

УДК 004.9

**В'ячеслав ЗОСИМОВ**

[zosimovvv@gmail.com](mailto:zosimovvv@gmail.com)

ORCID: 0000-0003-0824-4168

**Світлана САЧЕНКО**

[lana-avramenko@ukr.net](mailto:lana-avramenko@ukr.net)

ORCID: 0000-0001-7389-738X

м. Миколаїв

## АНАЛІЗ ЯКОСТІ ІНФОРМАЦІЙНОГО НАПОВНЕННЯ САЙТУ

*Дана стаття розкриває важливість аналізу якості інформаційного наповнення для сайту для забезпечення його повноцінного функціонування, корисності та зручності поданої інформації для відвідувачів, підвищити у списку видач пошукових систем.*

*В роботі визначено основні напрямки аналізу якості контенту, надана характеристика основних етапів пошукової оптимізації веб-сайтів.*

*Ключові слова: Page Rank (PR), тематичний індекс цитування, контент, пошукова оптимізація, релевантність, ключові слова, семантичне ядро.*

### Постановка проблеми

Щодо важливості контенту виникало так багато дискусій, що сама суть проблеми створення унікального контенту на сайті втратилась в безмежній кількості спречань про необхідність або, навпаки, обов'язковість унікального контенту для вдалої пошукової оптимізації сайту. В контексті даного питання треба зазначити, що саме по собі призначення Інтернету, як джерела швидкого отримання найбільш повної і змістовної інформації відповідно до запиту, не приймається до уваги. Пошукові системи ведуть неперервну боротьбу за якість пошуку між собою, підвищуючи будь-яким чином релевантність пошукових результатів, з оптимізаторами підвищення релевантності пошукової видачі, останні, в свою чергу, конкурують один з одним за підвищення позиції сайтів, намагаючись, суміжно обійти пошукові системи. Але користувача, як і раніше, ще до виникнення SEO, приваблює цінна інформація, тому задача розробника сайту подати її в найбільш зручному для сприйняття вигляді, саме звідси постає проблема оцінки якості інформаційного наповнення сайту.

### Аналіз останніх досліджень і публікацій

На сьогоднішній день, перехід ринкових відносин на онлайн-простір, сприяв появі нового методу – контент-маркетингу. Одним із завдань якого є аналіз інформаційного наповнення сайту. Виходячи з цього, дослідження маркетингових компаній в даному напрямку показали, що для аналізу якості маркетингових комунікацій важливо оцінити, перш за все, ступінь сприйняття цільовою аудиторією інформаційного контенту. У зв'язку з цим слід здійснювати оцінку вмісту сайтів і зручності роботи користувачів з інформацією, розміщеною на інформаційному ресурсі [1]. Треба зазначити, що, розглянуті джерела, надають тільки загальні принципи та методи здійснення аналізу контенту [2, 3, 4].

У проведених дослідженнях оцінюються такі параметри, як функціональність і дизайн. Деякі методики, використані в розглянутих дослідженнях, представляють простий експрес-аналіз [2], в основі інших лежать більш емкі і складні алгоритми отримання підсумкових кількісних характеристик і формування на їх основі рейтингів сайтів з однаковою тематикою. Рейтин-

нговий підхід дозволяє проводити порівняльний аналіз сайтів, причому підсумкова кількісна міра рейтингу показує, наскільки краще або гірше показники відповідного сайту в порівнянні з лідером рейтингу.

Порівняння сайту з іншими не надає повної оцінки якості інформаційного наповнення. Тому проблема визначення основних критеріїв, за якими можна дати більш адекватну оцінку веб-сайту залишається не вирішеною.

#### Постановка завдання

Метою даної статті є розглянути аналіз контенту сайту, основні етапи пошукової оптимізації веб-сайтів, визначити основні напрямки аналізу якості інформаційного наповнення сайту.

