

МОДИФІКАЦІЯ КОНТЕНТУ ДЛЯ ЕФЕКТИВНОГО ПОЗИЦІОНУВАННЯ ФОРУМУ В СЕРЕДОВИЩІ WORLD WIDE WEB

© Пелешишин А.М., Слобода К.О., 2010

Класифіковано основні елементи контенту форуму, здійснено їх аналіз та запропоновано способи їх модифікації, а також визначено поняття семантичного ядра форуму та описано етапи його створення. Здійснено аналіз позицій форуму в середовищі WWW після практичного застосування запропонованих методів: наведено динаміку зміни кількості відвідувачів та кількості переходів з пошукових систем.

Ключові слова: інтернет-спільнота, форум, контент, модифікація контенту, семантичне ядро, частотність запитів, позиціонування форуму.

In this article the basic elements of forum content are classified. Their analysis is conducted; the ways of their modification are suggested. The stages of forum semantic field creation are described. The analysis of the forum position in the WWW after the practical application of proposed methods is conducted. The dynamics of the visitors' number change and conversions from the search engines is shown.

Keywords: online-community, forum, content, modification of content, semantic field, forum positioning, requests frequency.

Постановка проблеми

Форум – це сайт з особливою структурою, який складається з множини тем та повідомлень, створених його учасниками. Дуже часто структура форуму побудована так, що вся множина тем форуму, для зручності учасників та підвищення ефективності модерування, розбита за семантичною ознакою на підмножини. Особливостями форумів є [6]:

- 1) тематично структуроване подання інформації;
- 2) тривалий період існування інформації;
- 3) зручне подання інформації;
- 4) усебічна участь членів спільноти в її житті.

Основною відмінністю між традиційними сайтами та сайтами інтернет-спільнот є те, що інформаційне наповнення сайту безпосередньо залежить від його відвідувачів. Тобто тематика форуму в такому разі лише частково залежить від його власників, які на початкових етапах створення спільноти визначають основну тематичну спрямованість і надалі лише модерують процес наповнення форуму контентом. Зазвичай інформаційне наповнення сайту створюють безпосередньо його користувачі, які самостійно формують, доповнюють та змінюють його тематику.

Форум надає засоби зручного представлення інформації та механізми взаємодії між учасниками спільноти. Можливість подавати різноманітну інформацію структурно дає змогу створювати на одному форумі різні за тематичною спрямованістю підфоруми, які містять теми з дописами користувачів.

Важливою перевагою форуму над соціальними мережами, як ще одного популярного типу інтернет-спільнот, є збережність інформації та її індексація пошуковими системами, що допомагає користувачам легко потрапляти на сторінки обговорень.

Використання пошукових машин є одним з найпопулярніших методів навігації користувачів у глобальному середовищі World Wide Web, яке складається з величезної кількості сайтів. Біль-

шість з цих сайтів взагалі не мають відвідувачів або мають 1–2 відвідувачів на день. Це означає, що власники сайта даремно витратили час і кошти на його створення та підтримку. Однією з основних та найважчих для вирішення проблем є ефективне позиціонування сайта в глобальному інформаційному середовищі та збільшення трафіку користувачів [5].

Для ефективного позиціонування форуму насамперед необхідно оптимізувати та модифікувати контент, який створюють його користувачі. Це забезпечить форуму високий ранг в пошукових системах та постійний пошуковий трафік.

Аналіз останніх досліджень

Питання ефективного позиціонування сайтів у глобальній мережі WWW є надзвичайно актуальним. Останніми дослідженнями цього питання є праці професора А.М. Пелешишина про позиціонування сайтів у глобальному інформаційному середовищі (2007), І.С. Ашманова та А.А. Іванова про просування сайтів у пошукових системах (2007), Н.В. Євдокімова про основи оптимізації контенту (2007). В цих працях розглянуто загальні методи покращання позицій сайта, але не розглянуто методів поліпшення позицій форумів як окремого підвиду сайтів за допомогою модифікації контенту, не класифіковано елементи контенту форуму та не запропоновано методів його оптимізації. Також недостатня увага приділяється процесу створення семантичного ядра форуму.

