

І.І. Кушнірецька², О.І. Кушнірецька², А.Ю. Берко¹
Національний університет "Львівська політехніка",
¹кафедра загальної екології та екоінформаційних систем,
²кафедра інформаційних систем та мереж

ЗАСТОСУВАННЯ ОНТОЛОГІЙ І МЕТАМОДЕЛЕЙ ДЛЯ ДИНАМІЧНОЇ ІНТЕГРАЦІЇ СЛАБОСТРУКТУРОВАНИХ ДАНИХ

© Берко А.Ю., Кушнірецька І.І., Кушнірецька О.І., 2014

Описано можливості використання онтологій і метамodelей для створення системи динамічної інтеграції слабоструктурованих даних. Описано процес перетворення слабоструктурованих даних на структуровану інформацію. Розглянуто динамічну інтеграцію слабоструктурованих даних на різних рівнях складності: фізичному, логічному і глобальному.

Ключові слова: динамічна інтеграція, онтологія, метамodelь, web-система.

This paper reviews the opportunities of the use of ontologies and meta-models for the creation of the system of dynamic integration of semistructured data. The process of the converting semistructured data into the structured information has been described. The dynamic integration of semistructured data at different levels of complexity (physical, logical and global) has been considered.

Key words: dynamic integration, ontology, meta-model, web-system.

Вступ. Загальна постановка проблеми

Постійні інновації, швидкий розвиток суспільного прогресу, відкриття все нових і нових технологічних новинок зумовлюють активне застосування web-систем у різноманітних сферах людської діяльності. Використання такого типу інформаційних ресурсів дає можливість швидкого пошуку потрібної інформації, не виходячи з дому чи офісу. Сьогодні внаслідок постійного розвитку інформаційних технологій у web-системах накопичено і продовжують швидко зростати досить значні обсяги слабоструктурованих даних різноманітного характеру. Виникає необхідність їх оперативної, динамічної інтеграції для зручності подання цих даних та подальшого використання.

Завдання інтеграції даних полягає в з'єднанні даних з різних джерел і наданні користувачеві єдиного (уніфікованого) подання цих даних, зокрема можливості виділення цікавої користувачеві інформації за запитом. Система інтеграції даних дає змогу звільнити користувача від необхідності самостійно відбирати джерела, в яких знаходиться потрібна користувачеві інформація, звертатися до кожного джерела окремо і вручну зіставляти і об'єднувати дані з різних джерел.

Зв'язок висвітленої проблеми із важливими науковими та практичними завданнями

Роль інтеграції даних зростає, коли збільшується обсяг даних і виникає необхідність спільно їх використовувати. Це стало фокусом широкої теоретичної роботи, а численні проблеми залишаються невирішеними. Проблема динамічної інтеграції даних надзвичайно багатоаспектна й різноманітна. Складність і характер використовуваних методів її вирішення істотно залежать від рівня інтеграції, який необхідно забезпечити, властивостей окремих джерел даних і всієї сукупності джерел та необхідних способів інтеграції.

Аналіз останніх досліджень та публікацій

Інтеграція даних в інформаційних системах розуміється як забезпечення єдиного уніфікованого інтерфейсу для доступу до деякої сукупності даних, які отримано з неоднорідних незалежних джерел [5, 6]. Отже, для користувача інформаційні ресурси всієї сукупності інтегрованих джерел утворюють єдиний сервіс. Система, що забезпечує користувачеві такі можливості, називається системою інтеграції даних.

Система інтеграції даних звільняє користувачів від необхідності знання, дані з яких джерел, крім інтегрованого, вони використовують, які властивості цих джерел і як здійснити доступ до них. Інтегрованими джерелами даних можуть бути традиційні системи баз даних, що підтримують різні моделі даних (реляційні, об'єктні, об'єктно-реляційні, графові і т.п.), різноманітні успадковані системи, репозиторії, web-сайти, файли структурованих даних. Забезпечення доступу до даних багатьох джерел через єдиний інтерфейс означає фактично, що йдеться про підтримку представлення сукупності даних з безлічі незалежних джерел в термінах єдиної моделі даних. Склад значної кількості джерел даних може бути наперед заданим або динамічно поповнюваним, джерела даних можуть володіти незмінним або оновлюваним змістом.

