

СУЧАСНІ СИСТЕМИ ТА МЕТОДИ ПОШУКУ ІНФОРМАЦІЇ В МЕРЕЖІ ІНТЕРНЕТ

Вінницький національний технічний університет

Анотація

Розглянуто сучасні системи та методи пошуку інформації в мережі Інтернет, проаналізовано популярні пошукові системи як у світі, так і в Україні, описано переваги та недоліки даних систем.

Ключові слова: інформація, Інтернет, пошук, системи, методи, критерії пошуку, ринок, переваги, недоліки.

Abstract

The paper considers modern systems and methods of information retrieval on the Internet, analyzes popular search engines both in the world and in Ukraine, describes the advantages and disadvantages of these systems.

Keywords: information, Internet, search, systems, methods, search criteria, market, advantages, disadvantages.

Вступ

З кожним роком кількість виробленої людством інформації збільшується, це зумовлено зростанням кількості користувачів у мережі Інтернет. Сучасне життя неможливе без інтернету, ми використовуємо його для пошуку будь-якої інформації. Наявність джерел інформації в мережі Інтернет є настільки великою, що може задовольнити практично будь-який інформаційний запит. Та споживачі інтернет-ресурсів часто стикаються з проблемою пошуку інформації у мережі Інтернет, через що цей процес займає багато часу, тому актуальним питання залишається аналіз сучасних методів пошуку якісної і точної інформації у мережі Інтернет.

Результати дослідження

Пошук інформації одне із найбільш затребуваних завдань, яке доводиться вирішувати будь-якому користувачу Інтернету. Він заключається у пошуку інформації з певної теми. Інформація з Інтернет сайтів може бути найактуальнішою та найцікавішою, саме за це інтернет цінують мільйони користувачів. Але, не дивлячись на загальноживані правила пошуку інформації, майже для кожного користувача Інтернет цей процес є особистим творчим підходом, який призводить до розробки нових нестандартних методів добування інформації з різних джерел.

Основним завданням пошуку являється швидке та точне знаходження потрібної інформації. Незалежно від типу інформації в загальному вигляді підходи до її пошуку припускають реалізацію наступної методики:

1. Визначення теми пошуку інформації, формулювання тематичного запиту.
2. Розробка критеріїв пошуку.
3. Вибір способів пошуку.
4. Обробка інформації.
5. Використання знайденої інформації [1]

Існують три основні способи пошуку інформації:

1. Вказання адреси сторінки;
2. Пересилання по гіперпосиланню;
3. Звернення до пошукової системи.

Найшвидшим засобом пошуку є звичайно перший спосіб – вказівка адреси сторінки, проте для його використання варто знати точну адресу документу чи сайту.

Пересування по гіперпосиланню є не доволі зручним способом пошуку інформації оскільки шукає документи лиш близькі за змістом, отже потрібну інформацію особливо якщо її потрібно знайти у доволі великих об'ємах слід шукати достатньо довго.

Найбільш популярним способом є звернення до пошукової системи. Для отримання потрібної інформації споживачі мережі Інтернет найчастіше використовують саме пошукові системи, адже кожна пошукова система - це велика база ключових слів, пов'язаних з Web-сторінками, на яких вони зустрілися. Щоб знайти адресу сервера з інформацією що цікавить споживача треба ввести в поле пошукової системи ключове слово, кілька слів або фразу. Тим самим надіслати пошуковій системі запит. Результати пошуку видаються у вигляді списку адрес Web-сторінок, на яких зустрілися ці слова. Найбільш популярні пошукові системами в світі відображено на рис. 1 [2].