Цілі статті

1. Визначення поняття модифікації контенту форуму та основних її етапів.
2. Класифікація елементів контенту форуму та визначення способів його модифікації.
3. Визначення поняття семантичного ядра форуму та основних етапів його створення.
4. Аналіз позиції у глобальному інформаційному середовищі сайта-форуму «Український клуб електронної книги» ebook – ua.org після застосування запропонованих методів модифікації контенту.

Основний матеріал

Завдання пошуку необхідної користувачеві інформації у гіпертекстових масивах даних є надзвичайно складним, оскільки пошуковим системам необхідно опрацьовувати величезні обсяги інформації. Результатами пошукової видачі на основі запитів, що складаються з найчастіше вживаних у певній мові слів та словосполучень, є величезна кількість сторінок. Проте звичайним користувачам такої кількості інформації просто не потрібно.

Хоча всі пошукові системи мають різні алгоритми роботи та пошуку за запитом, спільним є те, що більшість аналізує саме лінгвістичну інформацію, що дає змогу робити припущення щодо відповідності певної сторінки сайта пошуковому запиту. Зокрема, аналізується частота появи ключових слів запиту на сторінці, частота появи слів-форм ключових слів на сторінці, частота появи синонімів до ключових слів на сторінці, а також аналізується семантика запиту із семантикою сторінки [5]. Головним фактором ефективності форуму є його позиція в середовищі WWW – сукупність низки важливих системних показників, які об'єктивно відображають тематику, популярність та конкурентоспроможність форуму. Саме позиція визначає реальну значущість форуму, його суспільну вагу. А завдання покращання позиції сайта повинно бути пріоритетним для будь-якого інтернет-проекту. Тому для ефективного позиціонування форуму в інформаційному середовищі WWW надзвичайно важливою є оптимізація його контенту.

За статистикою користувачі схильні вибирати та переглядати інформацію на сайтах, що містяться на початку списку пошукової видачі. Тому, чим вище в списку сайт, тим більша ймовірність того, що користувач його перегляне. І навпаки, чим нижча позиція сайта в списку пошукової видачі за запитом, тим менша ймовірність, що на нього потраплять користувачі. Оскільки контент переважно створюється як учасниками форуму, так і адміністраторами та модераторами, важливим завданням є така модифікація створеного контенту, щоб пошукові системи могли за запитом користувачів видавати посилання на форум на перших сторінках результатів пошуку. На сучасному етапі розвитку WWW найоптимальнішим вирішенням проблеми ефективного позиціонування форуму є модифікація його контенту за принципами, які будуть розглянуті у статті.

Модифікація контенту форуму

Модифікація контенту – це цілеспрямовані дії, метою яких є зміна інформаційного вмісту форуму для покращання його індексації пошуковими системами. Вона складається з таких етапів:

1. Внутрішня оптимізація контенту форуму.
2. Створення семантичного ядра форуму.

Розглянемо детальніше ці два етапи.

1. Внутрішня оптимізація контенту форуму

Передусім необхідно визначити, з яких елементів складається контент форуму. В кожному окремому випадку вони можуть бути різними, проте існують спільні для всіх сайтів-форумів елементи.

