

## Поняття пошуку інформації в Інтернеті

Пошук інформації — завдання, яке найчастіше доводиться виконувати користувачу глобальної мережі. Але знайти у великій кількості сайтів і веб-сторінок потрібне джерело — дуже непросто. Для цього потрібно вміти використовувати різні способи пошуку інформації, правильно формулювати запити й критично оцінювати знайдену інформацію.

**Інтернет** (World Wide Web, WWW, що дослівно означає «всесвітня павутина») ще називають глобальною інформаційною системою.

### Основні способи пошуку у глобальній мережі такі:

**Вказання адреси веб-сторінки.** Це найшвидший спосіб пошуку. Його використовують, якщо точно відома адреса сторінки.

**Переміщення за допомогою гіперпосилань:** можна переходити зі сторінки на сторінку, шукаючи потрібну інформацію. Недолік очевидний: так можна довго і безрезультатно подорожувати мільйонами сторінок Інтернету.

**Використання спеціальних інструментів пошуку:** добірок посилань, пошукових каталогів та систем метапошуку. Ці інструменти мають спеціальні засоби організації пошуку, що забезпечує ефективний пошук потрібної інформації в Інтернеті.

Знайдену в результаті пошуку інформацію обов'язково потрібно проаналізувати. Публікації у глобальній мережі можуть містити застарілу інформацію або навіть помилки. Тому до використання знайденої інформації потрібно з'ясувати:

коли було створено сайт і як часто поновлюють повідомлення на ньому;

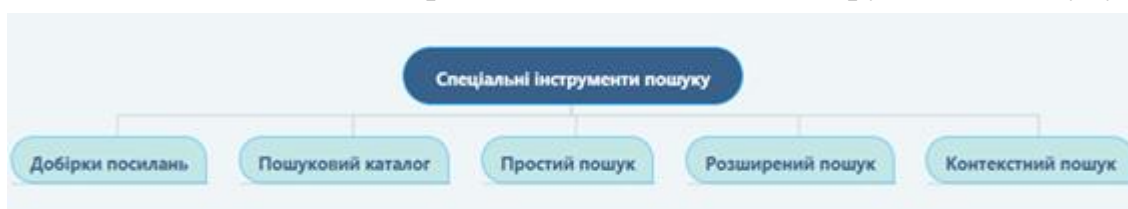
кому належить сайт — приватній особі чи організації;

чи є посилання на джерела інформації;

чи є зворотній зв'язок з адміністратором сайту або авторами статей.

При використанні матеріалів з Інтернету потрібно дотримуватися вимог авторського права: вказувати персональні дані автора, повну назву й адресу публікації.

Розглянемо детально використання спеціальних інструментів пошуку (Рис. 1).



## Рис. 1 Інструменти пошуку

**Добірки посилань** використовуються для пошуку інформації з конкретної теми. Вони містять списки посилань на сайти, які у деяких випадках згруповано в рубрики.

**Пошуковий каталог** — це структурований набір посилань на сайти з їх коротким описом. Його використовують для пошуку інформації з різноманітних тем. Його, як і добірку посилань, створюють «руками». Посилання на джерела інформації класифіковано у пошуковому каталозі за темами. Наприклад, освіта, наука, техніка, розваги, подорожі тощо. Кожна з цих тем може бути розбита на підтеми, що дозволяє звужувати область пошуку.

Пошукові системи Інтернету використовують такі алгоритми пошуку: простий пошук, розширений пошук та контекстний пошук.

**Простий пошук.** Під час цього пошуку у поле запиту вводять одне або декілька слів, які можуть характеризувати зміст документа. Під час введення одного слова машина видає, як правило, велику кількість посилань, з яких обрати потрібну інформацію буває досить складно. Тому простий пошук використовують для знаходження нескладних, однозначних питань чи теоретичних положень.

**Розширений пошук.** Такий пошук завжди включає запит із групи слів. Під час розширеного пошуку рекомендують зв'язувати ключові слова логічними операторами and (і), or (або), not (ні) тощо. Зазвичай записи ключових слів і логічних операторів у різних пошукових системах або однакові, або досить схожі. Тому, засвоївши один раз прийоми розширеного пошуку, можна ним користуватися де завгодно, переключивши машину в потрібний режим розширеного пошуку.