#### Виклад основного матеріалу

Деякі системи при видачі сайтів в пошукових результатах здатні аналізувати інформацію не тільки про частоту ключових слів, а й про найбільш популярні сторінки та кількість часу, який витрачається користувачами на їх перегляд. Чим ці параметри більші, тим вища ймовірність того, що на сайті розміщена дійсно якісна інформація. Інший підхід передбачає визначення кількості посилань і обчислення авторитетної ваги майданчика. Це такі показники, як ТІЦ, PR та інші, механізм визначення яких також є ще одним алгоритмом роботи пошукових систем. Кількість параметрів, що враховуються пошуковими алгоритмами, може досягати декількох сотень. Наприклад, система «Google» при ранжуванні сайтів звертає увагу на більш ніж 200 факторів.

Page Rank (PR) – дослівно перекладається з англійської як «ранг сторінки». Простіше кажучи, це оцінка, яку пошукова система Google ставить сайту в цілому та його окремим сторінкам. Оцінюється сайт за десятибальною системою та відображає «відношення» пошуковика до сайту: чим вище оцінка, тим вища «повага» Google до сайту, тим вище сайт буде стояти у видачі [5].

Ніхто точно не знає, як Google розраховує PR.

$$PR(A) = (1 - d) + d(PR(T_1)/C(T_1) + \dots + PR(T_n)/C(T_n)).$$

Так виглядає приблизна формула розрахунку PR, де  $T_1, \dots, T_n$  – сторінки, що посилаються на  $A$ ,  $C(T_n)$  – кількість вихідних посилань на відповідний сторінці,  $d$  коефіцієнт зазвичай рівний 0.85.

Оскільки посилань може бути багато, і загальна кількість сторінок в пошуковій системі Google досить велика (близько десятка мільйонів), а також їх кількість постійно зростає, то представляти вагу сторінки в абсолютних значеннях для вебмайстрів було б вельми неправильно. Для цього ввели поняття TLPR – Toolbar Page Rank – значення PR, яке приймає значення від 0 до 10 (шкала в панелі інструментів Google).

Для того, щоб укласти всі ваги сторінок між значеннями від 0 до 10 використовують логарифмічну шкалу. Визначається ToolBar Page Rank за формулою:

$$TLPR = \log_{base}(PR) \cdot a,$$

де  $base$  – основа логарифму, яка залежить від кількості сторінок в пошуковій машині (можливо і від ряду інших факторів), деякі приймають його рівним 7;  $a$  – коефіцієнт приведення, який задовольняє нерівність  $0 < a \leq 1$ .

У 2010 році на конференції представниками Google було заявлено про поступове переважання інших факторів і відмирання Page Rank (і не тільки тулбарного).

У жовтні 2013 року Метт Каттс припустив, що користувачі не побачать оновлення тулбарного Page Rank до кінця року. Востаннє Page Rank був оновлений 6 грудня 2013 р.

У жовтні 2014 року аналітик поведінки вебмайстрів Google Джон Мюллер (John Mueller) повідомив в ході відео-конференції для веб-майстрів Webmaster Hangout, що наступного оновлення тулбарного Page Rank не буде ніколи: «Ми не оновлювали Page Rank вже більше року і, ймовірно, не будемо оновлювати його в

майбутньому – принаймні, тулбарний Page Rank» («Page Rank is something that we haven't updated for I think over a year now, and we're probably not going to be updating it going forward, at least in the Toolbar Page Rank»).

У квітні 2016 року Google офіційно відключив показ тулбарного значення PR і прирівняв його до нуля всім сайтам.

