

## Вступ

З кожним роком кількість ресурсів глобального інформаційного простору невідмінно зростає. Якщо на початок 2010 року чисельність веб-сайтів становила близько 200 млн, то на початок 2019 року цей показник зріс до 1,8 млрд. Можна виділити дві основні проблеми мережі Інтернет: зростання обсягів інформаційного наповнення, породжені популярністю і дешевизною Web-технологій, а друга – формат подання інформації в мережі, який орієнтований переважно на людей і лише в окремих випадках допускає автоматичну обробку програмними агентами.

У результаті, по-перше, виникає проблема знаходження потрібної користувачу інформації в будь-якому вигляді – обсяги Web-простору не дають змоги оперативного оновлювати бази даних інформаційно-пошукових систем (не кажучи вже про неможливість досягнення 100-відсоткового охоплення), а, по-друге, в автоматизованому режимі практично неможливо виділити зміст інформаційних документів.

Тому подальший розвиток Internet багато дослідників пов'язують із концепцією семантичної мережі (Semantic Web), яка завдяки уніфікації обміну даними надає можливості із інтеграції в Internet об'єктів реального світу. Концепцію семантичної павутини висунув Тім Бернерс-Лі, один із основоположників World-Wide Web і голова консорціуму W3C на міжнародній конференції XML-2000, що проходила у Вашингтоні. Основна ідея цього проекту полягає в організації такого представлення даних у мережі, щоб допускалася не тільки їх візуалізація, але й ефективна автоматична обробка програмними агентами. Інакше кажучи, передбачається перетворення глобальної мережі в систему семантичного рівня. За задумом творців семантична мережа повинна забезпечити “розуміння” інформації комп'ютерами, виділення ними найбільш придатних даних відповідно до введених критеріїв, і вже після цього – надання інформації користувачам.

У разі автоматичної обробки інформації в межах Semantic Web передбачається взаємодія між сервісами на основі аналізу смислових зв'язків між об'єктами і поняттями, що зберігаються в мережі для вибору лише тієї інформації, яка буде реально корисна користувачам.

За визначенням консорціуму W3C, семантичний Web є розширенням мережі Internet, в якому інформація подана в чіткій та визначеній змістовій формі, що дає можливість людям і комп'ютерам працювати з вищим ступенем взаєморозуміння й узгодженості. Організація Semantic Web передбачає об'єднання різноманітних видів інформації в єдину структуру, де кожному смисловою елементу даних відповідатиме спеціальний синтаксичний блок

(тег). Теги становитимуть єдину ієрархічну структуру, на основі якої і повинна функціонувати семантична мережа. Згідно із рішенням консорціуму W3C, у межах проекту “Semantic Web” розробляють мови для подання інформації у формі, доступній для машинної обробки, якими можна буде описувати і дані, і принципи їх трактування. Це повинно зумовити те, що правила виведення, які існують у будь-якій окремій системі подання знань, будуть передаватися по мережі іншим подібним системам. У процесі реалізації концепції семантичної мережі широко розвинуті синтаксичні методи подання інформації мовними засобами XML і його доповненнями, призначених для опису типових властивостей елементів XML-документів, їх структури і семантики: рекомендації W3C, що регламентують DTD (Document Type Definition), XML Schema, XQuery (мова запитів до баз даних XML) тощо.

Інший напрям семантичної Мережі пов’язаний із напрямками, близькими до галузі штучного інтелекту, та отримав назву *онтологічного підходу*. Такий підхід охоплює засоби анотування документів, якими могли б скористатися комп’ютерні програми – Web-сервіси та агенти під час опрацювання складних запитів користувачів. З огляду на те, Semantic Web можна представити як симбіоз двох напрямів, перший з яких охоплює мови подання даних (розширювана мова розмітки XML (eXtensible Markup Language) та засоби опису ресурсів RDF (Resource Description Framework)); другий – концептуальний напрям, який відображає теоретичне подання моделей предметних галузей – онтологій та мову її специфікації OWL (Web Ontology Language). Отже, два напрями Semantic Web використовують три ключові мови (технології):

- специфікація XML, що дає змогу визначити синтаксис і структуру документів;
- механізм опису ресурсів RDF, що забезпечує модель кодування для значень, визначених в онтології;
- мова онтологій OWL, що дає змогу визначати поняття і відношення між ними.

З огляду на це, в цьому курсі вивчають особливості синтаксису та застосування мов подання онтологій та способи їх використання з метою застосування під час побудови семантичної мережі.