**Контекстний пошук.** Пошукові машини, що підтримують цей вид пошуку, видають посилання на інформацію, яка точно відповідає ключовим словам у пошуковому вікні. Для цього у більшості випадків ключову фразу потрібно взяти в лапки.

### **Пошукові системи та правила пошуку інформації**

Якщо адресу сторінки з цікавим для нас матеріалом невідомо і сторінки з відповідними посиланнями також немає, доводиться розшукувати матеріали у всьому Інтернеті. У цьому випадку застосовують пошукові системи — спеціальні web-вузли.

Існують два основних методи пошуку у глобальній мережі.

У першому випадку ви шукайте web-сторінки, які стосуються певної теми. Пошук здійснюють вибором тематичної категорії і поступовим її звуженням. Такі пошукові системи називають пошуковими каталогами. Вони зручні, коли ви знайомитеся з новою для себе темою чи хочете переглянути широковідомий «класичний» ресурс на цю тему.

Інший спосіб пошуку використовують, якщо тема носить вузький, специфічний характер або потрібні рідкісні, маловідомі ресурси. У цьому випадку потрібно знати, які ключові слова має містити документ з цікавої для вас теми.

Недоліком добірок посилань і пошукових каталогів є те, що вони пропонують невеликий за обсягом перелік посилань на сайти з потрібної теми, бо їх створюють люди. Для багатократного збільшення області пошуку цю роботу потрібно автоматизувати. Для цього були розроблені пошукові системи.

Пошук у таких системах здійснюють за допомогою спеціальних програм, які постійно переглядають веб-сайти Інтернету, автоматично створюють каталоги посилань та підтримують відповідність між створеними каталогами і наявними у мережі матеріалами. Робота користувача з пошуковою системою базується на формуванні запиту, згідно з яким відбувається добір потрібних документів. Запит формують за допомогою одного або кількох ключових слів. Результати пошуку надають у вигляді списку посилань і короткої анотації до них.

**Ключове слово** — це слово, яке найповніше відображає інформацію, яку потрібно знайти.

**Запит** — це набір ключових слів, за якими здійснюють пошук і відбір необхідних документів.

**Пошукова система** — це комплекс програм і потужних комп'ютерів, здатних приймати, аналізувати й обслуговувати запити користувачів з пошуку інформації у глобальній мережі.

Першою пошуковою системою стала Web Crawler, створена в 1994 році. На відміну від того, що були створено раніше, ця система дозволяла користувачам здійснювати пошук за будь-якими ключовими словами на будь-якій веб-сторінці. Це стало стандартом для всіх сучасних пошукових систем.

Адреси пошукових систем добре відомі всім, хто працює в Інтернеті. На сьогодні використовують такі пошукові системи: Google, Yahoo, Ask та інші. Розглянемо їх.

**Пошукова система Google, [google.com](http://google.com)**

Ця пошукова машина, основана на принципово новому алгоритмі пошуку, відрізняється гранично аскетичним інтерфейсом і прекрасними результатами пошуку з високим ступенем релевантності. На відміну від інших пошукових систем, у «першій десятці» результатів, виданих Google, зазвичай ви не зустрінете інформаційного сміття і випадкових сайтів: місце сайта в списку напряму зв'язано з кількістю посилань на нього з інших серверів аналогічної тематики.

Цікавою особливістю Google є наявність другої кнопки поряд з рядком пошуку. Перша кнопка запускає традиційний механізм. Друга кнопка одразу ж перекине користувача на сайт, який, на думку Google, максимально відповідає його запитам.

### **Пошукова система Yahoo, [yahoo.com](http://yahoo.com)**

Наявна база даних: новин, карт, рекламних і спортивних інформацій, бізнес, номери телефонів, персональні ваб-сторінки, електронні адреси (окрема база даних). Пошук можна обмежити певним проміжком часу. Діють оператори «і» та «або».

