektowania	nieznane potrzeby informacyjne pracowników poszczególnych jednostek składowych

Źródło: opracowanie własne.

WSPÓLDZIELENIE DANYCH W SYSTEMACH TRADYCYJNYCH I INTEGRACYJNYCH

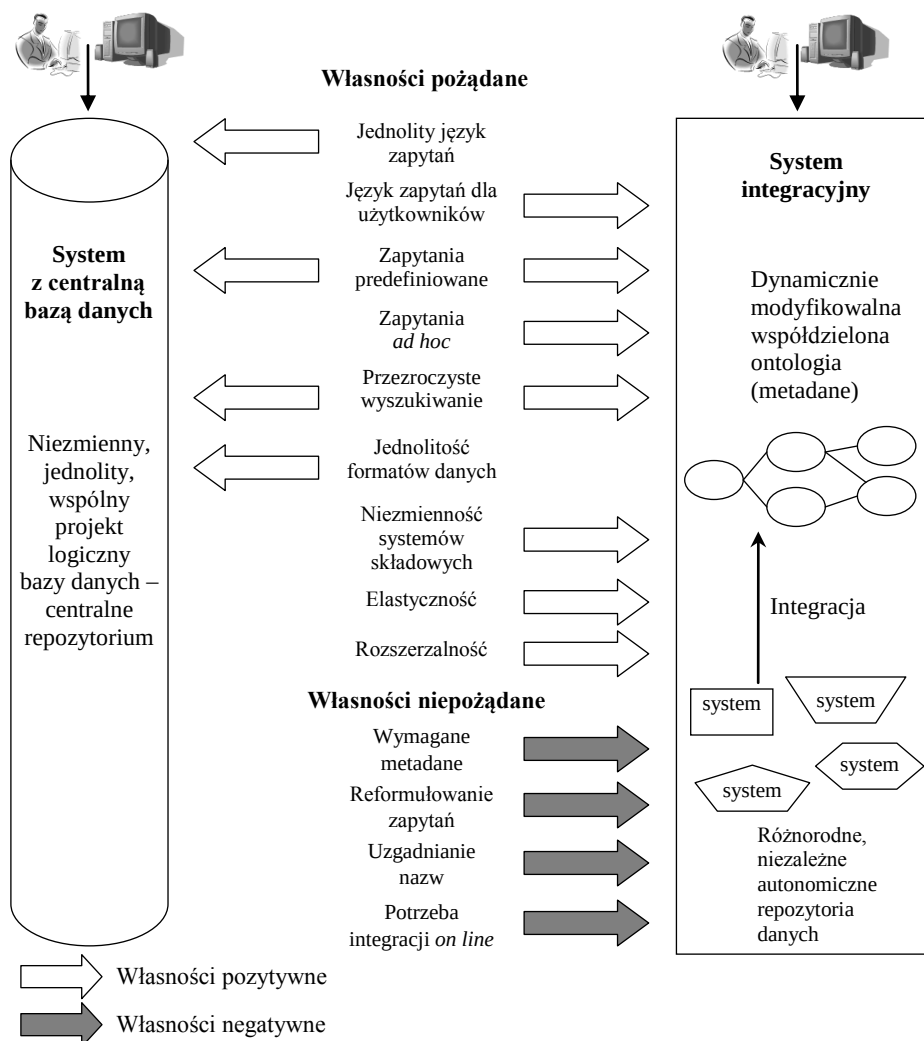
We współczesnych systemach transakcyjnych, w tym także w klasycznych systemach zintegrowanych umożliwiających dostęp do danych i informacji zgromadzonych w repozytoriach, w których warstwę zasobów informacyjnych stanowi jedna baza danych, nie istnieje problem różnorodności, gdyż¹⁰:

- wykorzystywany jest jeden język zapytań,
- dane mają jednolity, uzgodniony format, wszystkie koncepty (klasy, relacje) oraz charakteryzujące je atrybuty mają takie same, ustalone w trakcie projektowania systemu nazwy,
- więzy integralności i normalizacja minimalizują redundancję oraz możliwość występowania błędów, braków i niejednoznaczności w danych,
- kwerendy są kierowane do repozytoriów, a wybrane zasoby do aplikacji użytkownika, gdzie odpowiednie aplikacje prezentują informacje,
- ze względu na złożoność języka zapytań perspektywy są predefiniowane – na etapie projektowania identyfikowane są potrzeby informacyjne użytkowników

