

## ЕВОЛЮЦІЯ ПОШУКОВОЇ ОПТИМІЗАЦІЇ

*Визначено основні етапи еволюції пошукової оптимізації та сучасний рівень розвитку технологій. Виявлені основні компоненти зовнішнього інформаційного аналізу сайтів, які беруться до уваги пошуковими системами під час ранжування електронного ресурсу. Визначені майбутні напрями удосконалення функціонування технологій пошукової оптимізації.*

*The main stages of evolution of search engine optimization and modern level of development of technology are defined. The main components of the external information analysis of the sites, being considered are revealed by search engines when electronic resources will be ranking. Future directions of improvement of functioning of technology of search engine optimization are defined.*

### 1. ПОСТАНОВКА ПРОБЛЕМИ

Для всіх електронних ресурсів присутніх в мережі Інтернет питання пов'язані з технологіями пошукової оптимізації (ПО) є вкрай важливими. Саме вони визначають позицію, на якій знаходиться посилання на сайт в видачі результатів запиту пошукової системи (ПС). Верхні позиції в ПС вказують на конкурентоспроможність електронних ресурсів (ЕР), адже зазвичай, потенційний користувач переходить на ресурси саме з верхніх позицій рейтингу. Технології ПО розвиваються динамічно, кардинально змінюючи алгоритми та правила своєї роботи. Тому, важливо окреслити основні технологічні етапи еволюціонування технології ПО. Оптимізація ЕР за новітніми правилами вимагає ґрунтовних знань в цій галузі. Застосовувати, до сучасних ресурсів, застарілі технології та методи, які допомагали в просуванні ЕР ще декілька років тому, зараз вже недоцільно та може призвести до абсолютно протилежних наслідків, наприклад, вилучення з рейтингу ПС або до блокування ресурсу.

### 2. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ

Теоретико-методологічною базу дослідження склали фахові видання та блоги провідних спеціалістів в галузі ПО [1-6 та ін]. В наведених працях детально розглядається методологія просування сайтів, технологічні аспекти ПО та алгоритмам роботи ПС, проте приділено

---

<sup>1</sup> НТУУ «КПІ»

недостатньо уваги еволюційним аспектам розвитку технології ПО. Аналіз фахової літератури дозволив виявити основні етапи еволюціонування та перспективні напрями розвитку технології ПО.

### 3. МЕТА РОБОТИ

Визначити основні етапи еволюції та сучасний рівень розвитку пошукової оптимізації, виявити резерви удосконалення її функціонування.

### 4. РЕЗУЛЬТАТИ ПРОВЕДЕНИХ ДОСЛІДЖЕНЬ

Етапи еволюції ПО пов'язані з розвитком інформаційних технологій та штучним інтелектом, який застосовують ПС, тому стрімко змінні технології формування пошукової видачі результатів запитів призводять до удосконалення алгоритмів роботи ПО. При дослідженні етапів еволюції необхідно розглянути історичні аспекти питання, які призвели до найпомітніших змін.

#### *Перший етап*

В середині 1990-х років з'явилися перші пошукові системи, які застосовували прості алгоритми пошуку, на яких і базувались процеси пошуку інформації в мережі Інтернет та ранжування ЕР в рейтингах ПС. Типовими прикладами ПС, які виконували тільки інформаційні запити, в 1995-1997 роках були загальновідомі AltaVista, Excite, Webcrawler і т.д. Тоді процес пошукової оптимізації називали «розкручування». Методи «розкрути» були прості та прямолінійні. Вони базувались на маніпулюванні ключовими словами контенту і мета-тегів та на зовнішніх посиланнях на дорвеї – (англ. doorways), сторінки або ресурси створені спеціально для роботів пошукових систем. В результаті таких дій ПС видавали не релевантні посилання на сформований запит, а проранговані рекламні посилання. Для просування ресурсу потрібно було створити більший обсяг маси посилань ніж у конкурентів. Ці методи давали можливість говорити о гарантіях та строках «розкручування» ЕР. Можливо було «розкручувати» окремі сторінки, а не сайт в цілому. Семантичне ядро системи, в той період, зовсім залишалось поза увагою. Як наслідок, це знизило якість роботи ПС та призвело до занепаду більшості з них.

