

К. Слобода

Національний університет “Львівська політехніка”,  
кафедра соціальних комунікацій та інформаційної діяльності

## СТРУКТУРУВАННЯ СЕМАНТИЧНОГО ЯДРА ВЕБ-СПІЛЬНОТИ

© Слобода К., 2011

**Запропоновано новий підхід до структурування семантичного ядра веб-спільноти, а також до його формування. Запропоновано нову класифікацію типів структури семантичного ядра веб-спільноти.**

**Ключові слова:** semantic core, web-community, positionig, structuring.

**This article proposes a new approach to structuring the semantic core of web community, and to its formation. A new classification of types of web community semantic core structures is given.**

**Keywords:** семантичне ядро, веб-спільнота, позиціонування, структурування.

### Вступ

Формування семантичного ядра веб-спільноти є одним з найбільш трудомістких та складних завдань в процесі формування її позиції в глобальному інформаційному середовищі. Оскільки веб-спільноти відрізняються як за структурою, так і за аудиторією та інформаційним наповненням від звичайних сайтів, процес формування семантичного ядра істотно відрізняється від запропонованих для сайтів. Однією з основних проблем, що постають перед власниками веб-спільноти в процесі формування семантичного ядра, є великі обсяги інформаційного наповнення веб-спільноти. Що більший об'єм унікального інформаційного наповнення, то більше семантичне ядро. Велика кількість ключових слів у семантичному ядрі веб-спільноти створює проблеми з управління цим ядром, та, відповідно, з управління процесом її позиціонування. Тому для веб-спільноти доцільно розробити новий підхід до структурування семантичного ядра, який би забезпечив спрощення процесу позиціонування та зменшення витрат часу і ресурсів.

### Аналіз останніх досліджень

Завдання ефективного позиціонування веб-сайтів досліджено в працях І.С. Ашманова та А.А. Іванова, професора А. М. Пелещишина, проте підходи до ефективного формування та структурування семантичного ядра веб-спільнот не розглядаються в цих працях.

### Цілі статті

1. Визначити основні етапи процесу формування семантичного ядра веб-спільноти у таких випадках: якщо інформаційне наповнення вже створене і потрібно здійснювати процес оптимізації, якщо інформаційне наповнення не створене і якщо частина інформаційного наповнення створена, а частина не створена.
2. Визначити основні типи структури семантичного ядра.

### Структурування семантичного ядра веб-спільноти

Семантичне ядро веб-спільноти – це список слів, словосполучень та фраз, які відображають її тематику і використовуються в її інформаційному наповненні. Етап створення семантичного ядра веб-спільноти – один з основних етапів в процесі її оптимізації для пошукових систем.

Процес оптимізації може розпочинатися в одному випадку на етапі, коли інформаційне наповнення веб-спільноти ще не створене, а в іншому випадку, коли інформаційне наповнення вже створене. Процес формування семантичного ядра матиме відмінності у цих двох випадках.

Формування семантичного ядра:

1. Тексти (інформаційне наповнення) створені:

- Визначення тематики веб-спільноти.
- Визначення структури семантичного ядра.
- Формування базового семантичного ядра на основі слів з текстів.
- Розширення базового семантичного ядра іншими ключовими словами і фразами.
- Аналіз ключових слів з семантичного ядра, частот їх появи в запитах користувачів

пошукових систем.

- Розширення базового семантичного ядра словами і словосполученнями, підібраними автоматично з використанням бази даних запитів пошукових систем.

- Оптимізація інформаційного наповнення підібраними ключовими словами.

2. Тексти (інформаційне наповнення) ще не створені:

- Визначення потенційної тематики веб-спільноти.
- Поділ тематики на підкатегорії.
- Формування базового семантичного ядра на основі визначених тематичних

підкатегорій.

- Аналіз ключових слів з семантичного ядра частот їх появи в запитах користувачів пошукових систем.