¹⁰ Własności systemów z centralną bazą danych i integracyjnych zobrazowano na rysunku 1.

i na tej podstawie generowane są kwerendy będące, w trakcie eksploatacji systemu, źródłem formularzy i raportów,

- nie ma potrzeby konwersji zapytań,
- nie występuje problem uzgadniania nazewnictwa faktów czy charakteryzujących je atrybutów,
- nie jest konieczna publikacja metadanych o aktualnie dostępnych zasobach,
- dane wybrane z repozytorium nie muszą być łączone.



Rysunek 1. Własności systemów z centralną bazą danych i rozwiązań integracyjnych opartych na współdzielonej ontologii

Źródło: opracowanie własne.

Rozwiązanie takie posiada jednak wady. Tradycyjne systemy informacyjne, w tym także systemy zintegrowane, oparte na centralnej bazie danych, charakteryzują się niewielką elastycznością lub nawet jej brakiem. We współczesnej, szybko zmieniającej się rzeczywistości, w której niemal codziennie pojawiają się nowe źródła danych i informacji, struktury takich systemów muszą być modyfikowane, co wiąże się zwykle z dużym nakładem środków i niejednokrotnie koniecznością zaprojektowania całego systemu od początku. Należy więc poszukiwać rozwiązań bardziej elastycznych, które pozwalałyby na dynamiczne dostosowanie struktury repozytorium do szybko zmieniającej się rzeczywistości bez konieczności zmiany systemu. Do innych wad rozwiązań tradycyjnych trzeba zaliczyć brak skalowalności, brak możliwości formułowania zapytań *ad hoc* przez użytkowników końcowych (ze względu na złożoność języka zapytań) oraz konieczność przeprowadzania procesu projektowania uwzględniającego migrację systemów transakcyjnych do tworzonego rozwiązania, a co za tym idzie – newralgiczność etapu wdrożenia.

Odmierna sytuacja występuje w systemach, których zadaniem jest integracja heterogenicznych i rozproszonych zasobów informacyjnych, w których przyjmuje się założenie o nieingerencji w istniejące systemy transakcyjne (lub analityczne), co jest szczególnie istotne w przypadku dynamicznych organizacji sieciowych. W takiej sytuacji występuje wieloaspektowa różnorodność: modeli danych, systemów przechowywania danych i informacji, języków zapytań, rodzajów i formy współdzielonych zasobów, nazewnictwa obiektów i zdarzeń, metod identyfikacji faktów. Muszą więc istnieć rozwiązania niwelujące te niejednorodności. Model integracyjny, na podstawie którego budowane są systemy pozwalające na przeszukiwanie heterogenicznych zasobów informacyjnych, musi zawierać komponenty, które udostępniają metainformację o współdzielonych zasobach, umożliwiają ich wyszukanie, wybiorą żądane dane i informacje w sposób niewidoczny dla użytkowników, udostępnią im wyszukane zasoby oraz zaprezentują w czytelnej postaci. Systemy takie muszą realizować szereg zadań, wśród których najistotniejszymi są:

- komunikacja z zewnętrznymi heterogenicznymi źródłami danych – rozwiązanie powinno umożliwiać pobranie danych i informacji zapisanych w zewnętrznych repozytoriach. Należy wziąć pod uwagę różnorodność formatów danych i informacji i dostosować projektowane rozwiązanie do poszczególnych systemów, z których wybierane będą zasoby w trakcie eksploatacji,
- przetwarzanie zapytania wygenerowanego w języku zapytań wybranym dla projektowanego systemu integracyjnego na poszczególne języki zapytań implementowane w źródłach danych. Integrowane zasoby informacyjne mogą pochodzić z systemów opartych na różnych modelach danych z zaimplementowanymi różnymi językami zapytań. Dla każdego takiego systemu składowego należy utworzyć konwerter zapytań, którego zadaniem będzie reformułowanie kwerend na postać zrozumiałą w tym systemie. Problem ten może zo-