Search Engine Market Share Worldwide

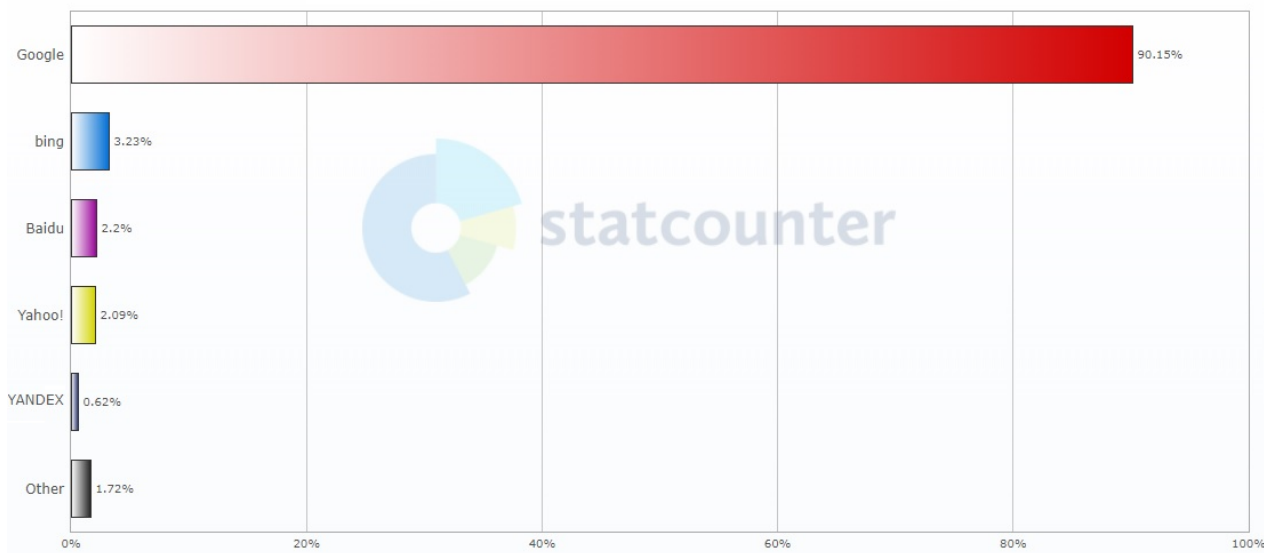


Рисунок 1. Частка ринку пошукових систем у всьому світі

Як бачимо, беззаперечним лідером серед пошукових систем у світі є Google - 90,15%. Також до найбільш популярних систем пошуку потрапили Bing - 3,23%, Baidu - 2,2% та Yahoo! - 2,09% однак існують і інші лідери в окремих країнах, наприклад в Китаї користувачі інтернет мережі найчастіше роблять пошук потрібної інформації через Baidu.com (74,51%). На другому місці в їхньому списку Shenma (11,25%), на третьому, зі своїми 7,3% стоїть пошуковик Haosou, на четвертому Sogou 3.76% [2]. Щодо України, то, як правило, користувачі інтернету України віддають перевагу Google. Статистика виглядає так: Google 87.56%, Яндекс 8,75% (не дивлячись на те, що згідно з рішенням РНБО від 28 квітня 2017 року доступ до «Яндекс» від 16 травня 2017 року було заблоковано в Україні [3] окремі несвідомі користувачі користуються даною системою через використання VPN), Bing 1,03% та Yahoo 0,74%. Найбільш розповсюдженими сервісами від українських розробників є: Sova.com.ua, Gala.Net, Ukr.net, Online.ua, MetaPin, Аванпорт, Мета, UA port, Атлас UA, I.ua, Meta.ua, Poshukach.com, Shukalka.com.ua, TopPING та ін. Пошукова система збирає інформацію з усіх сайтів інтернету, зберігає адреси сайтів у себе в сховище і у відповідь на запит запропонує посилання на найбільш підходящі сторінки.

У розпорядженні користувачів інтернету досить багато пошукових систем, які за охопленням сайтів можна розділити на дві групи:

- глобальні, які здійснюють пошук по всіх сайтах мережі (наприклад Google, Bing, і т.д.);
- локальні, вбудовані в один або кілька споріднених сайтів, які ведуть пошук тільки по ним.

Варто зазначити, що майже всі глобальні пошукові системи можуть використовуватися і в якості локальних, проте відносити їх до цієї групи неправомірно, оскільки пошук по окремому сайту для них є уточненням запиту.