Проблемою динамічної інтеграції інформаційних ресурсів наукові та інженерно-технічні спільноти світу переймаються вже багато років, але тільки в останні роки поява і активне розвинення сучасних інфраструктур (WEB та GRID середовища) і відкритих сервіс-орієнтованих архітектур у галузі інформаційних технологій (SOA, OGSA), а також значні просування в розробці відповідних міжнародних базових стандартів обміну інформацією (XML, RDF, TM, OWL) дозволяють створювати принципово нові моделі ІС. Такі моделі дають змогу будувати глобально розподілені застосування, що реалізують технологічні ланцюжки, в яких можуть використовуватися не тільки власні ІР, а й ті, що можуть бути запропоновані іншими організаційними структурами [3]. При цьому беруть до уваги те, що під час роботи з таким застосуванням можливо замінити один інформаційний сервіс іншим, видаляти ті, що втратили цінність або актуальність, та додавати нові.

Системи інтеграції даних можуть забезпечувати інтеграцію даних на фізичному, логічному і глобальному рівнях. Інтеграція даних на фізичному рівні з теоретичного погляду є найпростішим завданням і зводиться до конверсії даних з різних джерел у необхідний єдиний формат їх фізичного подання [2]. Інтеграція даних на логічному рівні передбачає можливість доступу до даних, що містяться в різних джерелах, в термінах єдиної глобальної схеми, яка описує їхнє спільне подання з урахуванням структурних і, можливо, поведінкових (при використанні об'єктних моделей) властивостей даних. Семантичних властивостей даних при цьому не враховуються. Підтримку єдиного представлення даних з урахуванням їхніх семантичних властивостей у контексті єдиної онтології предметної області забезпечує інтеграція даних на глобальному рівні.

Джерела даних можуть володіти різними властивостями, істотними для вибору методів інтеграції даних: вони можуть підтримувати подання даних у термінах тієї чи іншої моделі даних, можуть бути статичними або динамічними тощо. Безліч джерел інтегрованих даних можуть бути однорідними або неоднорідними щодо характеристик, відповідних використовуваному рівню інтеграції.

Серед способів інтеграції даних розрізняють два підходи: з віртуальним і актуальним (матеріалізованим) поданням інтегрованих даних. За першим підходом створюється механізм доступу, який при обробці користувачького запиту породжує дані в необхідному поданні безпосередньо з джерел даних. Повне матеріалізоване уявлення інтегрованих даних у термінах єдиного користувальницького інтерфейсу при цьому не підтримується. Віртуальний підхід найчастіше застосовується при використанні часто оновлюваних джерел даних. За другим підходом на стадії інтеграції формується повне матеріалізоване уявлення інтегрованих даних, відчужене від вихідних джерел і співіснуюче з ними.

Процесу інтеграції перешкоджає неоднорідність джерел даних, відповідно до рівня інтеграції.

– при інтеграції на фізичному рівні в джерелах даних можуть використовуватися різні формати файлів;

– на логічному рівні інтеграції може виникати неоднорідність використовуваних моделей даних для різних джерел або можуть розрізнятися схеми даних, хоча використовується одна і та сама модель даних. Одні джерела можуть бути web-сайтами, а інші – об'єктними базами даних і т. д.

– при інтеграції на глобальному рівні різним джерелам даних можуть відповідати різні онтології. Наприклад, можливий випадок, коли кожне з джерел представляє інформаційні ресурси, що моделюють деякий фрагмент предметної області, якому відповідає своя понятійна система, і ці фрагменти перетинаються.

Виділяють такі проблеми інтеграції даних:

- різнорідність (гетерогенність): джерела даних використовують різні моделі (і навіть метамоделі);

- автономність: джерела розроблені і експлуатуються незалежно один від одного, незалежно спроектовані під вирішення конкретних, різних завдань, різними методами;

- розподіленість: джерела фізично або логічно доступні тільки через мережеві протоколи віддаленого доступу, зокрема інформаційні джерела можуть бути розподілені в мережі Інтернет.

Існують такі типи невідповідності схем даних:

- конфлікти неоднорідності (використовуються різні моделі даних для різних джерел);

- конфлікти іменування (у різних схемах використовується різна термінологія, що призводить до омонімії та синонімії в іменуванні);

- семантичні конфлікти (вибрано різні рівні абстракції для моделювання подібних сутностей реального світу);

- структурні конфлікти (одні і ті самі сутності подають у різних джерелах різними структурами даних).

До основних засобів, що використовуються для забезпечення інтеграції інформаційних ресурсів, належать конвертори даних, що інтегрують моделі даних, механізми відображення моделей даних, об'єктні адаптери (Wrappers), посередники (Mediators), онтологічні специфікації, засоби інтеграції схем та інтеграції онтологічних специфікацій [3, 4], а також архітектура, що забезпечує взаємодію засобів, що використовуються в конкретній системі інтеграції ресурсів.

Інтеграція з використанням онтологічних специфікацій дає змогу виділити дані, збережені в кожному джерелі даних, взаємно їх пов'язати з логічними і глобальними онтологіями за допомогою відображення.