- stać rozwiązany globalnie – wtedy w systemie istnieje moduł umożliwiający konwersję zapytania na dowolny język zapytań lub lokalnie – wtedy w poszczególnych podsystemach występują konwertery dedykowane,
- łączenie i ujednolicenie zasobów informacyjnych pobranych z różnorodnych źródeł – dane i informacje otrzymane z systemów źródłowych, nawet po przetworzeniu do jednolitego formatu (np. XML), wymagają połączenia przed przesłaniem ich użytkownikom,
 - udostępnienie metainformacji o zasobach współdzielonych – należy dokonać wyboru modelu danych, który zostanie wykorzystany do opisu metadanych o współdzielonych zasobach w sposób jednolity. Metadane umożliwiają jednolite spojrzenie na integrowane zasoby. Ze względu na powszechne obecnie wykorzystanie w tym celu XML, do opisu danych współdzielonych należy rozważyć wybór jednego z modeli danych opartych na tej technologii. Umożliwi to współpracę i wymianę zasobów z różnymi systemami, zarówno integracyjnymi jak i innymi repozytoriami, o ile tylko systemy takie będą posiadały wspomaganie dla tej technologii,
 - umożliwienie dostępu dla różnych grup użytkowników – uniwersalne rozwiązanie pozwala na korzystanie z zasobów współdzielonych zarówno przez zdefiniowanych jak i niezdefiniowanych *a priori* użytkowników, a także przez różnorodne aplikacje w postaci agentów informacyjnych. Projektowane rozwiązanie powinno uwzględniać tę różnorodność uczestników i zawierać wsparcie dla każdego z nich, m.in. tekstowy i graficzny interfejs użytkownika, zapytania predefiniowane, interfejsy dla aplikacji,
 - podział zadania integracji na niezależne podzadania, aby modyfikacja jednej warstwy systemu lub jednego z jego modułów nie wymagała konieczności przeprojektowania pozostałych warstw lub modułów,
 - posiadanie różnorodnych form prezentacji zasobów informacyjnych, które mogą mieć niejednorodną formę. Wielorakie metody prezentacji umożliwiają zarówno różne spojrzenie na te same dane i informacje, jak i dostosowanie metody prezentacji do ich rodzajów (inne formy prezentacji są konieczne do przedstawienia plików zwartych a inne do danych pobranych z baz danych).

W związku z tym przygotowanie rozwiązania umożliwiającego wyszukiwanie zasobów w rozproszonych systemach heterogenicznych jest znacznie bardziej złożone, niż w rozwiązaniach tradycyjnych, gdyż wymaga zdefiniowania modułów realizujących te zadania.

Biorąc pod uwagę zadania, jakie muszą realizować systemy integracyjne dedykowane współdzieleniu zasobów informacyjnych w organizacjach sieciowych, można zidentyfikować pożądane własności tego typu rozwiązań:

- łączą autonomiczne podsystemy działające na dotychczasowych zasadach bez konieczności ich modyfikacji, dzięki czemu wdrożenie rozwiązania integracyjnego nie jest newralgiczne z punktu widzenia działania organizacji,