Яндекс розраховує тематичний індекс цитування сайтів за спеціальним алгоритмом, який враховує як подібність сайтів, що посилаються з тематики, так і рівень їх власного ТІЦ. На визначення показника ТІЦ не впливають посилання з безкоштовних каталогів та інших маловідомих ресурсів. ТІЦ сайту є динамічним коефіцієнтом і приблизно два рази на місяць відбувається його автоматичне оновлення (апдейт). Поширена думка, що на результати пошуку в Яндексі ТІЦ не впливає, проте, за деякими даними, з введенням алгоритму MatrixNet ТІЦ сайту став одним з факторів ранжування, що враховується при розкрутці сайтів, можливо, не дуже значущим. ТІЦ насамперед, застосовується з метою визначення порядку розташування сайтів в розділах Яндекс.Каталогу. Усі без винятку сайти мають бути проіндексовані даною пошуковою машиною. Причому на сторінках Яндекс.Каталогу наводяться округлені цифри – вони і дозволяють зробити висновок про статус того чи іншого веб-сайту в розділі.

Пошукові сервери при реєстрації Web-сервера, сайту або сторінки використовують інформацію, що зберігається в спеціальних тегах META, які містяться між тегами і більшість початківців Web-майстрів та Web-дизайнерів не приділяють особливої уваги цим тегам. Теги META не видно при перегляді документа і помітно збільшують його розмір, проте без них не обійтися. Рекомендації зі складання Мета-тегів.

Правильність складання META тегів робить можливим, щоб ваш сайт було легко відшукати за допомогою пошукових систем серед величезного числа інших

сайтів! Для цього необхідно слідувати наступним рекомендаціям:

1. Довжина вмісту тегів META «descriptions» не повинна перевищувати 200 символів, а «keywords» – 1000 символів. Це пов'язано з тим, що пошукові сервери, як правило, використовують саме такі величини при індексуванні інформації на вашому сайті. Можна отримати негативний результат, якщо пошукові машини відкинуть частина опису, що перевищує встановлені норми, в результаті чого в каталог потрапить не вся бажана інформація, або сервер просто відкине реєстрацію вашого сайту.

2. Потрібно включити всі найбільш часто використовувані в вашому сайті терміни та їх синоніми в тег «descriptions». Можете включити і терміни, що не мають відношення до ваших сторінок, однак в Internet така поведінка вважається некоректною. У список ключових слів можна вставляти і фрази. У цьому випадку у вашого сайту підвищується шанс потрапити в самий початок списку, виданого пошуковим сервером (у разі збігу фрази з тією, що вводить користувач).

3. Для збільшення рейтингу сайту та потрапляння його у верхню частину пошукового списку необхідно кожен термін, що включається в опис, повторювати близько 3 разів. Пошукові машини, як правило, працюють з тегами META, а в разі їх відсутності – вже з вмістом сторінки. Іноді це призводить до появи сайту у списку top 10, але може призвести і до негативного результату (перед реєстрацією перевірте пошукову машину, давши запит і переглянувши ключові слова у сайтів, що стоять на верхніх позиціях).

4. Використовуйте в описах і списку термінів як можна більше синонімів.

5. Не використовуйте в описах і в списку термінів поняття і слова, слабо пов'язані один з одним.

Склад META-тегів:

1) для того, щоб при перегляді документа браузер брав свіжу версію докумен-

та, а не версію з кешу використовується [6]:

```
<META HTTP-EQUIV = "Expires" Content =  
"Mon, 28 Mar 1999 00:00:01 GMT">
```

2) для визначення броузером мови, на якому написана сторінка використовується:

```
<META HTTP-EQUIV = "Content-Type"  
Content = "text/html; Charset =  
Windows-1251">
```

3) якщо недосвідчений Web-майстер скопіює Вашу сторінку, в якій залишить наступний тег, то при заході на цю сторінку броузер автоматично перейде на Вашу сторінку

```
<META HTTP-EQUIV = "Refresh" Content =  
"1; url = http://адреса_сторінки">
```

4) для зазначення імені автора та додаткової інформації про нього використовується:

```
<META Name = autor" Content = "ПІБ, E-  
Mail...">
```

5) для зазначення авторських прав використовується:

```
<META Name = "copyright" Content = "Ва-  
ша_фірма...">
```

6) для зазначення ключових слів і термінів використовується

```
<META Name = "keywords" Content = "Ва-  
ші_Ключові_слова">
```

7) для зазначення короткого опису використовується

```
<META Name = "description" Content =  
"Опис_сторінки">
```

Покладаючись на власні дослідження в цій області, а також на професійних розробників в цій галузі ми рекомендуємо вам при наявності великої кількості сторінок створити теги мета-інформації з різним змістом, що описують зміст даної конкретної сторінки чи розділу.