1.1. Контент форуму

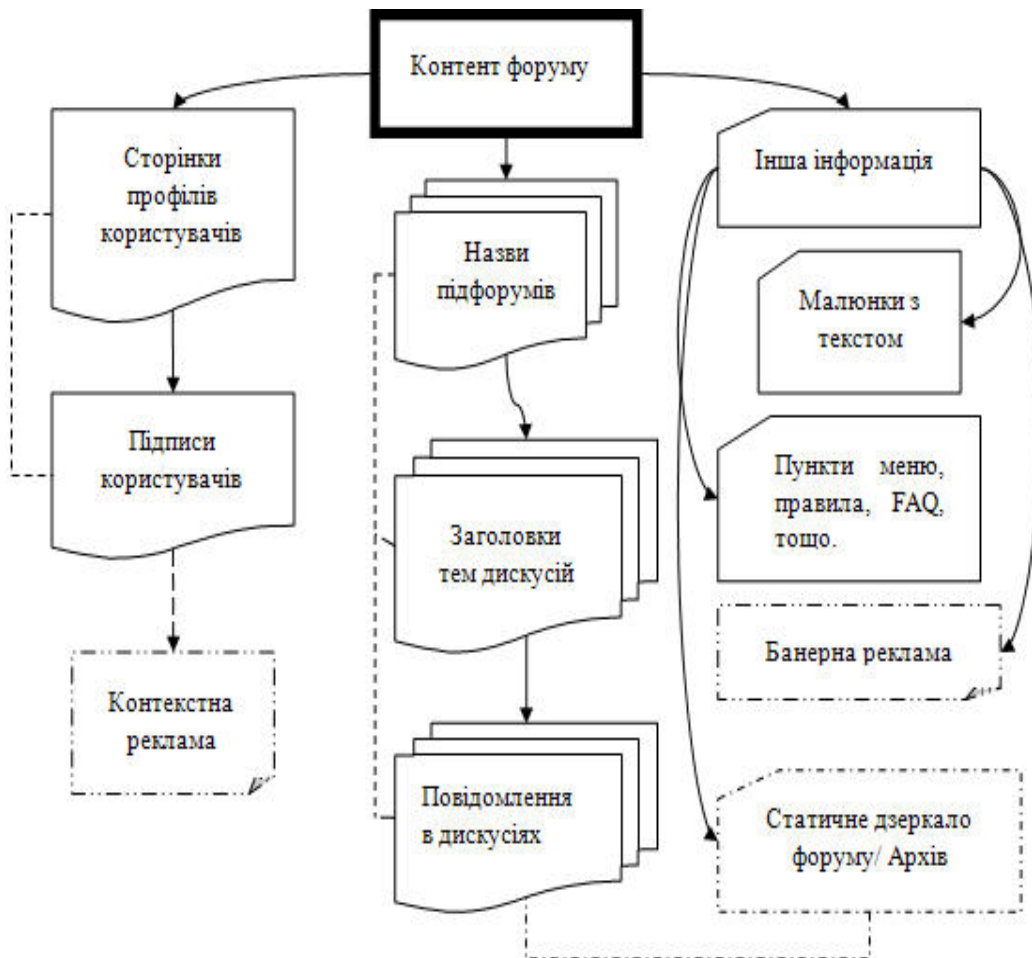


Рис. 1. Основні елементи контенту форуму

Оскільки пошукові системи аналізують здебільшого саме текст (до них належать пошукові системи лінгвістичного класу [5], що використовують лінгвістичний підхід до впорядкування інформації та для кожної сторінки оцінюють ступінь релевантності її тексту до запиту), для покращання позиції форуму в WWW доцільно модифікувати цей текст так, щоб у процесі ранжування сторінок форуму його сумарний ранг був високим. До кожного елемента контенту форуму доцільно застосовувати різні підходи до оптимізації та модифікації текстової інформації.

1.1.1. Аналіз елементів контенту форуму та способи їх модифікації

- **Назви підфорумів**

Назви підфорумів	Про проект	Topics	Posts	Last post
	 Загальна інформація Загальна інформація про е-книжки, читалки е-книг, технології, історію їх виникнення тощо.	11	12	Ел.книжка і ... by Doca 1 day 23 hours ago
	 Новини Найновіші тенденції на ринку е-читалок	4	4	Традиції ... by admin 02/05/2010 - 21:44
	 ЧаПи (FAQ)	26	26	Чим можна ... by Heldigvis Neste 01/09/2010 - 21:55

Рис. 2. Приклад назв підфорумів

Назагал, назви підфорумів відповідають певній тематиці, яка була визначена ще на етапі створення та розроблення форуму. Зокрема, назви підфорумів формують загальну тематику форуму, яка для пошукових машин є загальнішою метаінформацією про сайт, ніж семантика форуму [5]. Важливо правильно визначити назви підфорумів з самого початку під час створення форуму, оскільки потім кардинально змінити їх не вдасться. Основні вимоги до назв підфорумів такі:

1. Назви підфорумів повинні бути короткі, складатися з 1–3 слів, у межах одного форуму не повинно бути назв-синонімів.
2. Кожна окрема назва повинна чітко характеризувати підтему форуму.
3. Назви бажано давати іменниками та іменниковими словосполученнями, причому вибирати часто вживані в певній мові слова, які характеризуватимуть тему.

