



ДИНАМІЧНЕ ФОРМУВАННЯ СТОРІНКИ WEB-САЙТУ НА ОСНОВІ ВИКОРИСТАННЯ ОНТОЛОГІЇ ПРЕДМЕТНОЇ ОБЛАСТІ

Піскова Н.О.

*Національний технічний університет
«Харківський політехнічний інститут»
Харків, вул. Пушкінська, 79/2, тел. 096-123-31-88
e-mail: piskova.n@mail.ru*

Сьогодні стрімко розвиваються веб-орієнтовані інформаційні системи, зокрема інформаційні портали, системи електронної комерції і т.д. Розвиваються засоби і технології для програмування веб-додатків як розподілених клієнт-серверних (JSP, .NET і т.п.).

Всі сайти в мережі можна розділити на дві великі групи: статичні і динамічні сайти. Статичної прийнято називати сторінку, яка цілком зберігається на сервері і показується відвідувачеві у своєму незмінному вигляді (статична сторінка може містити деякі змінювані елементи, наприклад банери, проте вона все одно залишається статичною) [1].

Динамічна сторінка – веб-сторінка, яку згенеровано програмно на відміну від статичної сторінки, які є файлами, що лежать на сервері. Сервер генерує HTML код для обробки браузером або іншим агентом користувача. Динамічні сторінки зазвичай обробляють і виводять інформацію з бази даних. Вміст і зовнішній вигляд таких сторінок змінюється залежно від інформації, яка міститься в базі даних.

Для розробки динамічних сайтів використовуються такі технології:

– PHP (докладніше: www.php.net) - ця технологія, що має свою мову програмування, використовується практично на всіх операційних системах [2];

– ASP (Active Server Pages) – технологія від лідера світового програмування компанії Microsoft, використовує мови програмування VBScript, JScript, застосовується тільки на серверах, що працюють під операційною системою ОС Windows;

– бібліотека ARC була написана в 2007 році для використання в мові PHP. Вона була створена для написання динамічних сторінок, а точніше для вставки в них динамічного контенту. Дана бібліотека була створена для обробки онтологій, як у вигляді RDF, так і у форматі OWL. Бібліотека працює із запитамі SPARQL, і виводить отриманий результат в динамічну частину сайту.

У бібліотеці ARC розробники намагаються тримати речі простими і гнучкими. Всі функції з підтримки практичних випадків використання, не реалізують специфікації заради повного дотримання. Насправді, один з мотивів ARC навіть не створення більш семантичних додатків, а перенесення RDF в повсякденні web-розробки [3].

Термін "онтологія" вперше з'явився в роботі Томаса Грубера в 1993 році, в якій розглядалися різні аспекти взаємодії інтелектуальних систем між собою і с людиною. Інтелектуальними системами називаються програми, які моделю-



ють деякі аспекти інтелектуальної діяльності людини. Онтологія являє собою опис декларативних знань, призначене для читання людиною і виконане у вигляді класів з відношенням ієрархії між ними. Складання опису декларативних знань вимагає великої роботи і певних навичок. Для позначення цієї роботи, а також її результату, Грубер ввів спеціальний термін – «концептуалізація». Опис він назвав «специфікацією». Таким чином, онтологія по Грубер визначається як «специфікація концептуалізації».

Для машинного представлення різних предметних областей в Internet, використовуються онтології. Онтологія – специфікація концептуалізації, або явне, формальний опис предметної області. Як і в об'єктно-орієнтованому описі галузі знань, онтологія складається з класів та їх примірників, у класів і екземплярів виділяються властивості, на властивості можуть накладатися логічні обмеження.

Онтології вмісту Web-сторінок необхідні пошуковим програмам для поліпшення якості пошуку по Web. Ідея побудови специфікацій концептуалізацій вмісту Web-сторінок знаходиться в основі концепції так званого Розумного Web або Semantic Web. Semantic Web являє собою наступне покоління World Wide Web, в якому окрім гіпертекстових документів містяться описи семантики цих документів, а також опису семантики різних сервісів, що надають ці документи кінцевим користувачам [4].

Формальна специфікація вмісту Web-документа дає можливість пошуковій програмі робити висновки про відповідність пошукового запиту даному Web-документу не тільки на основі синтаксичної інформації, одержуваної з тексту цього документа, але і ґрунтуючись на семантиці вмісту даного документа. Це може кардинально поліпшити якість Web-пошуку, так як опис світу Web-сторінки, ясна пошуковій програмі, дає останньої набагато більше інформації, ніж вона може отримати з неструктурованого тексту.

Ідеї розумного Web давно були сприйняті спільнотою W3, в результаті чого вже протягом більше десяти років ведуться роботи по втіленню цих ідей у життя. Першим завданням, яке необхідно вирішити для цього, є розробка стандартного мови, який був би зрозумілий усім пошуковим програмам.

Онтології використовуються науковими спільнотами – для опису термінології, в електронній комерції – для опису товарів і послуг, і в інших додатках Internet. Через своєї популярності онтології стали використовуватися і в якості баз знань локальних інтелектуальних систем.

В Internet використовується безліч мов представлення даних, засновані на XML. В рамках проекту Semantic Web консорціумом W3 розроблено ряд специфікацій для роботи з машинним представленням знань. Зокрема, для представлення даних, що мають графову структуру, на основі XML, розроблена мова RDF (Resource Definition Framework – середа Описи Ресурсу). RDF надає засоби для запису триплетів, трійок даних – суб'єкт – предикат – об'єкт.

Для опису онтологій, доступних через Веб, створені мови RDFS (RDF Schema – RDF Схеми) і OWL (Ontology Web Language – мова Мережевих Онтологій). В якості своїх базових елементів дані мови використовують RDF ресурси. RDFS використовується для запису словників, а OWL – онтологій. Ме-



режеві онтології надають більш виразні можливості в порівнянні з RDF словниками, наприклад логічні операції над класами і логічні обмеження властивостей.

На сьогоднішній день знайти цікаву інформацію, що не переглянувши величезну кількість непотрібних сторінок сайтів, практично неможливо. Для вирішення цієї проблеми і були створені методи динамічного формування сторінок web-сайтів. При верстці будь-якого сайту виникає одна проста проблема, яка, незважаючи на свою простоту, потребує вирішення. Це – проблема статистичного контенту.

Список літератури

1. Стандарти W3C консорциума World Wide Web Consortium - Web Standards [Електронний ресурс] - Режим доступу: [www/ URL http://www.w3.org](http://www.w3.org) - 20.05.2007 - Загл. с екрана.
2. Semantic Web на PHP при помощи библиотеки ARC [Электронный ресурс] - Режим доступа:<http://sheremetov.com/php/semantic-web-php-arc/>- Загл. с екрана.
3. Library ARC [Электронный ресурс] – Режим доступа: <http://www.w3.org/2001/sw/wiki/ARC> - Загл. с екрана.
4. Глибовець М.М., Олецкий О.В. Штучний інтелект. – К. :Видавничий дім “КМ Академія”, 2002. – 366 с.