Оскільки просування сайтів це поліпшення становища ресурсу в пошуковій видачі, необхідно розуміти, за рахунок чого це відбувається. Тому пропонуємо розглянути етапи пошукової оптимізації веб-сайтів (табл. 1).

З іншого боку, кожен пошукову систему розробляли люди. А вони, задаючи програмі критерії відмінності якісних сайтів від неякі-

сних, спиралися аж ніяк не на фізичні закони, а на власне уявлення про інтернет-ресурс. Відповідно, ми можемо припустити, за якими ознаками людина, прийшовши на якийсь сайт визначить його корисність.

Щоб ефективно вести пошукове просування сайту, необхідно зібрати для нього масив ключових слів, що відповідають тематиці і контенту сайту, який просувається. Такий список слів отримав назву семантичне ядро, воно рівномірно розподіляється по всьому сайту відповідно до його структури.

Пропорційно масштабу сайту розмір семантичного ядра може бути величезним, нараховуючи десятки, а в деяких випадках навіть сотні тисяч ключових слів різної довжини, а може обмежитися лише кількома десятками ключів.

Весь процес складання семантичного ядра сайту можна розділити на дві стадії:

- підбір ключових слів (для отримання максимального ефекту він повинен бути досить грамотним і скрупульозним);

- розподіл зібраних ключових слів по сайту відповідно до його структури.

Кожну категорію сайту, будь то сторінка або розділ, потрібно заточити під відповідний пошуковий запит. В середньому для одного розділу або сторінки підбирається від п'яти до двадцяти п'яти ключових слів, але ця кількість може істотно змінюватися для різних типів сайтів.

Для складання семантичного ядра, необхідно вибрати найважливіші напрямки для подальшого пошукового просування. Потім під кожний з цих напрямків потрібно зібрати відповідні ключові слова.

Слід розуміти, що для підбору ключових слів потрібно певні ресурси часу і грошей. Для великих сайтів потрібно використовувати платні бази слів, для більш ефективного і швидкого складання семантичного ядра.

Наповнення сайтів текстом – важливий момент життєво необхідний для успішного існування сайту. Завдяки постійному наповненню сайту новим цікавим контентом можна не тільки привернути увагу

відвідувачів, а й утримати їх на сайті, змусити постійно відвідувати сторінки в пошуках інформації. Проведення аналізу на пряму пов'язано із перевіркою інформації, наскільки вона здатна впливати на корис-

тувачів і відповідає тематиці ресурсу. Дослідження охоплює також визначенні частоти слів, релевантність ключових слів. Спеціалісти виділяють дві методики проведення аналізу контенту:

Табл. 1. Основні етапи пошукової оптимізації веб-сайтів

№	Етап	Обсяг робіт
1.	Аудит сайту	<p>Наявність сайту в Яндекс.Вебмастер і Google Webmaster. Встановлення сервісів статистики: liveinternet Яндекс.Метрика, Google analytics (не обов'язково все). Перевірка на фільтри пошукових систем. Аналіз індексації сторінок. Вживання атрибуту rel = "canonical" на головний документ для ідентичних сторінок. Наявність і повнота карти сайту (sitemap.xml). Наявність і коректність robots.txt. Коректність обробки 404-помилки. Доступність ресурсу тільки з www або тільки без www. Аналіз показників ресурсу. Перевірка швидкості роботи сервера і завантаження сайту. Оцінка коду сторінок. Видалення зайвого і надмірного коду. Для магазинів і організацій наявність даних про компанію на картах Яндекс та Google. Наявність найбільш повних контактах даних. Відсутність на сайті чорних методів оптимізації. Оцінка якості контенту. Аналіз зовнішніх і внутрішніх посилань. Аналіз тегів title, description. Використання мікроформатів. Наявність YML-карти для Яндекс. Наявність на сайті кнопок соціальних мереж. Оцінка посилальної маси на ресурс. Наявність і оцінка фавікона.</p>
2.	Внутрішня оптимізація	<p>Структуризації контенту. Різноманітність контенту: текст, картинки, відео і так далі. Унікальність контенту. Використання тегів strong, b, em. Використання системи внутрішньої перелінковки. Ключове слово, що присувається, зустрічається в контенті документа мінімум 1 раз, але не більше 9% від всього тексту. Сторінка, що просувається, має текстовий контент не менш 500 символів.</p>
3.	Зовнішня оптимізація	<p>Плавне нарощування посилальної маси. Розміщення посилань на якісних майданчиках. Різноманітність посилальної маси. Релевантний title зворотного посилання. Організація конкурсів та інформаційних подій, як метод нарощування зворотних посилань.</p>