Рис. 1. Систематизація знань в області онтологій

Онтологія – формальна специфікація поділяючої концептуальної моделі. Онтологія складається з класів сутностей предметної області, властивостей цих класів, зв'язків між цими класами і тверджень, побудованих на основі цих класів, їхніх властивостей і зв'язків між ними. На рис. 1 наведено галузі застосування, ролі, типи, мови представлення і власників онтологій [1].

Моделювання онтології конкретних предметних областей (domain ontologies) є дорогою і трудомісткою задачею, що вимагає участі експертів. Однак ця робота повинна бути виконана тільки один раз основним співтовариством, яке вміє розробляти і підтримувати словники за допомогою засобів спільної роботи, розроблених співтовариством Semantic Web.

Але також не тільки експерти можуть додавати і відображати (map) нові джерела даних за допомогою широко визнаних допоміжних онтологій, наприклад, Dublin Core, FOAF, SOAP, DOAP і т.д.

Для конкретної предметної області розумним вважається такий підхід: спочатку здійснити пошук існуючих онтологій, а потім заповнити розриви між ними на основі підходу “знизу догори” (bottom-up) від власних понять до загальніших (рис. 2) [3].



Рис. 2. Інтеграція даних на основі онтологій

Виділення проблем

При створенні системи інтеграції виникають завдання, склад яких залежить від вимог до неї і використовуваного підходу. До них зокрема належать розроблення архітектури системи інтеграції даних та створення інтегруючої моделі даних, яка є основою єдиного користувацького інтерфейсу в системі інтеграції. Важливим є питання розроблення методів відображення моделей даних та побудови відображень у інтегровальну модель для конкретних моделей, підтримуваних окремими джерелами даних, а також інтеграція метаданих, що використовуються в системі джерел даних.

Актуальними проблемами залишаються подолання неоднорідності джерел інформаційних ресурсів та розроблення механізмів динамічної і семантичної інтеграції слабоструктурованих даних у web-системах.

Формулювання мети

Мета роботи полягає у використанні існуючих технологій, щоб створити інструмент, який дає змогу кінцевим користувачам отримувати уніфіковану інформацію з різних гетерогенних джерел даних, забезпечуючи необхідні перетворення даних з використанням концепцій, зрозумілих користувачеві.

Аналіз отриманих наукових результатів

Однією з основних сфер, в яких застосовують принципи та технології динамічної інтеграції даних, є Internet-торгівля. Аналіз способів діяльності Internet-магазинів в Україні та за кордоном показує, що структура Internet-магазинів, як правило, є однотипною. Internet-магазини мають широку і багаторівневу ієрархічну класифікацію, і кожен клас має ряд продуктів, якими оперує. Класифікація варіюється від одного Internet-магазину до іншого. Розглядаючи класифікацію кожного з Internet-магазинів, не можна сказати, правильною чи неправильною вона є в тому чи іншому Internet-магазині, тому що кожен з них був розроблений на основі доволі логічних міркувань. Не уникнути проблеми гетерогенності при класифікації:

- Один і той самий клас має різні назви в різних місцях;
- Один і той самий клас ділиться на підкласи різними способами;
- Є кілька класів з одним і тим самим ім'ям.

Кожен продукт має ряд атрибутів, що описують його. Продукти в одній групі в одному магазині, як правило, мають певні атрибути, в той же час продукти різних груп можуть мати різні набори ознак. Один продукт може мати різні атрибути в різних Internet-магазинах:

- Один магазин представляє атрибут об'єкта, інший ні;
- Один і той самий атрибут ідентифікується по-різному;
- Один і той самий атрибут містить різні значення;
- Один і той самий атрибут містить різні види даних.

Інтеграція даних розглядається на різних рівнях складності (рис. 3) [4].