- **Заголовки тем дискусій**

Заголовки тем дискусій	Загальна інформація				
	New Blog entry   Mark all topics read				
	Topic	Replies	Views	Created	Last reply
	 Sticky: Електронний папір	1	380	by admin 09/16/2009 - 23:03	by Katja 12/19/2009 - 00:17
	 Sticky: Електронна книга	0	315	by admin 09/17/2009 - 17:53	n/a
	 Ел.книжка і холод	0	14	by Doca 02/23/2010 - 20:21	n/a
	 Створення електронної книги з паперової	0	42	by admin 02/16/2010 - 11:32	n/a

Рис. 3. Приклад назв заголовків тем дискусій

Більшість пошукових систем спочатку індексують заголовки, а потім вже повідомлення, оскільки вважається, що інформація в заголовку теми є важливішою і окреслює загальну тематику. Основними вимогами до створених користувачами заголовків тем є:

1. Стислість — назва заголовка не повинна бути довшою за одне речення.
2. Репрезентативність — заголовки і підзаголовки повинні містити такі ключові слова, за допомогою яких користувачі могли б собі в загальних рисах уявити, про що йтиметься у певній гілці форуму.
3. Змістовність — заголовки тем повинні передавати конкретний зміст у логічно зрозумілих формах та містити ключові слова з семантичного ядра форуму.