- Розширення базового семантичного ядра словами і словосполученнями, підібраними автоматично з використанням бази даних запитів пошукових систем.

- Написання оригінальних текстів з паралельним наповненням ключовими словами.

- Подальше розширення семантичного ядра в процесі функціонування веб-спільноти, доповнення підкатегорій тематики тощо.

3. Частина текстів (інформаційного наповнення) створена, а частина створюється в процесі функціонування веб-спільноти:

- Аналіз вже створених текстів.
- Визначення тематики та тематичних підкатегорій.
- Формування структури семантичного ядра на основі структури тематики.
- Формування базового семантичного ядра на основі цих текстів.
- Розширення і доповнення семантичного ядра.
- Наповнення веб-спільноти новими текстами.
- Доповнення структури тематики та семантичного ядра.
- Розширення семантичного ядра ключовими словами, що відповідають доповненій

тематичі.

- Оптимізація текстів.

Типи семантичного ядра за інформаційним наповненням:

1. За кількістю охоплених мов:

- Одномовне (всі ключові слова складаються зі слів однієї мови).
- Двомовне (всі ключові слова складаються зі слів двох мов).
- Багатомовне (ключові слова складаються зі слів більш ніж двох мов).

2. За тематикою:

- Монотематичне.
- Політематичне.

Типи запитів, які користувачі задають пошуковим системам в процесі пошуку інформації:

1. Інформаційні (тип запитів, які користувач задає, щоби отримати інформацію, яка б відповідала на його запитання).

2. Навігаційні (тип запитів, які користувач задає, щоби знайти певний ресурс у WWW).
3. Транзакційні (тип запитів, які користувач задає з метою отримати посилання на сайт, де можна знайти товар, який його цікавить. Як правило, запити такого типу пов'язані з бажанням користувача купити товар).

Типи структури семантичного ядра:

- Проста структура (характерна для монотематичного семантичного ядра. Ключові слова розподілені за тематикою в межах одного семантичного ядра)

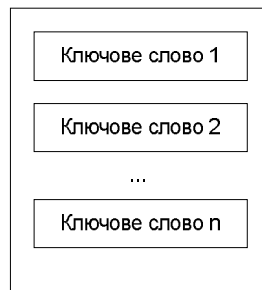


Рис. 1. Проста структура семантичного ядра

- Складна структура (характерна для політематичного семантичного ядра. Семантичне ядро структуроване відповідно до структури підтематик веб-спільноти і поділене на кілька окремих семантичних підядер)

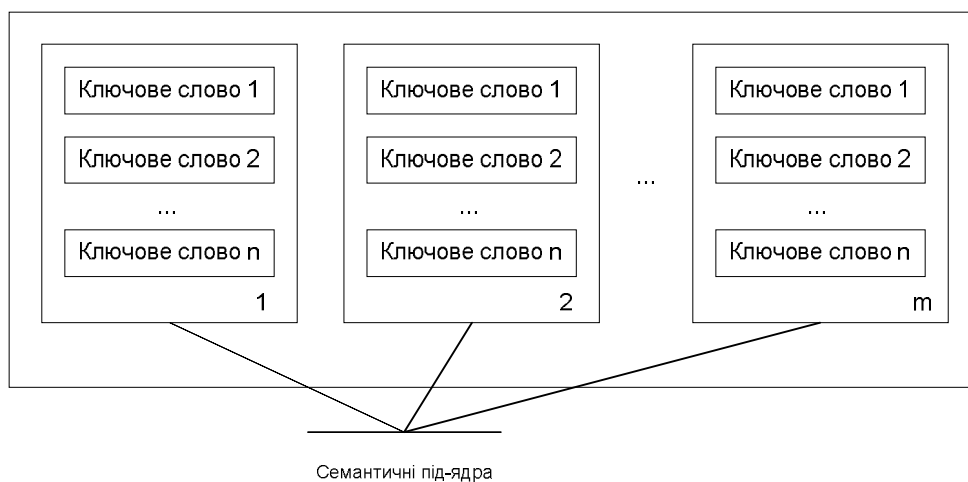


Рис. 2. Складна структура семантичного ядра

- Складена структура (характерна для політематичних семантичних ядер, в яких необхідний високий рівень деталізованості тематики)

