

УДК 004.043

В. Крамар, С. Лупенко

(Тернопільський національний технічний університет імені Івана Пулюя)

ПРОБЛЕМА СТРУКТУРИЗАЦІЇ ІНФОРМАЦІЇ В МЕРЕЖІ ІНТЕРНЕТ

У публікації розглянуто проблему структуризації інформації в мережі Інтернет для її інтелектуального аналізу та пошуку. Проаналізовано сучасний стан проблеми та підходи до її вирішення. Запропоновано використання моделей та методів теорії самоорганізації для розробки систем структуризації інформації [1].

Як відомо, сьогодні обсяг інформації в мережі Інтернет збільшується стрибкоподібним чином. Тому постає проблема структуризації цієї інформації для її інтелектуального аналізу та пошуку. Структуризація – це створення такої множини елементів, які знаходяться у взаємодії та специфічному порядку, необхідному для виконання певних функцій. Зокрема, структуризація інформації здійснюється насамперед для інформаційного пошуку.

Можна виділити два підходи до структуризації інформації в мережі Інтернет: шляхом розробки систем структуризації інформації та структуризації інформації в самій мережі, шляхом створення метаданих опису інформаційних ресурсів. Останній підхід є нічим іншим як створенням семантичного web. Семантичний web – це концепція розвитку мережі Інтернет, в якій основну роль відіграють метадані та їх машинна обробка. Ця концепція прийнята консорціумом W3. Однак застосуванню семантичного web перешкоджають декілька складних проблем. Це, насамперед, складність впровадження цієї концепції, незважаючи на існування цілої низки стандартів, таких як XML, RDF, OWL. На сьогодні величезний обсяг інформації в Інтернеті є беззмстовним з точки зору програм і орієнтований, в основному, на використання цих ресурсів людиною, що помітно обмежує розвиток інформаційних систем та інформаційних технологій [2]. Звичайно, практично вирішуватись більш ефективно ця проблема повинна з застосуванням цих двох підходів одночасно.

Щодо першого підходу, то такі системи є досить складними. Для їх розробки необхідно використовувати якісно нові принципи розробки таких систем. Ними можуть виступити принципи самоорганізації. Зокрема, пропонується використовувати підхід з використанням теорії Рамсея із застосуванням правил самоорганізації до неструктурованої інформації для її організації [3].

Перелік посилань

1. Крамар В.І. Самоорганізація в задачах проектування інформаційно-пошукових систем / В.І Крамар, С.А. Лупенко // Системний аналіз та інформаційні технології: матеріали 12-ї Міжнародної науково-технічної конференції SAIT 2010. – К.: ННК “ІПСА” НТУУ “КПІ”, 2010. – С. 276.

2. Марчук А.Г. Semantic Web - на пути к новому поколению информационных систем / А.Г. Марчук // Труды 7-ой Всероссийской научной конференции "Электронные библиотеки: перспективные методы и технологии, электронные коллекции", Ярославль, Россия, 4-6 октября 2005 г. – С. 122-125.

3. Крамар В.І. Новий підхід до формалізації процесів самоорганізації на основі теорії Рамсея / В.І. Крамар, С.А. Лупенко // Вісник Тернопільського національного технічного університету. – 2012. – №2. – С. 143-151.