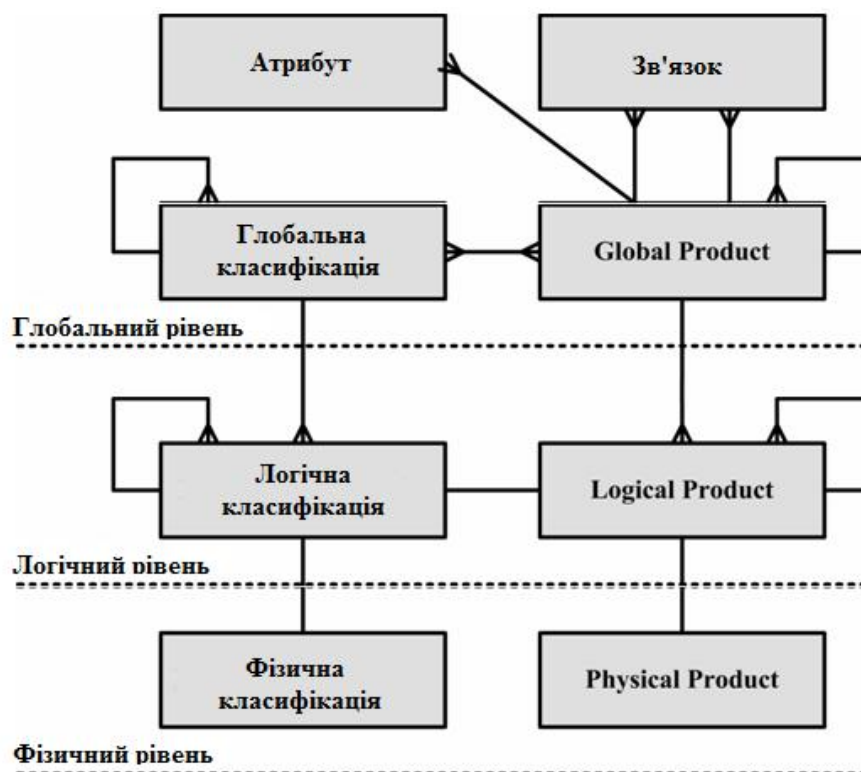


Рис. 3. Інтеграція даних на різних рівнях складності

Використання різних способів та методів інтеграції на кожному із рівнів, відповідно до складності рівня, значно спрощує процес інтеграції слабоструктурованих даних, особливо у web-системах.

На фізичному рівні, де всі дані з одного джерела, їх не перетворюють, а тільки розміщують у вигляді різних комбінацій або зображень. Ці функції виконує браузер бази даних. Другий рівень – логічний. Дані, як і раніше, отримуються з одного джерела, але за допомогою адаптерів. Адаптери (wrappers) – компоненти, пов'язані з джерелами даних для вирішення проблем технічної неоднорідності і неоднорідностей метамоделей. Їх основні функції: приймають запити до джерела деякою мовою, перетворюють запит мовою джерела, виконують запит, відправляють результати посереднику (Mediator). На посередника покладається завдання підтримки єдиного користувацького інтерфейсу на основі глобального представлення даних, що містяться в джерелах, а також підтримки відображення між глобальним і локальним уявленнями даних. Користувацький запит, сформульований в термінах єдиного інтерфейсу, ділиться на безліч підзапитів, адресованих до потрібних локальних джерел даних. На основі результатів їх обробки синтезується повна відповідь на запит. Засобами посередників підтримуються уніфіковані метаописання інтегрованих джерел даних. Як правило, семантичні посередники розробляються для конкретної вузької предметної області. Механізми посередників ґрунтуються на онтологічних специфікаціях джерел. Для посередника створюється інтегрована онтологія використовуваних джерел. У таких системах необхідна також інтегрована модель даних з розвиненими можливостями моделювання семантики даних.

Наступний рівень є рівнем інтеграції системи, де дані отримуються з безлічі джерел даних. Ці функції виконує універсальний браузер бази даних інформаційної системи. Він оснований на

принципі, що інформацію з кожного джерела даних отримано на основі принципу роботи логічного браузера бази даних. Інтеграції досягнуто за допомогою простої онтології, одна з яких об'єднує кілька найважливіших сутностей й атрибутів з різних джерел і вимагає переведення значення в єдиний формат.

На глобальному рівні (рівні інтеграції системи) пропонується використовувати Універсальний браузер структурованих даних [4], принципи роботи якого зображено на рис. 4.