- dostarczają użytkownikom niezbędnych, aktualnych lub historycznych informacji,
- proces wyszukiwania jest przezroczysty dla użytkowników,
- posiadają elastyczną współdzieloną ontologię, dzięki czemu raz zdefiniowany moduł metadanych nie musi być przeprojektowywany w przyszłości, możliwe jest też automatyczne modyfikowanie metainformacji o dostępnych zasobach,
- umożliwiają nie tylko predefiniowanie zapytań na etapie projektowania, ale także pozwalają na formułowanie kwerend *ad hoc* przez niezdefiniowanych uprzednio użytkowników, przy czym sposób formułowania zapytań powinien być możliwie prosty, co pozwoli każdej osobie, także nieznającej języka zapytań, na dostęp do współdzielonych zasobów,
- posiadają możliwie ekspresyjny język zapytań dla użytkowników, co pozwala na precyzyjne formułowanie kwerend wybierających poszukiwane zasoby informacyjne,
- umożliwiają integrację wszystkich rodzajów zasobów, a więc nie tylko baz danych ale także dokumentów, multimediów i innych plików o przeszukiwanej i nieprzeszukiwanej strukturze,
- mogą być zastosowane do integracji wewnątrzorganizacyjnej oraz do łączenia zasobów przechowywanych w systemach partnerów biznesowych,
- są rozszerzalne, co pozwala na dołączanie nowych repozytoriów w trakcie eksploatacji, w tym klientów i partnerów biznesowych, przy jednoczesnym zachowaniu możliwości ukrycia dostępu do danych poufnych,
- są zorganizowane w taki sposób, aby co najmniej nie obniżać efektywności działania łączonych podsystemów,
- są łatwe do zaprojektowania i konfiguracji,
- posiadają warstwową/modułową architekturę, co w konsekwencji ma pozwalać na modyfikację poszczególnych modułów w trakcie eksploatacji systemu bez konieczności dokonywania zmian w innych jego elementach.

Systemy posiadające wszystkie te cechy niełatwo zbudować, gdyż część z postulowanych własności pozostaje ze sobą w sprzeczności (np. ekspresyjność języka zapytań i jego prostota, złożoność systemu i łatwość jego projektowania), podejmując więc próby zdefiniowania ram dla modelu najbardziej zbliżonego do ideału niezbędne jest osiągnięcie szeregu kompromisów. Przed przystąpieniem do opracowania takiego systemu należy dokonać identyfikacji zadań, które muszą być wykonane podczas realizacji przedsięwzięcia, zdefiniować jego architekturę oraz protokoły komunikacyjne dla poszczególnych komponentów. Oprócz złożoności, do wad takiego systemu należy zaliczyć trudności w utrzymaniu aktualnych metainformacji o dostępnych zasobach informacyjnych, konieczność przeprowadzenia części procesu projektowania w gronie wszystkich uczestników, czy potrzebę ustalenia standardów, co może być utrudnione gdyż uczestnicy systemu prowadzą różnorodną działalność, a także ze względów językowych i kulturowych.

PORÓWNANIE WŁASNOŚCI WYBRANYCH MODELI INTEGRACYJNYCH

W literaturze opisano szereg modeli integracyjnych umożliwiających łączenie różnorodnych zasobów informacyjnych (por. tabela 2).

Do analizy wybrano trzynaście propozycji, które porównano według sześciu kryteriów:

- wykorzystywany język zapytań do formułowania potrzeb informacyjnych,
- architektura,
- metoda opisu metadanych,
- przeznaczenie,
- użyty model danych,
- metoda komunikacji z zewnętrznymi źródłami przechowującymi integrowane zasoby informacyjne.

Proponowane modele integracyjne charakteryzują się dużą różnorodnością. Dla większości zaproponowano dedykowany język zapytań (TSIMMIS, SOEM, SEAL i KAON, Strudel, SBA, Lore, Ozone, iMeMex), a w pozostałych przypadkach problem ten został pominięty (Net-Centric, CO, OntoPIM, WO) lub wykorzystano język SQL (PowerCenter). Modele mają różnorodną architekturę, głównie wielowarstwową (SEAL i KAON, SBA, Lore, WO, iMeMex, PowerCenter), przy czym przeważa architektura trójwarstwowa, ewentualnie uzupełniona o interfejs użytkownika. Kilka modeli ma budowę modułową (TSIMMIS, SOEM, Net-Centric, Ozone, CO), co oznacza, że system składa się z aplikacji realizujących zadania składające się na proces integracji i korzystających w tym celu z dedykowanych repozytoriów danych lub metadanych. Z kolei trójwarstwowa architektura OntoPIM jest wykorzystywana do reprezentacji zasobów informacyjnych i dostępna poprzez szereg modułów realizujących określone zadania podczas opracowywania systemu i, następnie, integracji zasobów.