- кількісна;
- якісна.

Якісний аналіз дозволяє визначити рівень сайту з точки зору самих відвідувачів:

1. Якість запропонованого матеріалу.
2. Рівень подання інформації.
3. Степінь наданих очікувань.

Кількісний аналіз оцінює рівень впливу інформації на читачів шляхом визначення кількості структурних одиниць, частоти слів, їх смислових характеристик. Грунтуючись на таких підрахунках, розуміють наскільки сайт буде успішно виводитися пошуковими системами. Наявність ключових слів і їх щільність, розмір тексту, а також інші характеристики визначають рівень сприйняття наявного контенту відвідувачами.

Оптимізація сайту напряму залежить від контенту (відео-контенту, тексту графічних матеріалів). Виявленні недоліки можуть призвести до серйозних проблем для сайту, в тому числі сторінки не будуть визначатися при запитах в пошукових системах, а відповідно про жодну цілісність пошукового трафіку не може бути і мови. У зв'язку з цим проведення аналізу контенту сторінки вкрай важливе, щоб упевнитись в повній його відповідності вимогам, які висувають пошукові системи. Аналіз контенту передбачає перевірку змісту текстових масивів та продуктів комунікативної кореспонденції.

В першу чергу перевіряється унікальність інформації, яка знаходиться на кожній із сторінок сайту. Для її перевірки існує багато програм, а також онлайн-ресурсів, які дозволяють отримати дані практично моментально. Унікальність повинна бути не менше 98%, перед тим як виставити інформацію з нижчими показниками унікальності, її коригують, досягаючи таким чином кращих показників.

Якісно проведений контент-аналіз дозволяє створити список відповідності сторінок запитам і при необхідності внести відповідні корективи. Розглянемо два способи, які дозволяють створити такий список:

- «від пошукових систем»;
- «від сайту».

За допомогою методу «від пошукових систем» можна визначити релевантні сторінки сайту і запиту, за яким сайт буде легше знайти пошуковою системою. Він дозволяє виявити, якого контенту не вистачає з точки зору пошукової системи, а також автоматизувати процес проведення аналізу сайту.

Перевірка інформаційного наповнення сайту другим способом являє собою розгляд окремо кожної сторінки, аналізується група запитів із семантичного ядра, завдяки використанню яких сайт буде успішно просуватися в пошуковій системі.

Пошукові системи мають особливі вимоги до змісту текстової інформації: чіткість, лаконічність, інформативність і достовірність. Пошукові системи мають перелік рубрик веб-сайту, основні розділи інформаційних новин та посилання на свої найбільш популярні служби. Для більшої привабливості, сайт може містити кілька зображень та заголовки основних подій економічних, політичних, спортивних, культурних, тощо. Як правило, такі системи дають можливість користувачеві за власними уподобаннями дещо змінити вигляд сайту. Обов'язковою вимогою є наявність рубрик служби підтримки інформаційної системи.

