

АРХІТЕКТУРА ПОБУДОВИ СЕМАНТИЧНОГО WEB

Анотація. У статті проаналізовано основні структурні елементи та архітектуру семантичної мережі, як базовий рівень для створення майбутніх інтелектуальних семантичних Web-технологій

Ключові слова. Семантичний WEB, онтологія, семантичні інформаційні технології.

Аннотация. Статья посвящена анализу основных структурных элементов и архитектуры построения семантической сети, как базового уровня для создания будущих интеллектуальных семантических Web-технологий

Ключевые слова. Семантический WEB, онтология, семантические информационные технологии.

Актуальність. Сучасний стан досліджень в області інтеграції інформаційних мережних технологій дозволяє зробити висновок про те, що наступним етапом такої інтеграції стане подальша автоматизація, направлена на інтелектуальний аналіз і програмне вирішення поставлених завдань на якісно новому рівні. Багатокомпонентні розподілені засоби та інформаційні системи повинні без участі людини або з мінімальною її участю аналізувати поставлене завдання і ухвалювати рішення про те, які ресурси і дані слід використовувати для його реалізації. Все це, також, вимагає інтелектуального підходу для вирішення конфліктів, викликаних неповнотою, надмірністю або суперечністю отриманих даних, для розв'язку проблем їх класифікації і встановлення логічних зв'язків між ними, для забезпечення складного багатокритерійного нечіткого пошуку і т.д. Вирішення багатьох подібних завдань неможливе без введення нових інформаційних систем, які ґрунтуються на семантичному аналізі та пошуку інформації, при застосуванні гнучких глобальних семантичних мереж, впровадженні сучасних систем управління знаннями, структуризації баз та архівів знань в мережі Інтернет.

Семантичний Web є розширенням існуючої мережі Інтернет, в якому інформація представляється в чіткому і визначеному смислового значенні, що дає можливість людям і комп'ютерам працювати з вищим ступенем взаєморозуміння та узгодженості.

Побудова семантичного Web-ресурсу опирається на створенні мереж нового типу, які містять інформацію не тільки про самі дані і програми, але й несуть смислову інформацію про їх зміст. Такі мережі стали все більш популярними і відносяться до класу семантичних мереж.

У спрощеному вигляді семантична мережа — це структура представлення знань у вигляді орієнтованого графа, в якому вершини — змістові поняття, а дуги — їх відношення.

Найперші семантичні мережі були розроблені для систем машинного

перекладу у якості електронних перекладачів. Однак останні версії семантичних мереж стали більш потужними і гнучкими і можуть становити конкуренцію для логічного програмування, фреймових систем та ін. інформаційних мережних технологій.

На сьогоднішні існує багато варіантів семантичних мереж. Їх термінологія і структура різняться, але існують подібні ознаки, властиві усім семантичним мережам:

1. Вузли семантичних мереж представляють концепти предметів, подій, станів.

2. Дуги семантичних мереж створюють відносини між вузлами-концептами (позначки над дугами вказують на тип їх відносин).

3. Деякі відносини між концептами представляють лінгвістичні відмінки, такі як агент, об'єкт, реципієнт та інструмент (інші означають тимчасові, просторові, логічні відносини і відносини між окремими пропозиціями).

4. Концепти організовані по рівнях відповідно за ступенем узагальненості їх семантичних значень.

Незважаючи на деякі відмінності, сучасні мережі зручні для читання та обробки на персональних комп'ютерах, а також вони досить потужні, щоб представити семантику природної мови. Найбільш часто в семантичних мережах використовуються такі відносини:

- зв'язок типу «частина-ціле» («клас-підклас», «множина-підмножина» тощо);
- функціональні зв'язки, які, зазвичай, визначаються дієсловами (наприклад, «робить», «належить» і т.д.);
- кількісні («більше», «менше», «одне» і т.д.);
- просторові («близько від», «далеко від» і т.д.);
- тимчасові («раніше», «пізніше» і т.д.);
- логічні зв'язки («і», «або» і т.д.);
- лінгвістичні зв'язки.

Головна відмінність семантичної мережі від звичайної полягає в тому, що кожна веб-сторінка семантичної мережі містить інформацію на двох мовах: на звичайній, зрозумілій людині, що відображається браузером, і на спеціальній, інформація на якій прихована від користувача, але зрозуміла інтелектуальним програмам-агентам. Цю спеціальну мову описує представлений на сторінці матеріал за допомогою тегів і атрибутів, зрозумілих комп'ютеру.

Концепцію семантичної мережі як нової форми представлення веб-сервер контенту голова W3C Тім Бернерс-Лі представив ще у 2001 р., виступивши на восьмій конференції консорціуму W3C і опублікувавши відповідний трактат в журналі «Scientific American». Він детально описав переваги нової форми представлення інформації в WWW, яка передбачає максимальний ступінь класифікації будь-якої інформації, роблячи спільну роботу людей і машин ефективнішою на порядок [2].

Семантичні веб-сервер технології допомагають виділяти корисну інформацію з даних, змісту документів або коди програм, спираючись на відкриті стандарти. Якщо комп'ютер розуміє семантику документа, то це не означає, що він просто інтерпретує набір символів, що містяться в документі, а розкриває зміст, суть документа.