- **Повідомлення в дискусіях**

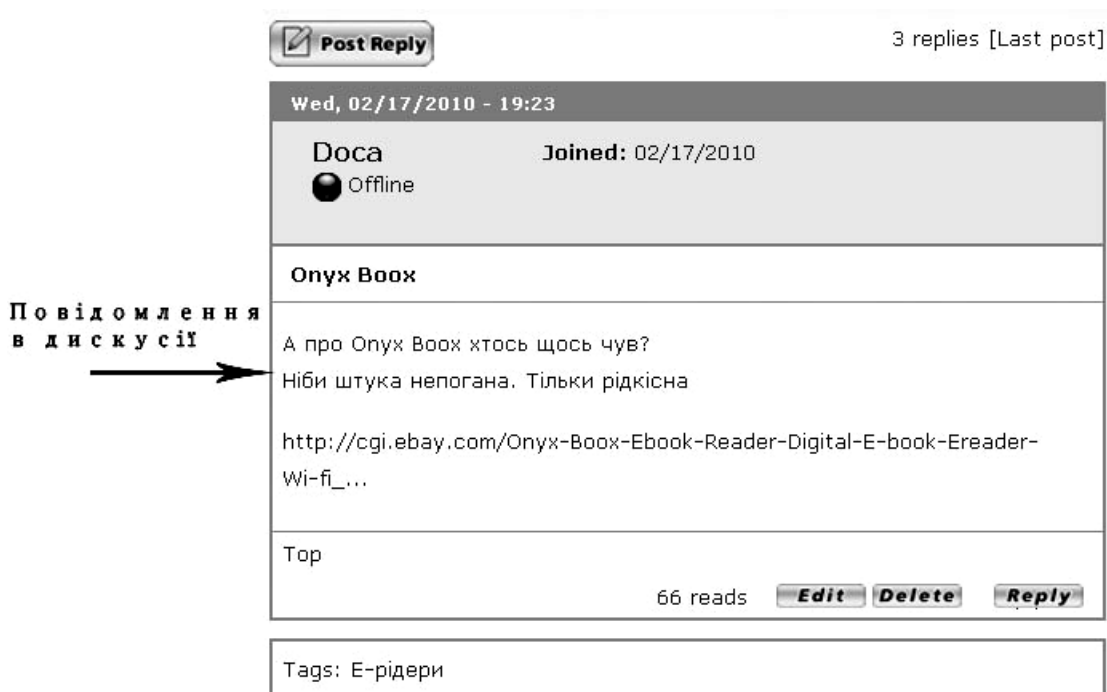


Рис. 4. Приклад повідомлення в дискусії

- Основні вимоги до створених користувачами повідомлень та дописів можна класифікувати так:
1. У перших повідомленнях в дискусіях необхідно повідомляти основну інформацію.
 2. У повідомленнях повинні з'являтися слова та фрази, які є синонімами до ключових слів з потрібної тематики і входять до семантичного ядра форуму або безпосередньо стосуються до теми.
 3. Частота вживання ключових слів на сторінці повинна бути в межах 3–7 %. Проте текст, в якому занадто часто вживаються ключові слова, може трактуватися пошуковими системами як спам і позиція документа в результатах пошуку може бути автоматично знижена [1].
 4. Важливим є "точне входження" шуканої фрази в текст повідомлень для запитів з декількох слів, тобто ідентичний запит із збереженням порядку слів. Тобто, чим більше оригінальних ключових слів вживається в дописах, тим вищий ранг надає пошукова машина під час ранжування.
 5. З повідомлень необхідно вилучати слова, які є небажаними в описуванні тематики сайту.
 - **Сторінки профілів користувачів** містять інформацію, яку користувачі хочуть надати про себе. Проте модифікація цієї інформації істотно не впливає на покращання позицій форуму.
 - **Підписи користувачів** відображаються одночасно і на сторінках профілів, і в повідомленнях користувачів. Дуже часто вони містять зовнішні посилання на інші сайти та контекстну рекламу, а також цитати з творів або авторські тексти самих користувачів.
 - **Статичне дзеркало форуму** відображає всю текстову інформацію з форуму без зображень і призначене для економії в оплаті за інтернет-трафік. Модифікація заголовків тем та повідомлень форуму означає автоматичну модифікацію текстів зі статичного дзеркала.
 - **Архів.** На деяких форумах немає статичного дзеркала, проте є архів, який також зберігає всю інформацію з обговорень на форумі.

1.2. Семантичне ядро форуму

Семантичне ядро – це список слів і словосполучень, які відображають тематику форуму і використовуються в його контенті. 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 може вважатися 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).

1.2.1. Частотність запитів

Залежно від частоти вживання у пошукових системах запити поділяються на:

1. Високочастотні запити – здебільшого складаються з одного слова загального іменника (найчастіше це збірні, речовинні та загальні назви предметів) або словосполучення – іменника, власної назви (антропоніми, топоніми та культуроніми). Використовуються в процесі пошуку інформації великою кількістю користувачів, що створює конкуренцію між форумами 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 застосовуються зрідка, але забезпечують форуму аудиторію користувачів, які точно знають, яка інформація їм потрібна, і задають відповідні запити.

1.2.2. Обсяг семантичного ядра форуму

Обсяг семантичного ядра залежить від кількості контенту форуму. Чим більше дописів та тем містить форум, тим більшим є семантичне ядро. Створюючи семантичне ядро, необхідно також враховувати, які типи запитів користувачі задають пошуковим системам в процесі пошуку інформації. Вони поділяються на:

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

2. Навігаційні (тип запитів, які користувач задає з метою знайти певний ресурс у WWW).

3. Транзакційні (тип запитів, які користувач задає з метою отримати посилання на сайт, де можна знайти товар, що його цікавить. Як правило, запити такого типу пов'язані з бажанням користувача купити товар)

1.2.3. Створення семантичного ядра форуму

На етапі створення семантичного ядра здійснюється аналіз ключових слів і фраз, які використовуються або ймовірно можуть використовуватися користувачами для знаходження інформації, яка відповідає тематиці форуму. Це один з найтрудомісткіших етапів роботи з оптимізації сайта. Сторінки форуму, оптимізовані під певні ключові слова, стабільно отримуватимуть 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 форуму. До семантичного ядра сайта необхідно внести всі ключові слова та вирази, які користувачі можуть задавати як запити пошуковим машинам.

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 підставі цих даних робиться висновок про включення фрази в семантичне ядро форуму.