W zakresie wykorzystywanych metod opisu metadanych występuje duża różnorodność. W niektórych propozycjach problem ten, kluczowy z punktu widzenia całego projektu integracyjnego, został potraktowany marginesowo (TSIMMIS, SOEM, Lore, PowerCenter), choć w większości tych projektów w charakterze modelu danych wykorzystany został OEM, co sugeruje, że schemat metadanych może być zdefiniowany w postaci ontologii, do opisu której użyty zostanie ten właśnie model danych. W pozostałych przypadkach do stworzenia jednolitego dostępu do integrowanych zasobów wykorzystywane są różnorodne mechanizmy: baza danych współdzielonych, w której gromadzone są integrowane zasoby oraz rejestry i katalogi metadanych (Net-Centric), ontologie (SEAL i KAON, OntoPIM) lub zbiór połączonych ontologii (SBA, WO), struktury grafowe (Strudel, iMeMex), klasyczny model obiektowy (Ozone) oraz połączone encje (CO). Tak duża różnorodność sprawia, że połączenie rozwiązań opartych na poszczególnych modelach integracyjnych może sprawiać duży pro-

blem i wymagać będzie tworzenia mechanizmów mediacyjnych, czyli kolejnej warstwy nad istniejącymi rozwiązaniami integracyjnymi.

Tabela 2. Własności wybranych modeli integracyjnych

Model \ Własność	Język zapytań	Architektura	Metoda opisu metadanych	Zastosowanie	Model danych	Metoda komunikacji z zewnętrznymi źródłami
TSIMMIS ¹¹	OEM-QL	cztery rodzaje komponentów	brak	integracja różnorodnych zasobów	OEM	translatory, mediatory
Oparty na OEM (SO-EM) ¹²	OEM-QL	modułowa	brak	integracja różnorodnych zasobów	OEM	translatory, mediatory
Net-Centric ¹³	bd	modułowa	baza danych, katalog metadanych	integracja różnorodnych zasobów	bd	dedykowane aplikacje
SEAL i KA-ON ¹⁴	QEL	trójwarstwowa oraz interfejs użytkownika	ontologia	integracja różnorodnych zasobów internetowych	OEM lub RDF	wrappery, mediatory
Strudel ¹⁵	StruQL	mediacyjna	graf danych	automatyczne generowanie witryn internetowych	OEM	wrappery, mediatory
SBA ¹⁶	SBQL	warstwowa / modułowa	lokalne schematy	integracja różnorodnych zasobów	obiek-towy	protokoły kom., perspektywy
Lore ¹⁷	Lorel	trójwarstwowa	brak	integracja danych semistrukturalnych	OEM	wrappery
Ozone ¹⁸	OQL ^S	modułowa	zgodna z modelem obiekto-wym	integracja danych obiektowych i semistrukturalnych	OEM	za pomocą obiektowego języka zapytań
CO ¹⁹	bd	modułowa	encje	integracja danych semistrukturalnych	RDF	dodatkowe komponenty

¹¹ S. Chawathe i in., *The TSIMMIS Project: Integration of Heterogeneous Information Sources*, Proceedings of IPSJ Conference, 7–18, Tokio 1994.

¹² Y. Papakonstantinou, H. Garcia-Molina, J. Widom, *Object Exchange across Heterogeneous Information Sources*, Proceedings of the IEEE International Conference on Data Engineering, Taipei Taiwan 1995, s. 251–260.

¹³ J.P. Stenbit, *DoD Net-centric Data Strategy*, Department of Defense, Washington 2003.

¹⁴ A. Maedche i in., *SEAL – Tying up Information Integration and Web Site Management by Ontologies*, Technical Report, Institute AIFB, University of Karlsruhe, Karlsruhe 2002.

¹⁵ M. Fernandez i in., *Catching the Boat with STRUDEL: Experiences with a Web-Site Management System*, SIGMOD Conference, 1998, s. 414–425.

¹⁶ H. Kozankiewicz, K. Stencel, K. Subieta, *Intelligent Data Integration Middleware Based on Updateable Views*, Springer, 3490/2005, Berlin / Heidelberg 2005, s. 29–39.

¹⁷ J. McHugh i in., *Lore: A Database Management System for Semistructured Data*, SIGMOD Record, 26 (3), 1997, s. 54–66.