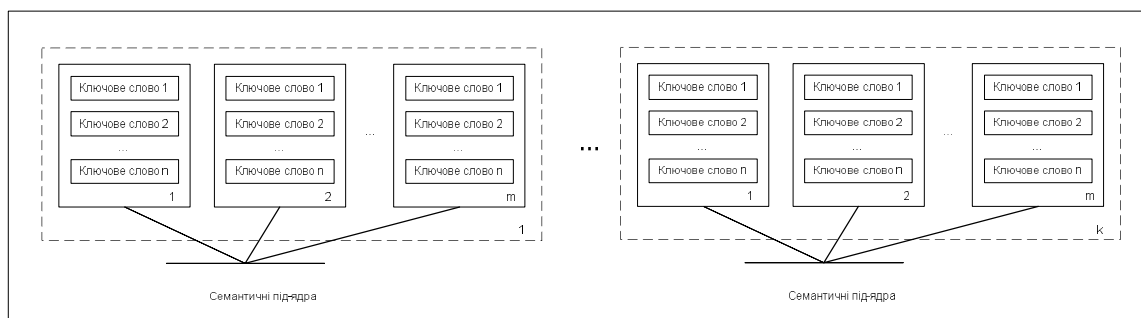


Рис. 3. Складена структура семантичного ядра

Сайти веб-спільнот, як правило, мають політематичні семантичні ядра зі складною або складеною структурою.

Семантичне ядро звичайного сайту vs. Семантичне ядро сайту веб-спільноти:

Семантичне ядро звичайного сайту формується відповідно до тематики, цілей його власників та завдань сайту. Базове ядро сайту є вже основним ядром сайту. Таке ядро є статичним, тобто воно не зазнає суттєвих змін в процесі функціонування та оптимізації сайту.

Семантичне ядро сайту веб-спільноти є динамічним, тобто базове семантичне ядро, розроблене на етапі створення спільноти, істотно доповнюється і змінюється в процесі її функціонування (створення нових обговорень, дискусій та дописів).

Основними показниками семантичного ядра форуму є:

1. Частотність запитів.
2. Обсяг семантичного ядра.

Пошукові запити користувачів до пошукових машин є списками ключових слів з можливими додатковими обмеженнями. Використання ключових слів є традиційним методом визначення тематики текстової інформації. Проте ключовим словом у глобальному інформаційному середовищі можна вважати як одне слово природної мови, так і словосполучення з кількох слів [5]. Тому доцільно розділити два поняття: ключове слово (одне слово або логічно нерозривне коротке словосполучення) та ключова фраза (сукупність ключових слів та словосполучень).

Залежно від частоти вживання у пошукових системах запити поділяються на:

1. Високочастотні запити – в переважній більшості випадків складаються з одного слова загального іменника (найчастіше це збірні, речовинні та загальні назви предметів) або словосполучення – власного іменника (антропоніми, топоніми та культуроніми). Використовуються в процесі пошуку інформації великою кількістю користувачів, що створює конкуренцію між веб-спільнотами з подібною тематикою.

2. Середньочастотні запити – зазвичай це словосполучення з двох-трьох слів, які можуть містити високочастотні запити з додатковими словами для уточнення.

3. Низькочастотні запити – використовуються зрідка, але забезпечують веб-спільноті аудиторію користувачів, які точно знають, яка інформація їм потрібна і задають відповідні запити.

Обсяг семантичного ядра залежить від кількості інформаційного наповнення веб-спільноти.

