

УДК 004.056.53

М.М. Герасимчук, Н.Я. Шингера канд. техн. наук

Тернопільський національний технічний університет імені Івана Пулюя, Україна

ПОБУДОВА МЕРЕЖІ ІНТЕРНЕТ В КОНЦЕПЦІЇ SEMANTIC WEB

M.M. Herasymchuk, N.Y. Shynhera, Ph.D.

CONSTRUCTION OF THE INTERNET IN THE CONCEPT OF SEMANTIC WEB

Сьогодні комп'ютери беруть досить обмежену участь у формуванні й обробці інформації в мережі Інтернет. Функції комп'ютерів в основному зводяться до збереження, відображення і пошуку інформації. У той же час створення інформації, її оцінка, класифікація й актуалізація — усе це як і раніше виконує людина. Як включити комп'ютер у ці процеси? Якщо комп'ютер поки не можна навчити розуміти людську мову, то потрібно використовувати мову, що була б зрозумілою комп'ютеру. Тобто, в ідеальному варіанті, вся інформація в Інтернеті повинна розміщуватись двома мовами: людською мовою для людини і комп'ютерною мовою для розуміння комп'ютера. Семантична павутина — це концепція мережі, у якій кожен ресурс людською мовою був би доповнений описом, зрозумілим комп'ютеру. [1]

Семантична павутина — це надбудова над сучасною Всесвітньою павутиною, яка покликана зробити інформацію, що розміщена в мережі, зрозумілішою для комп'ютерів. Відомо, що майже вся інформація в Інтернеті знаходиться в текстовій формі. Не секрет також, що прогрес в галузі обробки людської мови (Natural Language Processing, NLP) йде дуже повільно. Комп'ютери не можуть сприйняти й осмислити словесну інформацію, розміщену в Інтернеті, і в найближчий час, мабуть, не зможуть. Тоді постає питання — як змусити комп'ютери розуміти зміст розміщеної в мережі інформації і навчити їх користуватися нею? На це питання і покликана відповісти концепція семантичної павутини. Слово «семантична» у цьому випадку означає «осмислена», «зрозуміла». [2]

Основна ідея Semantic Web полягає у тому, щоб зробити інформацію в мережі більш формалізованою і зручною для машинного сприйняття, зокрема, для того щоб її можна було ідентифікувати і класифікувати. На думку авторів технології Semantic Web, це можна досягти за допомогою введення метаданих, які повинні супроводжувати будь-яку інформацію і розповідати про її походження, формати та багато іншого, що має полегшити пошук інформації в мережі та її обробку. [3]

Ґрунтуючись на відкритих стандартах, технологія Semantic Web дозволяє описувати і виділяти смислову інформацію (семантику) з довільних даних, зокрема змісту документів або коду додатків. Говорячи, що машина розуміє семантику документа, мається на увазі не тільки інтерпретація набору символів, що містяться в документі, але і те, що машина розуміє сенс (зміст) документу. Наступні елементи є основними в технології Semantic Web: [4]

- глобальна схема імен (URI — Uniform Resource Identifier);
- модель опису даних (RDF — Resource Description Framework);
- мова опису словників (RDFS — RDF Schema);
- засоби опису зв'язків між об'єктами даних (онтології, і мова їх опису OWL — Web Ontology Language).

Ключовим елементом технологій Semantic Web є унікальна система ідентифікації об'єктів. URI — це ідентифікатор об'єкту (ресурсу) в глобальній мережі. Будь-який елемент, схема чи модель даних семантичної мережі повинні мати власний унікальний адресу (URI). Зараз використовуються два типи ідентифікаторів.

1. Універсальний покажчик ресурсів (Uniform Resource Locator, URL) — це URI, який, крім ідентифікації ресурсу, вказує на спосіб поведінки з ресурсом шляхом опису способу доступу до нього або його положення в мережі.

2. Універсальне ім'я ресурсу (Uniform Resource Name, URN) — це URI, який ідентифікує ресурс за допомогою імені в певному просторі імен. Це дозволяє посилатися на ресурси без використання інформації про його розташуванні.

Другий базовий компонент Semantic Web — це модель даних Resource Description Framework (RDF), яка дозволяє об'єднати інформацію з довільних джерел. Формат RDF найбільш корисний у забезпеченні спільного використання інформації, зміст якої може однаково інтерпретуватися різними програмними агентами. RDF заснований на синтаксисі XML і використовує ідентифікатори URI для позначення ресурсів. RDF — це система опису мережевих ресурсів, зрозуміла комп'ютеру. Формат RDF призначений для збереження метаданих. Він описує предметну область в термінах ресурсів, властивостях ресурсів і значеннях властивостей. Відповідно до концепції семантичної павутини, опис у форматі RDF повинен прикріплюватися до кожного мережевого ресурсу. Документи RDF повинні оброблятися комп'ютером автоматично, RDF не призначений для читання і використання людиною. RDF-дані можна розцінювати як сукупність тверджень — суб'єкт, предикат і об'єкт твердження, і представляти у вигляді спрямованого графа, утвореного такими твердженнями. На сьогодні формат RDF вже сформувався й одержав широке поширення, він служить каркасом для створення семантичної павутини.

Наступний рівень у піраміді технології Semantic Web займає RDF Schema — мова опису словників RDF-термінів. RDFS служить фундаментом для багатших мов опису онтологій предметної області, які дозволяють адаптувати до Web системи логіки й забезпечити семантичну обробку даних. Схема RDF являє собою систему типів для Semantic Web і дозволяє визначити класи ресурсів і властивості як елементи словника, зокрема задати, які властивості з якими класами можуть бути використані.

Наступним важливим напрямком концепції семантичної павутини є мова OWL (Web Ontology Language). Ця мова побудована на форматах RDF і RDFS, вона призначена для обробки інформації в мережі. Мова OWL має 3 ступені деталізації, що є новим словом у комп'ютерних технологіях. Вона також легко масштабується й узгоджується з найсучаснішими мережевими стандартами. Теоретичною основою OWL є Описова логіка.

Останнім елементом концепції є SPARQL (Protocol And RDF Query Language) — нова мова запитів для швидкого доступу до даних RDF. Використовуючи звичайний протокол і мову SPARQL, програми можуть аналізувати RDF-описи ресурсів і отримувати з мережі необхідну інформацію.

Література

1. Tim Berners-Lee. The Semantic Web. Tim Berners-Lee, James Hendler and Ora Lassila. Scientific American, May 17, 2001. – 8 с.
2. Глибовець М.М. Застосування Semantic Web для створення колаборативного освітнього простору/М.М. Глибовець//Інформаційні технології в освіті: збірник наукових праць. – 2010. – Вип. 8. – 141-148 с.
3. Ghaleb G. E-learning model based on Semantic Web technology / G. Ghaleb, S. Daoud, A. Hasna and others//International Journal of Computing & Information Sciences. 2006. – Vol. 4. – No. 2. – 63-71 с.
4. Пантелеев М.Г., Пузанков Д.В., Татаринев Ю.С. Перспективы использования технологий Семантического Web [Електронний ресурс] Режим доступу: URL: <http://www.ict.edu.ru/ft/005511//portal3-18.pdf>.