#### *Другий етап*

Дослідження щодо створенню ПС з новою ідеологією пошуку стали логічним кроком розвитку технології ПО. У 1996 році С. Брін і Л. Пейдж, тоді ще аспіранти Стенфордського університету, почали роботу над дослідницьким проектом BackRub – пошуковою системою по Інтернету, що використовувала нову тоді ідею про те, що веб-сторінка

повинна вважатися тим важливішою, чим більше на неї посилається інших сторінок, і чим більше важливими в свою чергу є ці сторінки. Через деякий час BackRub була перейменована в Google. На початку 1998 року було опубліковано опис алгоритму ранжирування, PageRank та опис архітектури самої ПС. За якістю повернених пошукових результатів ця система значно перевершувала всі існуючі. Усвідомлення потенціалу нової ПС призвело до заснування у вересні 1998 року компанії Google Inc. та її подальшого розвитку як комерційного продукту. У 1998 році Google був однією з перших ПС, що впровадили посилальне ранжирування, завдяки чому значно поліпшилась якість пошуку запитів в порівнянні з конкурентами. Надалі, інші ПС розробили і впровадили свої аналоги PageRank та інші методи запито-незалежного ранжирування документів [1].

В 1998 році увійшов у широкий обіг термін «пошукова оптимізація» (англ. search engine optimization, SEO).

Цей етап характеризується використанням інформації, яка існує поза веб-сторінкою. Збиралась інформація про аналіз посилань (англ. link analysis); текст якорів (англ. anchor-text), що має вирішальне значення для навігаційних запитів; відстежувались данні, що передаються з http-запитом (англ. click-through data). Таким чином ПО стала брати до уваги структуру глобальної мережі та семантику запитів. Основними компонентами зовнішнього інформаційного аналізу стали: відстеження кліків (англ. tracking clicks), репутація сторінки (page reputation), індекс популярності (link popularity), кількість часу, що проводять відвідувачі на сторінці – темпоральні спостереження (англ. temporal tracking) та якість посилань (англ. link quality) [2]. Потім, до переліку аналізуємої інформації додали вектори термів (англ. term vectors), аналіз статистики відвідування (англ. stats analysis), кеш-дані (англ. cache data) і контекст. В якості аналізу контексту розглядався пошук пар ключових слів. Це дозволило покращити категоризацію EP та побудувати ранжування на аналізі даних, що передаються під час http-запиту. ПС другого покоління підтримували як інформаційні, так і навігаційні запити. Нові алгоритми роботи ПС значно підвищили показник ефективності пошуку інформації, але ще мало місце використання методів «розкрути».

#### *Третій етап*

Розвиток технологій третього покоління ПО базувався на спробі ПС поєднати дані з різних джерел для досягнення головної мети – видачі результату, що відповідає потребі користувача. Завданням третього етапу було вийти за рамки обмежень фіксованих баз даних за допомогою семантичного аналізу, визначення контексту пошуку, вибору динамічної бази даних і т.п., щоб забезпечити інформаційні, навігацій-

ні та транзакційні запити. Третє покоління пошукових технологій було покликане об'єднати масштабованість існуючих ПС з новими та удосконаленими моделями пошуку релевантності. ПО прийшлося мати справу з новими технологіями ПС які збагатились картами Вебу, які застосовуються для фільтрації – видалення дублікатів сайтів, а також багатьох самостійних сторінок, які привертають трафік на всього лише декілька ключових слів. Це означало, що сторінки типу дорвеев, гейтвеев (англ. gateways) та вхідної (англ. entry, splash), тобто, спеціально створених спам-сторінок для цільової розкрутки сайту на визначення позицій ключових слів, після фільтрації було заблоковано. Також, значна увага була приділена профілюванню пошукових запитів конкретних користувачів. Всі основні ПС запланували створення персональних профілів та агентів, які мають накопичувати знання про користувача протягом певного періоду часу та використовувати їх виходячи з минулих пошукових звичок.

Поява Семантичного Вебу в 2001 році надала нові можливості та ще більше диференціювало поняття інформаційного пошуку. Семантична павутина – це надбудова над сучасною Всесвітньою павутиною, що покликана зробити інформацію, розміщену в мережі, зрозумілішою для комп'ютерів. Семантичний Веб надав можливість використовувати існуючу семантичну інформацію. Змінився алгоритм ранжування результатуючих документів – з'явилося поняття семантичного ранжування ER.