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

1. Аналізування тематики форуму та створення початкового списку ключових слів. Для сайта-форуму доцільно визначити 10–50 ключових слів для кожної з підтем форуму.

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 (Наприклад: слово Форум – Aļhev), оскільки дуже часто користувачі вводять пошукові запити, не перемкнувши розкладки клавіатури.

4. Аналізування частот запитів, виявлених на форумі за списком зовнішніх посилань.

5. Видалення багатозначних ключових слів, невживаних у певній мові та таких, що неточно відповідають тематиці форуму.

Якщо форум великий і його семантичне ядро складається з великої кількості ключових слів, доцільно згрупувати їх за тематикою. Це дасть змогу розробити чітку структуру семантичного ядра форуму та забезпечить можливість легко знайти потрібні ключові слова.

Name	Type	Operations		
+ Forums	Blog entry, Forum topic, Poll	edit vocabulary	list terms	add terms
+ Про е-читалки книг	Blog entry, Page	edit vocabulary	list terms	add terms
+ Програмне забезпечення	Blog entry, Forum topic	edit vocabulary	list terms	add terms
+ Формати е-книг	Blog entry, Forum topic	edit vocabulary	list terms	add terms

Рис. 5. Приклад групування тем з семантичного ядра форуму

На рис. 5 зображено приклади згрупованих за темами частин семантичного ядра. Це, зокрема, «Про е-читалки книг», «Програмне забезпечення» та «Формати е-книг». Кожна з тем містить списки ключових слів з семантичного ядра форуму. Один з таких списків, а саме «Про е-читалки книг», наведено на рис. 6.

Terms in <i>Про е-читалки книг</i>		List	Add term
<i>Про е-читалки книг</i> is a free-tagging vocabulary. To change the name or description of a term, click the <i>edit</i> link next to the term.			
[more help...]			
Name	Operations		
Asus DR-950	edit		
Asus Eee	edit		
Bookeen CyBook Opus	edit		
e-book reader	edit		
e-paper	edit		
електронна книжка	edit		
Insdream SX-601	edit		
iRiver	edit		
Kindle 2	edit		
Kindle DX	edit		
LBook V3	edit		

Рис. 6. Приклади ключових слів із семантичного ядра форуму

2. Аналіз позицій форуму в середовищі WWW після практичного застосування запропонованих методів модифікації контенту

Запропоновані у статті методи модифікації контенту форуму застосовувались для покращання позицій новоствореного сайта-форуму «Український клуб електронної книги» ebook-ua.org протягом 20 тижнів.

2.1. Динаміка зміни кількості відвідувачів форуму

На рис. 7 наведено динаміку зміни популярності форуму «Український клуб електронної книги» (кількості відвідувачів на тиждень) після практичного застосування запропонованих методів модифікації контенту.



Рис. 7. Динаміка зміни популярності форуму

2.2. Динаміка зміни кількості зареєстрованих користувачів форуму

На рис. 8 зображено динаміку зміни кількості користувачів, які після перегляду сторінок обговорень виявили бажання стати членами спільноти форуму та зареєструвалися на форумі.



Рис. 8. Динаміка зміни кількості користувачів, які виявили бажання стати членами спільноти форуму

2.3. Кількість переходів на форум з пошукових систем

На рис. 9 зображено загальну кількість переходів за ключовими словами з чотирьох найпопулярніших та найпотужніших пошукових систем – Google, Yandex, Bing та Nigma за 20 тижнів (п'ять місяців). В цьому випадку не враховувалися переходи на форум користувачів за посиланнями, розміщеними на сторінках сайтів-партнерів, та за посиланнями контекстної реклами.