¹⁸ T. Lahiri, S. Abiteboul, J. Widom, *Ozone: Integrating Structured and Semistructured Data*, Proceedings of the Seventh International Conference on Database Programming Languages, Kinloch Rannoch, Szkocja 1999, s. 297–323.

¹⁹ T. Novotný, *A Content-Oriented Data Model for Semistructured Data [w:] Proceedings of the Dateso 2007 Workshop*, red. J. Pokorný, V. Snášel, K. Richta, Amphora Research Group, Czechy 2007, s. 55–66.

Model \ Własność	Język zapytań	Architektura	Metoda opisu meta-danych	Zastosowanie	Model danych	Metoda komunikacji z zewnętrznymi źródłami
OntoPIM ²⁰	bd	trójwarstwowa, dostęp poprzez moduły	ontologia	integracja różnorodnych zasobów personalnych	bd	wrappery pierwszego i drugiego poziomu
System oparty na ontologiach (WO) ²¹	bd	trójwarstwowa i informacje personalne	zbiór ontologii	integracja różnorodnych zasobów personalnych	RDF	bd
iMeMex ²²	iQL	trójwarstwowa/modułowa	graf iDM	integracja zasobów personalnych	iDM	poprzez moduł pośredniczący
PowerCenter ²³	SQL	warstwowa	bd	integracja fizyczna i wirtualna	bd	bd

Źródło: opracowanie własne.

Omówione modele mają także różnorodne zastosowanie:

- integracja zasobów internetowych i realizacja idei sieci semantycznej (SEAL i KAON),
- automatyczne generowanie witryn internetowych prezentujących różnorodne zasoby informacyjne pochodzące z wielu źródeł (Strudel),
- integracja danych semistrukturalnych (Lore, CO) oraz semistrukturalnych i obiektowych poprzez dodanie do modelu obiektowego dodatkowych funkcjonalności umożliwiających takie łączenie tych rodzajów zasobów (Ozone),
- integracja różnorodnych zasobów: plików zwartych, danych opartych na schemacie, danych semistrukturalnych (TSIMMIS, SOEM, Net-Centric, SBA),
- integracja zasobów personalnych, gromadzonych w pamięciach masowych użytkowników (OntoPIM, WO, iMeMex),
- integracja danych fizycznych i wirtualnych (PowerCenter).

W większości opisanych rozwiązań w charakterze modelu danych wykorzystywanego do opisu integrowanych zasobów proponuje się użycie OEM (TSIMMIS, SOEM, SEAL i KAON, Strudel, Lore, Ozone), rzadziej RDF (SEAL i KAON, CO, WO). W modelu iMeMex zaproponowano dedykowany model danych iDM, w SBA wykorzystano model obiektowy, a w trzech przypadkach nie sprecyzowano tej własności (Net-Centric, OntoPIM, PowerCenter). Do

²⁰ V. Katifori i in., *OntoPIM: How to Rely on a Personal Ontology for Personal Information Management*, Proceedings of the First Workshop on the Semantic Desktop, International Semantic Web Conference, Galway Irlandia 2005.

²¹ H. Xiao, I.F. Cruz, *A Multi-Ontology Approach for Personal Information Management*, Proceedings of 4-th International Semantic Web Conference, 6–10.11.2005, Galway Irlandia 2005.

²² J.P. Dittrich, M.A. Vaz Salles, *iDM: A Unified and Versatile Data Model for Personal Dataspace Management*, 32nd VLDB Conference, Seul Korea 2006, s. 367–378.