З 2007 року Google використовували вже 200 сигналів ранжування.

У 2009 році в Яндекс, одночасно із введенням пошукового алгоритму «Сніжинськ», з'явився новий метод машинного навчання, який назвали MatrixNet. Особливість цього методу полягає в тому, що ПС піддається навчання і може враховувати величезну кількість різних факторів ранжування. Завдяки цьому відсіюються нетематичні та нерелевантні сторінки, котрі містять тематичні запити. Також, система вміє змінювати формулу ранжування для різних класів запитів, завдяки чому досягаються більш точні результати пошуку [3].

Поява рубрикації пошуку за різними типами подання інформації (веб-документи, відео, рисунки, новини, книги, блоги, патенти, карти, місця та ін.) спростило для користувача процес пошуку, але поставило нові задачі для ПО. З'явилась потреба в розробці нових методів ПО. Модернізовані технології ПО врахували зміни та почали звертати увагу на вік домену, динаміку та історію посилань сайту. З'явився копірайтинг, який дозволив обійти алгоритми ранжування. Як наслідок, значно зросла складність ПО.

#### Четвертий етап

Рейтинг пошукових систем світу виявив що найбільш затребуваною сьогодні є система Google, її частка складає 47% (див. рис.1).

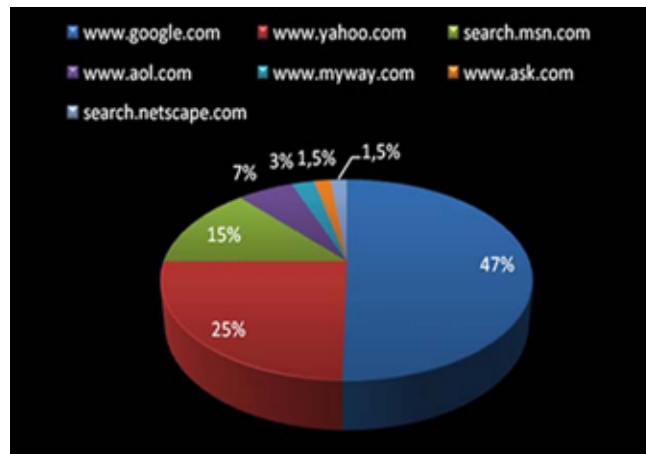


Рис.1. Рейтинг пошукових систем світу

Маючи на меті виявити сайти з неякісним контентом та пониження їх позицій у ранжуючих списках, у квітні 2011 року Google запусив новий алгоритм пошуку в Інтернет під назвою Panda. Це змусило спеціалістів в галузі ПО переоптимізувати ресурси відповідно до новітніх вимог.

В середині березня 2012 стало відомо, що Google готує до запуску новий фільтр проти переоптимізованих ресурсів. Основна ідея нововведення – бажання надати сайтам з якісним контентом можливість займати вищі позиції, ніж ресурси з поганим змістом, але з ознаками активної ПО. Новий алгоритм Google Penguin (саме так був названий новий фільтр Google), був запусканий інтернет-гігантом 24 квітня 2012 відразу у всьому світі. Основною метою фільтра стала боротьба з «пошуковим спамом». В якості основних ознак пошукового спаму були позначені надлишкове заповнення сторінок сайту ключовими словами, а також наявність на ресурсі «неприродних» посилань [4].

Два нових фільтри розділили сфери контролю: за якість контенту відповідає Google Panda, а Google Penguin – за кількість і якість посилань.

Регулярні оновлення фільтрів змусили веб-майстрів коригувати роботи пов'язані з ПО ресурсів. Застосування нових фільтрів ПС здебільшого торкнулось сайтів, які займаються електронною комерцією. Деякі ресурси втратили чверть трафіку, який надходив з Google, а де-

які ресурси більше не видавались в результатах пошукового запиту Google, оскільки посилання в футері, наскрізні посилання, посилання з нових і відвідуваних сайтів Google Penguin блокує. Спрощення алгоритму роботи фільтра не планується, бо Google визнав Google Penguin успішним фільтром.