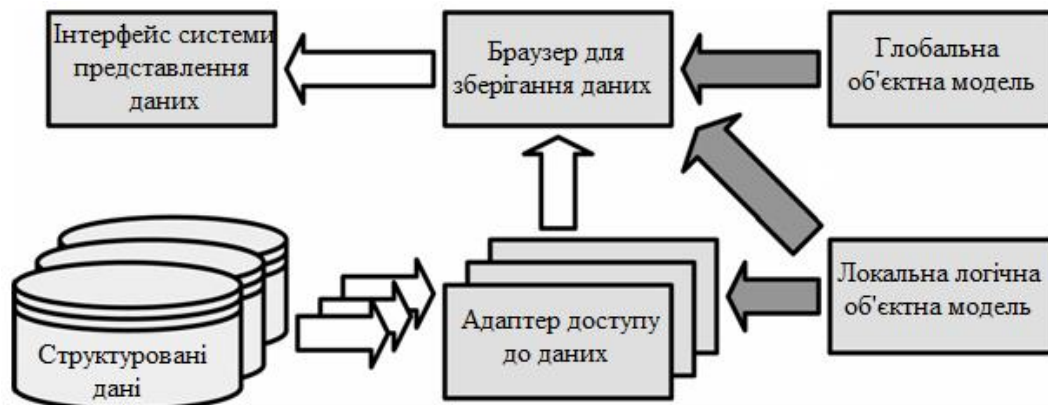


Рис. 4. Система роботи Універсального браузера структурованих даних

Запропонована технологія заснована на поширеному використанні локальних онтологій, що описують кожне джерело даних як частину всього процесу [4]. Також використовуються глобальні онтології, які пов'язують локальні онтології і відкривають різні види одних і тих самих даних. Використовується також принцип сховища даних – слабоструктуровані дані конвертуються і зберігаються в інструментальній базі даних як структурована інформація. Це дає змогу інтегрувати інформацію та переглядати її на основі вже протестованої інформаційної системи Універсального браузера бази даних.

Застосування Універсального браузера структурованих даних надзвичайно важливо саме у web-системах, адже навіть структурована інформація, яка взята з Інтернету, не завжди правильно подана і зображена. Для того, щоб правильно подати користувачеві інформацію, отриману з декількох web-сайтів, її передусім потрібно правильно синтаксично і семантично оформити.

Інтеграція інформації відбувається на трьох рівнях. Фізичний рівень визначає структуру і синтаксис інформації, яка була поміщена в джерело. Існує позначення тих компонент, що містять дані, які нам потрібні. Визначається структура, за допомогою якої дані зберігаються (окремі одиниці інформації, списки, таблиці, картинки, гіперпосилання та зв'язки між цими різними елементами). Тут немає потреби розуміти семантику інформації. Дані, які було отримано, мають дуже загальний характер: рядок, URL, картинка.

Логічний рівень описує семантику джерела даних. Необхідна інформація береться з текстів. Вона перетворюється на потрібний тип даних, а потім пов'язується в логічній структурі. Робота на цьому рівні потрібна для того, щоб представити форму інформації в джерелі даних на найвищому рівні генералізації, зберігаючи логіку, яку було використано у джерелі даних. Водночас необхідно наблизитися до опису об'єктів даних, які було визначено на глобальному рівні, а також взаємозв'язку між цими стандартами.

На глобальному рівні описуються стандарти даних і стандарти класифікації (patterns). Тут метою є зв'язати воедино інформацію, яка зберігається в різних джерелах даних.

На додаток до описаної онтології можна також динамічно створювати віртуальні онтології, оскільки глобальний рівень забезпечує зв'язок онтологій, які існують на набагато нижчих рівнях, використовуючи різні аспекти.

Наприклад, метод потенціалу, який повинен бути перевірений на практиці, полягає в об'єднанні багатьох фізичних онтологій в єдину онтологію. Схожі вузли онтології зібрано разом, і всі назви залишаються на місці – так, щоб вони могли бути ідентифіковані. Це схоже на граф, і

користувачам потрібно надати спосіб легко маневрувати через нього так, щоб знайти відповідний вузол.

Як вже було зазначено, інтегрувати слабоструктуровані дані потрібно на трьох рівнях. На верхньому, глобальному рівні дані інтегруємо за допомогою універсального браузера структурованих даних. Отже, перед використанням універсального браузера структурованих даних дані потрібно структурувати. Перетворення слабоструктурованих даних на структуровані – доволі трудомісткий процес, який вимагає високої точності, тому для роботи із слабоструктурованими даними використовується так званий універсальний браузер слабоструктурованих даних, який детально описано у [4] і зображено на рис. 5.