Пошукові системи інтернету використовують: простий, розширений та контекстний пошук. Всі ці системи володіють певними перевагами, в числі яких:

- простота і зручність використання, що дозволяє непідготовленому користувачеві відразу приступити до пошуку інформації;

- ранжування або сортування результатів пошуку від найбільш релевантних до менш релевантних;
- відображення заголовка сторінки і невеликого екстракту (зазвичай 2-3 рядки) поряд з посиланням на сайт, що дозволяє скласти перше враження про релевантність сайту або виданого результату.

Разом з тим всі ці системи мають загальні недоліки:

- більшість цих систем комерційні, основна їх мета - приносити прибуток, тому вони часто і не завжди в тему розміщують рекламу, а також «просувають» сайт, тобто штучно підвищують його релевантність;
- вразливість: оскільки механізми індексації пошукових систем автоматичні, це дозволяє творцям сторінок вводити для підвищення релевантності ключові слова, які не мають відношення до утримання сторінки, але при цьому видно тільки пошуковим системам (тобто при завантаженні сторінки вони не відображаються);
- сортування тільки за релевантністю: до уваги береться дата створення сторінки, тому дуже часто на перших сторінках результатів пошуку йдуть посилання на релевантні, але застарілі матеріали;
- надлишок релевантних посилань, число яких іноді доходить до декількох мільйонів;
- відсутність уточнення запиту з потрібних областей;
- іноді відсутні релевантні посилання [4].

Кожна пошукова система намагається поліпшити результати пошуку і позбутися від недоліків або хоча б мінімізувати їх. Одні системи намагаються удосконалювати алгоритми пошуку, другі - пропонують користувачеві уточнити пошуковий запит.

Багато пошукових систем реалізували функцію «підказок», які при наборі тексту в пошуковому полі видають невеликий список найбільш частих запитів. В більшості глобальних пошукових систем надана можливість уточнити запит по типу інформації, наприклад: Google - вибрати зі списку: Усі, Зображення, Карти, Покупки, Новини, Відео, Книги, Авіарейси. А також окремою вкладкою для більш точного пошуку користувача є вибір мови та налаштування часових термінів, що особливо є зручним, коли треба отримати найновішу інформацію, наприклад за останній тиждень, день чи навіть годину. Також деякі пошукові системи для уточнення пошуку пропонують скористатися спеціальними операторами і пунктуацією.

При використанні знайденої інформації в Інтернеті, її слід проаналізувати, тому що бувають застарілі дані та можливі помилки у подачі інформації в мережу Інтернет.

Висновки

Кожна пошукова система, як глобальна, так і локальна, є унікальною і має свої алгоритми пошуку або їх поєднання. Підходи при формуванні пошукового запиту також відрізняються, кожен з них має як переваги, так і недоліки. Однак виявлені недоліки не повинні бути причиною відмови від їх подальшого використання. Спільне використання різних підходів пошуку представляється оптимальним для користувача, який сам зробить вибір залежно від розв'язуваних ним завдань.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Operating System | Thread. URL: <https://www.geeksforgeeks.org/operating-system-thread>
2. Юдін А. Пошукові системи світу. Статистика. URL: <https://marketer.ua/ua/search-engine-stat>
3. Указ президента України. URL: <https://www.president.gov.ua/documents/1332017-21850>
4. How the browsers compare. URL: <http://www.microsoft.com>

Дарморіс Богдан Геннадійович — студент гр. МВКД-196 факультет менеджменту та інформаційної безпеки, Вінницький національний технічний університет, e-mail: darmoris02@gmail.com

Нікіфорова Лілія Олександрівна - канд. екон. наук, доцент, доцент кафедри менеджменту та безпеки інформаційних систем, Вінницький національний технічний університет, e-mail: nikiforovalilia@gmail.com

Darmoris Bohdan Hennadiiovych - student gr. MPCA-19b, Faculty of Management and Information Security, Vinnytsia National Technical University.

Nikiforova Liliia Oleksandrivna - Ph.D. (Econ.), Associate Professor, Associate Professor of the Department of Management and Security of Information Systems, Vinnytsia National Technical University, Vinnytsia.