²³ *Informatica PowerCenter Data Federation Option*, Informatica – The Data Integration Company, Technical Report, USA 2008, http://www.informatica.com/INFA_Resources/ds_data_federation_6667.pdf (dostęp 23.03.2009).

komunikacji z zewnętrznymi źródłami danych zaproponowano *wrappery* komunikacyjne (SEAL i KAON, Strudel, Lore, OntoPIM) umożliwiające łączność z różnorodnymi systemami bazodanowymi lub innymi repozytoriami. Spośród innych rozwiązań zastosowano translatory konwertujące zapytania przed skierowaniem ich do poszczególnych repozytoriów (TSIMMIS, SOEM), mediatory, których zadaniem jest ustalenie jednolitego formatu zasobów otrzymanych ze źródeł zewnętrznych (TSIMMIS, SOEM, SEAL i KAON, Strudel), biblioteki wspomagające komunikację pomiędzy poszczególnymi modułami (SOEM), moduły zarządzania spójnością, których zadaniem jest zapewnienie jednorodności łączonych zasobów (TSIMMIS), modyfikowalne perspektywy (SBA), czy, jak w systemie Ozone, obiektowy język zapytań. W kilku przypadkach do rozwiązania tego problemu wykorzystano więcej niż jeden mechanizm, dzieląc w ten sposób zadanie łączności ze źródłami na oddzielne podzadania.

PODSUMOWANIE

Rozwój technologii teleinformatycznych jest podstawą tworzenia nowych form organizacyjnych, w tym struktur sieciowych charakteryzujących się m.in. czasowością istnienia oraz dynamiką zmian liczby wchodzących w ich skład jednostek. Jednym z istotnych warunków działania tego typu organizacji jest współdzielenie i wymiana danych i informacji gromadzonych przez poszczególne podmioty składowe. Wymaga to użycia rozwiązań charakteryzujących się własnościami, których nie posiadają tradycyjne systemy gromadzenia danych oparte zwykle na modelu relacyjnym, które umożliwią zarządzanie zasobami informacyjnymi i dostęp do danych i informacji gromadzonych w poszczególnych jednostkach wchodzących w skład organizacji sieciowych. W artykule scharakteryzowano struktury sieciowe oraz zidentyfikowano własności jakie powinny posiadać systemy integracyjne wspomagające zarządzanie zasobami informacyjnymi w tego typu organizacjach. Porównano istniejące rozwiązania zwracając przede wszystkim uwagę na występujące między nimi różnice. Z jednej strony ta różnorodność umożliwia wybór rozwiązania dostosowanego do własnych potrzeb, z drugiej strony utrudnia łączenie systemów integracyjnych implementowanych w poszczególnych jednostkach.

Wydaje się, że, ze względu na rozpowszechnienie XML coraz więcej systemów integracyjnych będzie opartych na ontologiach tworzonych w oparciu o tę technologię. Projektując odpowiedni system należy więc nie tylko dopasować go do własnych potrzeb informacyjnych, ale także mieć na uwadze potencjalną konieczność dołączania nowych jednostek organizacyjnych w przyszłości, a wykorzystanie w charakterze opisu metainformacji modułu opartego na ontologii zdefiniowanej w technologii XML z pewnością ułatwi realizację tego zadania.