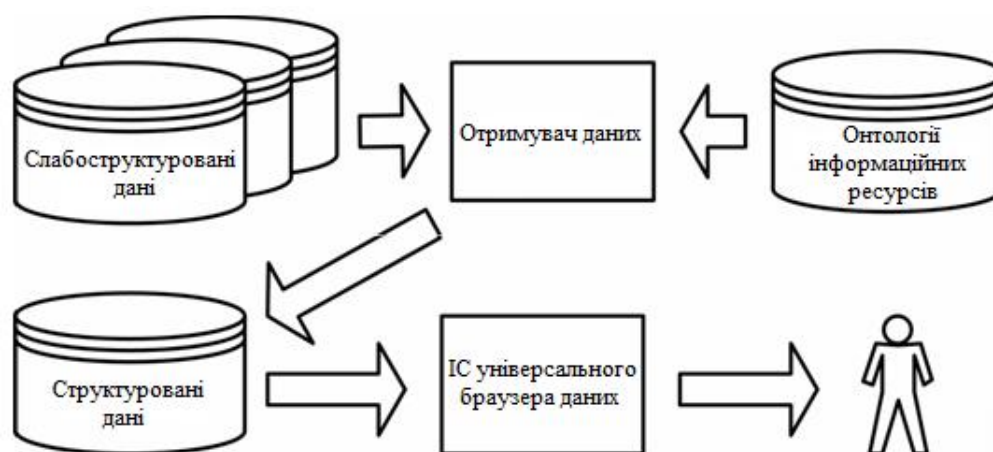


Рис. 5. Концептуальна модель універсального браузера слабоструктурованих даних

Універсальний браузер слабоструктурованих даних призначений для пошуку, аналізу, перекладу, зберігання, інтеграції та перегляду частково структурованих даних. Основна ідея системи полягає в перетворенні слабоструктурованих даних на структуровані дані. Багато різних типів онтологій використовуються для поєднання інформації. Джерелом частково структурованих даних, як правило, є web-сайти або деякі типи документів (PDF, DOC, XLS). Ці джерела даних, як правило, не пропонують доступу до функцій, які забезпечують доступ до конкретних даних. Зазвичай тільки можливо отримувати і аналізувати весь документ.

Інтеграція даних в інформаційній системі природним чином передбачає і інтеграцію в тій чи іншій формі метаданих, що визначають їх джерела. Одним із традиційних завдань інтеграції метаданих у системах інтеграції структурованих даних є завдання інтеграції схем. Труднощі його вирішення в конкретних ситуаціях можуть бути пов'язані з наявністю конфліктів, наприклад:

- Конфліктів неоднорідності (використовуються різні моделі даних для різних джерел);
- Конфліктів іменування (в різних схемах використовується різна термінологія, що призводить до омонімії та синонімії в іменуванні);
- Семантичних конфліктів (вибрані різні рівні абстракції для моделювання подібних сутностей реального світу);
- Структурних конфліктів (одні й ті самі сутності представляються в різних джерелах різними структурами даних).

Розглянемо роботу Універсального браузера слабоструктурованих даних і те, як для нашого випадку за його допомогою можна перетворювати слабоструктуровані дані на структуровані. Отже, для нас вхідною інформацією є певний клас слабоструктурованих даних. Усередині цього класу ми знаходимо дані, які складаються з двох частин: ієрархічної класифікації і даних конкретних об'єктів. Ці дані можуть бути частково структуровані. Як приклад візьмемо каталог продукції виробника або продавця.

Універсальний браузер структурованих даних глобального рівня не може працювати з частково структурованими даними, – вони повинні бути перетворені на структуровані дані. Загалом

це є проблемою, яку ще не було повністю вирішено. Попередньо ми структурували тільки ту інформацію, яка може бути структурована відносно легко – класифікація структурованої частини даних. Частина залишається неструктурованою, і пошук в ній можливий лише на підставі ключових слів (текст).

Як приклад використаємо каталог продукції. У цьому каталозі можна легко структурувати продукт за класифікацією і призначити конкретні властивості до конкретних продуктів.

Класифікатори даних є, по суті, прикладами онтологій (онтології класифікацій). Дані моделі можуть бути встановлені для джерел даних, які є по суті онтологіями (продукт онтології).

Отримані структуровані дані повинні зберігатися в базі даних. Класифіковані продукти, які отримуються з джерела даних, зберігаються в онтології цих класифікаторів.

Основна проблема полягає в тому, що структури продукту бази даних ніколи не відомо заздалегідь, і тому проблематично встановити базу даних. Тому дані про продукт зберігаються в певній універсальній структурі.

Універсальна структура означає, що продукт або його частина є сутністю з конкретним набором ознак. Кожен атрибут має безліч значень. Сутності можуть бути зв'язані відношеннями. Для того, щоб зберігати дані в універсальній структурі і отримувати їх з універсальної структури, ми використовуємо інформацію про модель даних, яка знаходиться в продукті онтологій.

Механізм для перетворення слабоструктурованих даних на структуровані зображено у роботі [4] (рис. 6).