Семантичні технології окреслюють загальні рамки, що дозволяють здійснювати обмін даними і їх багатократне використання в різних галузях, організаціях і навіть співтовариствах. Семантичні веб-сервер технології — це ефективний спосіб представлення даних і програм в мережі Інтернеті. Семантичні технології представляють дані за допомогою онтологій і забезпечують аргументацію, використовуючи зв'язки, правила, логіку і умови, обумовлені в онтології. [3]

Онтологія (у інформації) — це повна і детальна формалізація деякої області знань за допомогою концептуальної схеми. Зазвичай, така схема складається з ієрархічної структури даних, що містить всі релевантні класи об'єктів, їх зв'язки і правила (теореми, обмеження), прийняті в цій області. Онтології призначені для опису наочної області в термінах відносин між суттю і їх обмеженнями (constraints), які необхідні комп'ютерам для сприйняття семантики з метою аналізу, порівняння, зіставлення даних, а також виведення нових знань.

Можна визначити такі найпоширеніші семантичні веб-технології:

Глобальна схема імен (URI, Universal Resource Identifier). URI — це ідентифікатор адреси веб-сервера-ресурсу. Будь-який користувач може створити URI, але права власності на них чітко організовані, тому вони представляють ідеальну базову технологію для побудови глобальної мережі. Фактично, Інтернет є саме такою структурою: все, що має URI, вважається таким, що знаходиться в глобальній мережі.

Стандартний синтаксис опису даних (RDF, Resource Description Framework). RDF — це специфікація, яка визначає модель представлення даних і синтаксис для обміну даними. RDF забезпечує послідовний стандартний спосіб опису і роботи практично з будь-якими інтернет-ресурсами: від текстових сторінок і графіків до аудіо-файлів і відео-кліпів. Він пропонує синтаксичні можливості для взаємодії мереж і формує базовий шар для створення семантичної мережі. RDF визначає керовані графи зв'язків, представлені трійками «об'єкт-атрибут-значення». Наприклад, об'єкт X має атрибут Y із значенням Z.

Стандартні способи опису властивостей даних (схема RDF). Схема RDF — це семантичне розширення RDF. Вона забезпечує механізми опису пов'язаних ресурсів, а також самих зв'язків. Описова мова словника RDF визначає властивості ресурсу в термінах того класу ресурсів, до якого цей ресурс відноситься.

Стандартні способи опису зв'язків між об'єктами даних (онтологія, визначається за допомогою онтологічної мови веб-сервера (Ontology Web Language, OWL)). Синтаксична взаємодія — необхідна умова того, що

програми можуть працювати з даними як з інформацією. Це також необхідна умова для коректної перевірки даних. Синтаксична взаємодія часто вимагає перетворення між термінами, для чого, у свою чергу, необхідний контент-аналіз. Такий контент-аналіз вимагає формальних і докладних специфікацій моделей доменів, які визначають використовувані терміни і їх зв'язки. OWL покликана полегшити вираз онтологій і додає більше словникових можливостей для опису властивостей і класів, ніж RDF або схема RDF. Зокрема, вона дозволяє описувати зв'язки між класами (наприклад, не перетин), повнота множини (наприклад, «рівно одна»), рівність, багатшу типологію властивостей і їх характеристики (наприклад, симетрію). Загальна схема побудови семантичної мережі представлена на рис. 1. [4]

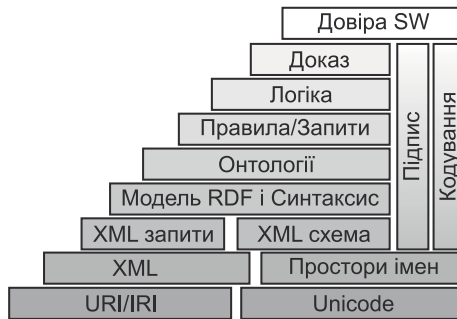


Рис. 1. Архітектура семантичної мережі

Розвиток семантичних мереж сприяє розробкам програмних агентів, що виконують завдання збору і аналізу інформації, багатокритерійного пошуку, зчитування релевантної інформації, накопичення знань і т.п., вирішення яких утруднене в рамках сучасного слабоструктурованого Інтернету.

Такі агенти відіграють найважливішу роль в семантичній мережі. Агенти по завданнях користувачів (або інших агентів і сервісів) знаходять джерела інформації, запрошують дані, зіставляють і перевіряють їх на відповідність критеріям пошуку, а потім видають відповідь у формі, зручній для відправника запити. [5]

Висновки. Створення семантичної мережі дозволяє значно підвищити ступінь автоматизації обробки даних і сприяє розвитку динамічних інформаційних систем, що характеризуються гнучкістю, масштабованістю, незалежністю від операційних платформ ПК.

Таким чином, семантична мережа успішно вирішує наступні важливі для інтеграції даних і програм завдання:

- забезпечує незалежність даних від програм;
- реалізує семантичну інтеграцію даних;
- створює основи для використання інтелектуальних сервісів (веб-сервісів), без яких не може бути реалізований семантичний Web.

1. *Balani N.* The future of the Web is Semantic, 18 Oct. 2005, (<http://www-128.ibm.com/developerworks/library/wa-semweb/index.html>).
2. *Tim Berners-Lee.* Semantic Web Overview, Hong-Kong, May 1-5 2001, (<http://www10.org/program/w10-prog-w3c.html>)
3. *Gomez-Perez A. Corcho O.* Ontology Languages for the Semantic Web. IEEE Intelligent Systems, vol. 17, no. 1, January/February 2002. — С. 54–60.
4. *Ланде Д. В.* Семантический веб: воплощение идей / Д. В. Ланде. — М. : Телеком, 2005. — № 6. — С. 60–65.
5. *Кальченко Д.* Интеллектуальные агенты семантического Web'a / Д. Кальченко. — М. «Компьютер-Пресс», 2004. — № 10. — С. 45-48.

Поступила 25.9.2013г.