На етапі створення семантичного ядра аналізують ключові слова і фрази, які використовують або ймовірно можуть використовувати користувачі для знаходження інформації, яка відповідає тематиці веб-спільноти. Це є один з найбільш трудомістких етапів роботи з оптимізації сайту. Сторінки веб-спільноти, оптимізовані під певні ключові слова, будуть стабільно отримувати постійних відвідувачів. Також створення семантичного ядра дасть змогу точно визначити найпопулярніші запити відвідувачів, що використовуються в процесі пошуку інформації, а також оцінити сукупні пошукові запити за тематикою веб-спільноти. До семантичного ядра необхідно внести всі ключові слова та вирази, які користувачі можуть задавати як запити пошуковим машинам та які точно описують тематику ресурсу.

Створюючи семантичне ядро, слід враховувати не тільки правильно написані ключові слова, а й можливі запити з найчастішими помилками, на сленгу та слова, написані за допомогою транслітерації. На цьому етапі оцінюється популярність кожної ключової фрази, ступінь її відповідності інтересам цільової аудиторії і на основі цих даних робиться висновок про введення фрази в семантичне ядро веб-спільноти.

Сформувавши семантичне ядро, можна оцінити приблизний розмір цільової аудиторії веб-спільноти, що користується пошуковими машинами для пошуку інформації, а, отже, і потенційну відвідуваність, яка може забезпечити успішне просування в пошукових системах. Для створення семантичного ядра необхідно провести:

1. Аналіз тематики та створення початкового списку ключових слів.
2. Пошук синонімів, спільнокоренових слів, слів з подібним змістом, популярних запитів для всіх ключових слів. На цьому етапі слід використовувати словники синонімів та бази запитів пошукових систем.
3. Виявлення варіантів написання термінів, які використовуються як в українській, так і в англійській мовах, жаргонізмів та типових орфографічних помилок. Також доцільно ввести в

семантичне ядро найпопулярніші ключові слова, надруковані на англійській розкладці клавіатури, оскільки дуже часто користувачі вводять пошукові запити, не переключивши розкладку клавіатури.

4. Аналіз частот запитів, виявлених за списком зовнішніх посилань.

5. Видалення багатозначних ключових слів, не вживаних у певній мові та таких, що неточно відповідають тематиці веб-спільноти.

### Висновки

Запропоновано підхід до структурування семантичного ядра веб-спільноти, запропоновано три підходи до формування структури семантичного ядра веб-спільноти, а також визначено типи структури семантичного ядра. Чітким структуруванням семантичного ядра можна спростити процес позиціонування та зменшити витрати часу і ресурсів. Запропоновані в статті підходи успішно застосовані в процесі створення та розвитку веб-спільноти “Клуб любителів електронних книг”

1. Ашманов И. С., Иванов А. А. Продвижение сайта в поисковых системах / И. С. Ашманов, А. А. Иванов – М.: “Вильямс”, 2007. – 304 с. 2. Евдокимов Н. В. Раскрутка Web-сайтов. Эффективная Интернет-коммерция / Н. В. Евдокимов – М.: “Вильямс”, 2007. – 160 с. 3. Пелецишин А. М. Моделювання аудиторії Веб-сайта // Вісник Нац. ун-ту “Львівська політехніка”: Комп’ютерні системи проектування: теорія та практика. – 2004. – №522. – С.136 – 141. 4. Пелецишин А. М. Позиціонування сайтів у глобальному інформаційному середовищі: Монографія. – Львів: Вид-во Нац. ун-ту “Львівська політехніка”, 2007. – 260 с. 5. Пелецишин А. М. Оптимізація форумів та інших форм спільнот користувачів WWW // Вісник Національного університету “Львівська політехніка”: Інформаційні системи та мережі. – 2005. – №549. – С.138–146. 6. Серов Ю. О. Методи аналізу ефективності Веб-форумів / Ю. О. Серов, Р. Б. Кравець, А. М. Пелецишин // Інформаційні системи та мережі: Вісник Нац. ун-ту “Львівська політехніка”. – 2009. – № 653. – С.197–206.