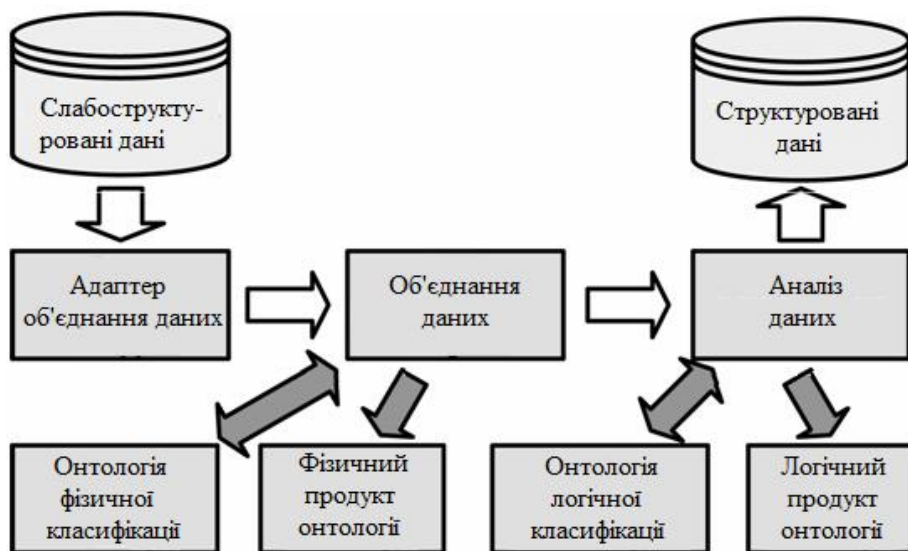


Рис. 6. Принцип перетворення слабоструктурованих даних

Покажемо, як цей механізм можна застосувати для роботи нашого Універсального браузера слабоструктурованих даних. Переважно кожне з джерел слабоструктурованих даних має інший формат. Сайти аналізуються за допомогою адаптерів, які можуть аналізувати web-сторінки і отримувати із них необхідну інформацію. Якщо це сторінка з класифікацією якогось продукту, то адаптер отримує цю класифікацію. Якщо сторінка містить загальну інформацію про продукт, то будуть отримані дані про конкретний продукт. Web-сайт зазвичай має головну сторінку, з якої різноманітні посилання можуть бути використані для того, щоб, наприклад, переглянути весь каталог інформації про класифікацію конкретного продукту, а також інформацію опису цього виробу.

Перед тим, як проаналізувати джерела даних і отримувати дані з них, необхідно проаналізувати web-сторінки і встановити онтології для локальних продуктів. Крім того, повинен бути встановлений адаптер, який може аналізувати конкретне джерело даних.

Заповнення і оновлення даних виконується за допомогою інформації про модель джерела даних, яка зберігається в продукті онтології. Потім відбувається процес запису точних даних у базі даних продукції. Це вже структуровані дані, працює з якими універсальний браузер структурованих

даних. Отже, основна задача полягає в тому, щоб перетворити слабоструктуровані дані на структуровані. Слабоструктуровані дані після того, як потрапляють до системи, перетворюються за допомогою адаптера об'єднання даних, який об'єднує дані, користуючись онтологією фізичної класифікації. Дані після перетворення переходять на стадію аналізу. Під час аналізу даних система універсального браузера слабоструктурованих даних за допомогою використання онтології логічної класифікації перетворює їх на вже структуровані дані, працює з якими універсальний браузер структурованих даних.

Робота із вже структурованими даними полягає в наступному: за допомогою адаптера отримуються дані з бази даних, а потім отримані дані передаються у браузер для зберігання даних, який потім подає дані для користувача. Браузер для зберігання даних далі отримує і передає користувацький запит до системи.

Для роботи із слабоструктурованими даними і перетворення їх на структуровані дані використовується система універсального браузера слабоструктурованих даних, яку зображено на рис. 6. Із вже структурованими даними працює система універсального браузера структурованих даних, яку зображено на рис. 4. Якщо у системі універсального браузера структурованих даних використовується тільки одна база метаданих, то система універсального браузера слабоструктурованих даних використовуватиме інформацію з декількох онтологій класифікації і продуктів онтологій для пошуку та перегляду даних.

Структурована база даних слугує джерелом даних для системи універсального браузера структурованих даних. Адаптер доступу до даних дає змогу системі отримати доступ до інформації, яка зберігається в структурованій базі даних.

У системі немає будь-яких конкретних функцій доступу, які дозволять отримати доступ до структурованої бази даних, тому адаптер інтерпретує дані з продукту онтології так, щоб генерувати динамічні запити для вибору інформації із бази структурованих даних.

Отже, основна ідея динамічної інтеграції слабоструктурованих даних у web-системах полягає в тому, що використовуються онтології та мета-моделі, а інтеграція відбувається на трьох рівнях складності. Робота web-системи, яка працює за принципом динамічної інтеграції слабоструктурованих даних, використовуючи онтології та мета-моделі, полягає в такому: користувач посилає запит через інтерфейс до системи, система аналізує запит і динамічно отримує дані з різноманітних web-джерел. Це слабоструктуровані дані, з якими надалі працює універсальний браузер слабоструктурованих даних, який перетворює їх на структуровану інформацію. Із вже структурованими даними працює універсальний браузер структурованих даних, який їх семантично і синтаксично оформлює і виводить вже оформлені дані через користувацький інтерфейс.