У 2012-му році істотно зросла частка мобільного пошуку. Стрімко зростаюча популярність планшетів і смартфонів і зміна технології 3G на 4G, змусили переоптимізувати дизайн ресурсів для коректно відображення на невеликих дисплеях дивайсів. Також, почали враховувати мобільний трафік, оскільки з кожним роком кількість замовлень, які надходять з пристроїв на iOS, Windows Phone і Android, збільшується.

Істотно збільшилось використання голосового пошуку і пошуку на карті. Оскільки Siri і Voice Search від Google стали точнішими і зручнішими, все більше користувачів смартфонів використовують цю опцію. Відповідно, оптимізатори змушені переглядати підходи до оптимізації в роботі з ключовими запитами. При голосовому пошуку важливо враховувати, що користувачі говорять не літературною, а розмовною мовою.

Характерною ознакою цього етапу стало прагнення ПС нівелювати вплив «нечесних» стратегій просування. Google навіть залучив до цієї боротьби самих оптимізаторів, запропонувавши їм кнопку, що дозволяє відключати небажані посилання. У той же час вага дійсно корисних для користувачів посилань істотно зросла. Це змінило ідеологію розуміння ПО. Відтепер посилання мають набути максимально природнішого вигляду, а швидкість приросту маси посилань має стати сталою, також враховується різноманітність джерел ЕР.

Зміни в алгоритмах фільтрації ЕР призвели до збільшення часу просування ЕР в видачі списків пошукових запитів та до здороження робіт пов'язаних з оптимізацією сайтів. До позитивних сторін нововведень слід віднести застосування комплексного підходу до оптимізації ЕР та планомірної стратегії просування сайту в мережі. Поступово сайти, які обирають за основу якість, витісняють з пошукової видачі ресурси, що зробили ставку на кількість.

На даний час четверте покоління розвитку ПО проходить етап свого становлення. Одне з ключових завдань цього етапу повністю виключити методи «сірої» та «чорної» оптимізації, які базуються на застарілих та грубих методах «розкрути».

## 5. РЕЗЕРВИ УДОСКОНАЛЕННЯ ФУНКЦІОНУВАННЯ ПО

Світова статистика за 2012-й рік показала, що значення SEO не зменшилось. Щомісяця 450 млн. людей вводять в пошукових системах близько 11 млрд. запитів. Близько 90% користувачів приходять на ре-

сурси з сайтів пошукових систем. До 75% користувачів переходять виключно за посиланнями з органічної видачі, ігноруючи контекстну рекламу. Серед покупців інтернет-магазинів приблизно 42% складають люди, що прийшли з пошукових машин. Про нові сайти близько 46% користувачів дізнаються з результатів пошукової видачі. Стає очевидним, що люди, котрі ігнорують ПО не матимуть притоку трафіку з Інтернету.

Враховуючі останні зміни технології, можна окреслити напрями удосконалення функціонування ПО. Вони стосуються наступного: писати статті з використанням ключових слів, при аналізі ключових слів (семантичного ядра) засобами Google key word tool уникати слів загального вжитку; прописувати назви зображень, адже вони також індексуються; аналізувати сайти конкурентів, а саме структуру, контент, технічні фактори; враховувати регіон; визначитись з аудиторією; систематизувати інформацію та слідкувати за якістю контенту, адже корисний контент сам по собі притягує посилання; визначитись з акцентами; виокремити категорії та підкатегорії. При оптимізації сайтів необхідно користуватись засобами аналітики, які надають ПС.

Робота з технічною частиною матиме наступні особливості: потрібно робити дублі URL-адреси з www та без www; структурувати URL-адресу та налаштувати зрозумілу для людини URL-адресу, яка складається з транслітерації латинськими літерами тем рубрик, а не з цифр та символів.

Поява нових фільтрів ПС сприятиме добросовісній конкуренції в мережі, а результати видачі пошукових запитів стануть дійсно відображати авторитетні EP, які дбають про свою репутацію, відслідковують зміни пов'язані з ПО та демонструють сталий розвиток.

Роботи, пов'язані з ПО ресурсу, мають стати невід'ємною складовою його функціонування під час всього життєвого циклу та комплексними, тобто стосуватись, як внутрішньої так і зовнішньої частини сайту.