LITERATURA

- Barczak B., Bartusik K., Kozina A., *Modele strukturalne organizacji uczącej się* [w:] *Doskonalenie struktur organizacyjnych przedsiębiorstw w gospodarce opartej na wiedzy*, red. A. Stabryła, Wydawnictwo C.H. Beck, Warszawa 2009.
- Binsztok A., *Organizacja fraktalna* [w:] *Zarządzanie przedsiębiorstwem w turbulentnym otoczeniu. Ku superelastycznej organizacji*, red. R. Krupski, PWE, Warszawa 2005.
- Chawathe S. i in., *The TSIMMIS Project: Integration of Heterogeneous Information Sources*, Proceedings of IPSJ Conference, Tokio 1994.
- Dittrich J.P., Vaz Salles M.A., *iDM: A Unified and Versatile Data Model for Personal Dataspace Management*, 32nd VLDB Conference, Seul Korea 2006.
- Dziembek D., *Struktura podmiotowa organizacji wirtualnej* [w:] *Technologie i systemy komunikacji oraz zarządzania informacją i wiedzą*, L. Kiełtyka, Difin, Warszawa 2008.
- Fernandez M. i in., *Catching the Boat with STRUDEL: Experiences with a Web-Site Management System*, SIGMOD Conference, 1998.
- Informatica PowerCenter Data Federation Option*, Informatica – The Data Integration Company, Technical Report, USA 2008, http://www.informatica.com/INFA_Resources/ds_data_federation_6667.pdf (dostęp 23.03.2009).
- Jaruga A., *Struktury sieciowe przyczynkiem kreowania wirtualnej organizacji przedsiębiorstwa* [w:] *Zarządzanie organizacjami gospodarczymi. Koncepcje i metody*, red. J. Lewandowski, t. 1, Wydawnictwo Politechniki Łódzkiej, Łódź 2005.
- Katifori V. i in., *OntoPIM: How to Rely on a Personal Ontology for Personal Information Management*, Proceedings of the First Workshop on the Semantic Desktop, International Semantic Web Conference, Galway Irlandia 2005.
- Korenik S., *Dysproporcje w rozwoju rejonów Polski – wybrane aspekty*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław 2003.
- Kozankiewicz H., Stencel K., Subieta K., *Intelligent Data Integration Middleware Based on Updateable Views*, Springer, 3490/2005, Berlin / Heidelberg 2005.
- Lahiri T., Abiteboul S., Widom J., *Ozone: Integrating Structured and Semistructured Data*, Proceedings of the Seventh International Conference on Database Programming Languages, Kinloch Rannoch, Szkocja 1999.
- Maedche A. i in., *SEAL – Tying up Information Integration and Web Site Management by Ontologies*, Technical Report, Institute AIFB, University of Karlsruhe, Karlsruhe 2002.
- Matusiak K.B. (red.), *Ośrodki innowacji w Polsce. Analiza krajowych instytucji wspierających innowacyjność i transfer technologii*, PARP, SOOIPP, Warszawa 2005.
- McHugh J. i in., *Lore: A Database Management System for Semistructured Data*, SIGMOD Record, 26 (3), 1997.
- Mikuła B., Pietruszka-Ortyl A., *Organizacje sieciowe*, Zeszyty Naukowe Akademii Ekonomicznej w Krakowie, Prace z zakresu zachowań organizacyjnych, nr 715, Kraków 2006.
- Novotný T., *A Content-Oriented Data Model for Semistructured Data* [w:] *Proceedings of the Dateso 2007 Workshop*, red. J. Pokorný, V. Snášel, K. Richta, Amphora Research Group, Czechy 2007.

- Papakonstantinou Y., Garcia-Molina H., Widom J., *Object Exchange across Heterogeneous Information Sources*, Proceedings of the IEEE International Conference on Data Engineering, Taipei Taiwan 1995.
- Perechuda K. (red.), *Zarządzanie wiedzą w przedsiębiorstwie*, WN PWN, Warszawa 2005.
- Stenbit J P., *DoD Net-centric Data Strategy*, Department of Defense, Washington 2003.
- Xiao H., Cruz I F., *A Multi-Ontology Approach for Personal Information Management*, Proceedings of 4-th International Semantic Web Conference, 6–10.11.2005, Galway Irlandia 2005.

Streszczenie

Celem artykułu jest omówienie problemu współdzielenia zasobów informacyjnych w środowisku dynamicznym, w którym działają organizacje sieciowe oraz istniejących rozwiązań w tym zakresie. Omówiono własności i rodzaje organizacji sieciowych zwracając uwagę na różnice między tego typu formami organizacyjnymi a organizacjami tradycyjnymi. Scharakteryzowano zagadnienie współdzielenia danych w systemach opartych na centralnej bazie danych i w tych opartych na współdzielonej ontologii. Zidentyfikowano pożądane własności systemów integracyjnych oraz porównano własności kilkunastu rozwiązań dedykowanych zarządzaniu rozproszonymi i heterogenicznymi zasobami informacyjnymi w dynamicznym środowisku sieciowym.

Data integrating systems for contemporary networking organisations

Summary

The aim of this paper is to discuss the problem of information resources sharing in dynamic environment, in which network organisations operate as far as existing solutions in this area. The nature of network organisations is discussed, their types are described and their characteristics are compared with the features of traditional organisations. The problem of data sharing in systems based on central databases as well as in solutions based on co-shared ontologies is discussed. The most important attributes of integrating systems are also identified and over a dozen solutions dedicated to the integration of distributed and heterogeneous information resources in dynamic network environment are talked over.