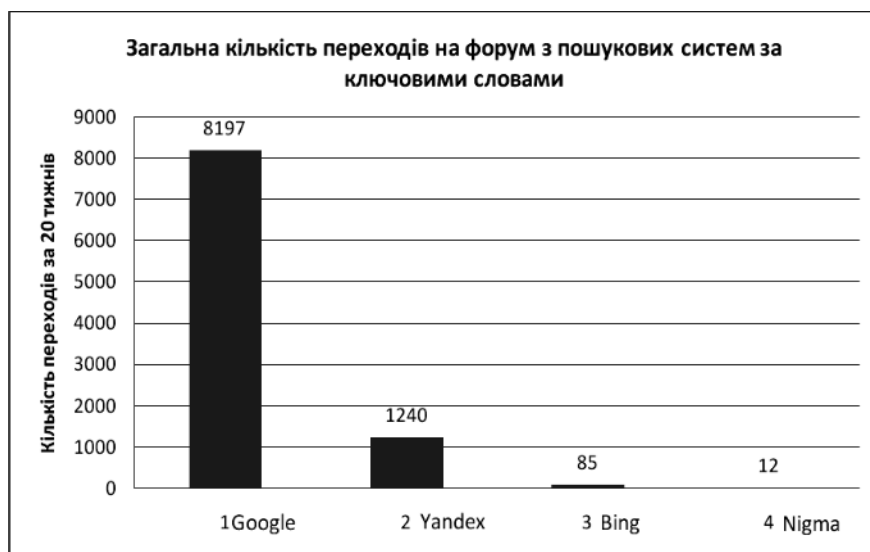


Рис. 9. Загальна кількість переходів на форум з пошукових систем за ключовими словами

Висновок

Мережа Інтернет надає безмежні можливості користувачам взаємодіяти та спілкуватися між собою. Проте з розвитком Інтернету створюється безмежно велика кількість сайтів, на які звичайним користувачам не завжди вдається потрапити, оскільки під час видавання результатів пошуку за певними запитами ці сайти потрапляють на соті, а то й тисячні сторінки результатів пошукової видачі. Те саме стосується і форумів. Бездоганно виконаний з програмно-технічного погляду проект буде гарантовано безрезультатним без формування позиції форуму в глобальному інформаційному середовищі. Однак це одне з найскладніших завдань в процесі оптимізації сайту для пошукових машин. Основним завданням, яке постає перед адміністраторами та модераторами форуму, є оптимізація контенту форуму, який міг би, з одного боку, бути цікавим та корисним користувачам, а, з іншого боку, розглядався пошуковими машинами як специфічний серед потоку подібної інформації. Для цього необхідно виділити основні елементи контенту форуму та вимоги до його якісної модифікації. Надзвичайно важливим є створення його семантичного ядра.

Запропоновані у статті методи модифікації контенту форуму використовувалися для покращання позиції сайту-форуму «Український клуб електронної книги» ebook-ua.org протягом 20 тижнів (п'ять місяців). За цей час вдалося досягти значних результатів, зокрема, значно зросла кількість відвідувачів за тиждень та кількість членів спільноти, що свідчить про поліпшення позицій форуму у глобальному інформаційному середовищі WWW.

Наповнення контенту форуму ключовими словами з семантичного ядра форуму забезпечило постійний потік користувачів з найпопулярніших пошукових систем, зокрема Google та Yandex.

1. Ашманов И.С. Продвижение сайта в поисковых системах / И.С. Ашманов, А.А. Иванов – М.: Вильямс, 2007. – 304 с. 2. Гусев В.С. Google. Эффективный поиск / В.С. Гусев – М.: Вильямс, 2006. – 240 с. 3. Евдокимов Н.В. Основы контентной оптимизации. Эффективная Интернет-коммерция и продвижение сайтов в Интернет / Н.В. Евдокимов – М.: Вильямс, 2007. – 160 с. 4. Зуев М.Б. Продвижение сайтов в поисковых системах. Спасательный круг для малого бизнеса. / М.Б. Зуев, П.А. Маурис, А.Г. Прокофьев – М.: Бином. Лаборатория знаний, 2006. – 304 с. 5. Пелецишин А.М. Позиціонування сайтів у глобальному інформаційному середовищі: Монографія. – Львів: Вид-во Нац. ун-ту «Львівська політехніка», 2007. – 260 с. 6. Серов Ю.О. Аналіз комунікативних процесів у Веб-спільнотах середовища Веб 2.0 / Ю.О. Серов, А.М. Пелецишин, К.О. Слобода // Східно-Європейський журнал передових технологій. – 2009. – № 1/2 (37). – С. 38–41.