Таким чином можна виділити такі п'ять основних напрямків аналізу якості інформаційного наповнення сайту:

1. Відповідність семантичному ядру сайту.
2. Розгляд сайту з точки зору відвідувачів – це визначає цінність для користувача контенту сторінок сайту, доступність і корисність поданих на сайті текстів.
3. Визначення кількості необхідного контенту. В даному випадку визначається оптимальна кількість сторінок сайту і аналізуються об'єми тексту на сторінках, як з точки зору пошукових систем, так і – відвідувачів.
4. Змістовний та лінгвістичний аспект реурсу. Дуже важливим фактором є перевірка

унікальності тексту на сторінках сайту, а також його орфографічних та стилістичних помилок.

5. Графічний аналіз включає в себе перевірку правильності оформлення текстів на сторінках, шрифти, заголовки, виділення, посилання тощо.

### Висновки і перспективи досліджень

В даній роботі розглянуть основні методи та принципи на яких базується якісне інформаційне наповнення сайту та його просування. Детально розглянуто аналіз семантичного ядра сайту, визначено дета-

лі, на які необхідно звертати увагу при аналізі контенту, також звернули увагу на якість інтернет-ресурсу з точки зору користувача. Подано основні напрямки аналізу якості інформаційного наповнення сайту.

Перспективою подальших досліджень може бути деталізація оцінювання ефективності інформаційного наповнення веб-ресурсів. Оскільки Інтернет-бізнес розвивається надзвичайно швидко, то й кількість комерційних сайтів збільшується стрімко, що зумовлює потребу у формуванні універсальної системи кількісної оцінки якості Інтернет-ресурсу.

### СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Контент-маркетинг – ефективний шлях просування вашого бренду [Електронний ресурс]. – Режим доступу: <http://www.semesprit.com/uk/kontent-marketing-efektivnij-shlyah-prosuvannya-vashogo-brendu/> від 23 жовтня 2017 року.
2. Комплексний аналіз сайту [Електронний ресурс]. – Режим доступу: <http://www.cy-pr.com/>
3. Контент сайту та його складові [Електронний ресурс]. – Режим доступу: <http://webstudio2u.net/ua/studio-web/686-kontent-saita.html>
4. Полянко, А. Особенности поисковой системы Google [Електронний ресурс] / А. Полянко // Site-seo.ru – поисковая оптимизация и интернет-маркетинг – Режим доступу: <http://site-seo.ru/info/seostati/GoogleMSNYa/40/>
5. Просування сайтів [Електронний ресурс]. – Режим доступу: <http://igroup.com.ua/seo-articles/pagerank-pr/>
6. PHPist [Електронний ресурс]. – Режим доступу: <http://phpist.com.ua/html/33-metatags/>

**Viacheslav ZOSIMOV, Svitlana SACHENKO**  
Mykolayiv

### THE ANALYSIS OF THE WEB-SITE CONTENT QUALITY

*The article explains the importance of the web-site content quality analysis for the purpose of its unimpaired operation, usefulness, usability of information given for its visitors, and upward movement in search engine results.*

*The research defines main evaluation categories of content quality, gives characteristics of main stages of search engine optimization.*

**Keywords:** Page Rank (PR), thematic citation index, content, search engine optimization, relevance, keywords, semantic core.

**Вячеслав ЗОСИМОВ, Светлана САЧЕНКО**  
Николаев

### АНАЛИЗ КАЧЕСТВА ИНФОРМАЦИОННОГО НАПОЛНЕНИЯ САЙТА

*Данная статья раскрывает важность анализа качества информационного наполнения сайта для обеспечения его полноценного функционирования, полезности и удобства представленной информации для посетителей, повысить в списке выдачи поисковых систем.*

*В работе определены основные направления анализа качества контента, дана характеристика основных этапов поисковой оптимизации веб-сайтов.*

**Ключевые слова:** Page Rank (PR), тематический индекс цитирования, контент, поисковая оптимизация, релевантность, ключевые слова, семантическое ядро.

Стаття надійшла до редколегії 12.02.2018