Тенденції розвитку в сфері ПО дозволяють спрогнозувати майбутні напрями досліджень та робіт. Наприклад, отримає розвиток алгоритм семантичного пошуку Knowledge Graph (Граф знань) – розпізнавання мови і природної мови запитів. Бо саме на ці аспекти мобільного пошуку і пошуку Google в цілому сьогодні спрямовані основні зусилля розробників [5]. В майбутньому SEO-фахівцям доведеться поєднувати оптимізацію пошуку для користувачів стаціонарних комп'ютерів і мобільних пристроїв, орієнтуватися не тільки на текстові, а й на голосові запити.

Стратегії, які поєднують ПО і просування в соціальних мережах, стають все більш затребуваними. Безумовно, це стане одним з ключових факторів успіху ЕР в майбутньому.

Оскільки ПС можуть визначити якість контенту, тільки покладаючись на думку користувачів, то в майбутньому на верхніх позиціях виявляться сайти, які дадуть найбільш цікаві та актуальні статті. Перевагу надаватимуть великим, за обсягом, публікаціям, бо в стислих складно надати вичерпну інформацію, через це великі за обсягом статті ранжуватимуться краще.

Посилити роботу з Google+ буде дуже корисно, адже оптимізатори відзначають, що «+» і пости в цій соціальній мережі впливають на позицію ресурсу. Через переплетення з пошуковими алгоритмами Google ця соцмережа стане в майбутньому році ще більш важливим фактором, який вплине на ранжування ресурсів [6].

ПС приділятимуть підвищену увагу чинникам, які знаходяться поза сторінкою сайту. Майбутнє ПО за аналізом поведінки цільових користувачів. Механічний підрахунок ключових слів і посилань залишиться в минулому.

Прогнозується збільшення значення SEO. Використання прозорих методів ПО в подальшому значно полегшить просування ЕР, які мають природні посилання на свій ресурс та тривалий час присутні в глобальній мережі.

Враховуючи, що основний прибуток і Google, і Yandex отримують завдяки продажу контекстної реклами, можна припустити, що ПС будуть шукати способи збільшити частки платного пошукового трафіку, зменшуючи органічний. Звідси висновок: конкуренція між оптимізаторами, які за основу маркетингової стратегії прийняли просування сайту в ПС, стане ще гостріше.

## 6. ВИСНОВКИ

Проведений аналіз літературних джерел дозволив визначити чотири основні етапи еволюції технології пошукової оптимізації.

Окреслено завдання сучасного етапу розвитку технології пошукової оптимізації – повністю виключити методи «сірої» та «чорної» оптимізації, які базуються на застарілих та грубих методах «розкрути».

Вивчення алгоритмів роботи сучасних фільтрів ПС дозволило виявити перспективні напрями оптимізації електронних ресурсів, що сприятиме добросовісній конкуренції в мережі, а результати видачі пошукових запитів стануть дійсно відображати авторитетні ЕР, які відслідковують зміни пов'язані з еволюцією технології ПО та демонструють сталий розвиток технологічності ресурсу.

Встановлено резерви удосконалення функціональності ПО.



1. PageRank [Електронний ресурс]: 2013 – Назва з екрана. URI: <http://ru.wikipedia.org/wiki/PageRank>. 2. Еволюція пошукових систем інтернет [Електронний ресурс]: 2013 – Назва з екрана. URI: [http://in-search-of-semantic.blogspot.com/2010/04/blog-post\\_8716.html](http://in-search-of-semantic.blogspot.com/2010/04/blog-post_8716.html). 3. Штучний інтелект у пошукових технологіях [Електронний ресурс]: 2013 – Назва з екрана. URI: <http://seolife.in.ua>. 4. Фільтри Google 2012: Panda и Penguin [Електронний ресурс]: 2013 – Назва з екрана. URI: <http://www.searchengines.ru>. 5. Амит Сингал розказав про майбутнє пошуку і про ролі пошукової оптимізації в просуванні ресурса [Електронний ресурс]: 2013 – Назва з екрана. URI: <http://www.searchengines.ru/news/archives>. 6. Основні тренди пошукової оптимізації на 2013-й рік [Електронний ресурс]: 2013 – Назва з екрана. URI: <http://seo.viento.ru/stati/311/137/>.