Висновки

Сформульовано основні положення динамічної інтеграції слабоструктурованих даних у web-системах з використанням онтологій і метамоделей. Показано функціонування універсального браузера динамічної інтеграції слабоструктурованих даних у web-системах. Розглянуто принципи перетворення слабоструктурованих даних на структуровану інформацію. Зображено інтеграцію даних на різних рівнях складності: фізичному, логічному і глобальному. Описано концептуальну модель універсальної структури зберігання даних. Принципи роботи універсального браузера, який описано у статті, можна застосувати при створенні web-систем динамічної інтеграції слабоструктурованих даних різноманітного характеру.

1. Гаврилова Т.А. *Онтологический инжиниринг // Сб. докладов Восьмой научно-практической конференции "Реинжиниринг бизнес-процессов на основе современных информационных технологий. Системы управления знаниями" (РБП-СУЗ-2005).* – М., 2005. – С. 79–82.
2. Козаловский М.Р. *Методы интеграции данных в информационных системах. Институт проблем рынка РАН // М.Р. Козаловский – М., 2010. – 74 с.*
3. Briukhov D.O., Shumilov S.S. *Ontology Specification and Integration Facilities in a Semantic Interoperation Framework. Proc. of the Second Intern. Workshop ADBIS'95. Moscow, 1995. - P. 195-200.*
4. Arnicans, G., Karnitis, G. *Intelligent*

Integration of Information From Semi-Structured Web Data Sources on the Basis of Ontology and Meta-Models, 7th International Baltic Conference on Databases and Information Systems Vilnius, Lithuania, 2006, P. 177 – 186. 5. Levy A.Y. *Logic-Based Techniques in Data Integration. Logic-based Techniques in Data Integration. In: Logic Based Artificial Intelligence. Edited by J. Minker. Kluwer Publishers, 2000.* 6. Manolescu I., Florescu D., Kossman D. *Answering XML Queries over Heterogeneous Data Sources. Proc. Of the 27th VLDB Conference, Roma, Italy, 2001.*

УДК 004.89

В.В. Литвин

Національний університет “Львівська політехніка”,
кафедра інформаційних систем та мереж

МЕТОД ВИКОРИСТАННЯ ОНТОЛОГІЙ У ПЕТЛІ OODA

© Литвин В.В., 2014

Досліджено поведінку інтелектуального агента у конкурентному середовищі. Для моделювання поведінки обрано петлю OODA. Розглянуто взаємодію етапів петлі OODA (спостереження, орієнтація, прийняття рішення, дія) із онтологією задач та предметної області, в межах якої функціонує цей агент.

Ключові слова: петля OODA, спостереження, орієнтація; аналіз; синтез; прийняття рішення; дія, онтологія.

In the paper the behavior of an intelligent agent in a competitive environment is investigated in the paper. The OODA loop is chosen for behavior simulation. The interaction of OODA loop stages (observation, orientation, decision support, action) with the ontology of tasks and subject area in which the agent operates was explored.

Key words: OODA loop, observation, orientation, analysis, synthesis, decision support, action, ontology.

Вступ. Постановка проблеми у загальному вигляді

Відповідно до ідей Джона Бойда та його послідовників будь-яку діяльність у конкурентному середовищі (наприклад, у військовій сфері) з певним ступенем наближення можна подати у вигляді кібернетичної моделі OODA [1]. Зазначена модель передбачає багаторазове повторення петлі дій, складеної з чотирьох послідовних взаємодіючих процесів (рис. 1.): спостереження (observation), орієнтація (orientation), прийняття рішення (decision), дія (action).

Така модель з успіхом почала застосовуватися для моделювання діяльності та прийняття рішень у бізнесі, політиці, соціології тощо, тобто у тих сферах, де наявна конкуруюча сторона.

Згідно з теорією Бойда, кожна людина або організація при вирішенні поставлених перед ними завдань має свою петлю прийняття рішень і діяльності.

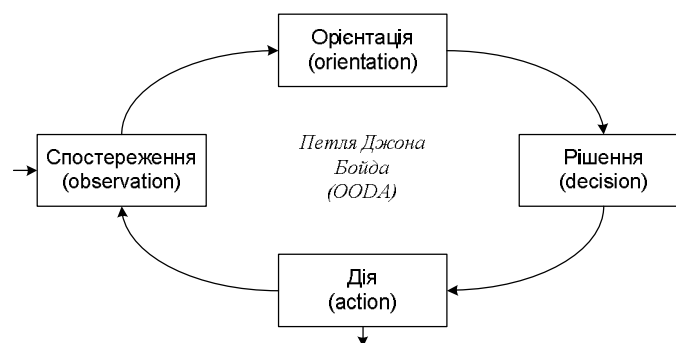


Рис. 1. Процеси петлі OODA