### **Пошукова система Ask (колишня назва Ask Jeeves), [ask.com](http://ask.com)**

Цю пошукову систему засновано 1996 року Гарреттом Гурнером і Девідом Вартемом у місті Берклі, штат Каліфорнія, США. З самого початку користувачі могли отримати відповіді на питання, сформульовані у теперешньому часі, а також використати традиційний пошук ключових слів. Сучасний Ask.com також підтримує це. Але долучено додаткову підтримку математики, словник і конверсію питання.

#### ***Пошук потрібної інформації здійснюють у 3 етапи***

Спочатку необхідно продумати тему запиту та проаналізувати, що ви хочете отримати в результаті пошуку. Правильно визначена тема пошуку допоможе вам при доборі ключових слів.

Другий етап — вибір пошукової системи — залежить переважно від уподобань користувача. Перед тим, як вибрати конкретну пошукову систему, потрібно ознайомитися з можливостями кожної з них. Вікна всіх пошукових систем виглядають по-різному, але містять однакові за призначенням об'єкти.

На третьому етапі треба визначити, які саме слова є характерними (ключовими) для шуканого документа. Правильно підібрані ключові слова і правильно сформульований запит значно розширяють область пошуку й підвищують його ефективність.

Після того, як визначено тему пошуку, вибрано конкретну систему пошуку, підібрано ключові слова та виконано опрацювання запиту, саме час перейти до

аналізу отриманого списку посилань. Якщо знайдені документи вас не задовольняють, варто звужити область пошуку. Для цього можна вказати додаткові ключові слова, терміни, визначення, які будуть включені до наступного пошукового запиту. У деяких пошукових системах для звуження області пошуку можна також встановити позначку прапорця «Шукати у знайденому» та ввести додаткові ключові слова.

### *Формування запитів і поліпшення результатів пошуку*

Дуже часто пошуковий сервер повертає меншій, ніж хочеться, перелік. Або навпаки, отриманий список посилань величезний, але вони зовсім не стосуються заданої теми. Перш ніж засмучуватися і/або переходити до іншого пошукового сервера, спробуйте поліпшити результати пошуку. Завантажте заголовну сторінку пошукового серверу у вікні перегляду. Уважно вивчіть ті сторінки (пошукового сервера), де опубліковано щодо організації пошуку й різного роду підказки. Хоча їх подано чужою мовою (у випадку міжнародних вузлів — англійською), але їх викладено дуже простими словами, і головне, супроводжено багатьма прикладами. На жаль, майже на кожному сервері застосовують свою систему правил (синтаксис вказівок).

### ***Незалежно від використаного сервера, керуйтеся такими настановами:***

Перед початком роботи уважно **ознайомтеся з довідковим розділом пошукової системи** щодо того, як краще скласти запит.

**Шукайте більше, ніж за одним словом**, щоб зменшити кількість посилань за рахунок підвищення відповідності їх темі пошуку. Якщо запит містить лише одне слово, ймовірність отримання необхідної інформації мала. У випадку, коли термін чи визначення використовують у різних предметних галузях, ви отримаєте багато зайвої інформації. Наприклад, при використанні у запиті ключового слова «архітектура» серед результатів пошуку ви отримаєте посилання сайти за темами: архітектура у містобудівництві, історія архітектури, архітектура обчислювальної техніки, архітектура комп'ютерних мереж тощо.

**Суворо дотримуйтеся правил орфографії.** Якщо в результаті пошуку за ключовими словами не знайдено жодного документу, можливо, ви допустили орфографічну помилку в написанні цих слів.

**Використовуйте малі літери.** Починаючи у запиті ключове слово з великої літери, серед результатів пошуку ви не знайдете слів, написаних з маленької літери, якщо це слово стоїть не першим у реченні. Наприклад, якщо ви використали ключове слово «презентація», то серед результатів пошуку ви отримаєте посилання на

сторінки зі словами: «презентація», «Презентація», «резентація» тощо. На запит «резентація» ви отримаєте посилання на сторінки, які містять лише слово «резентація», а решта сторінок з іншими варіантами написання цього слова буде проігноровано. Великі літери в запиті рекомендують використовувати лише в назвах і власних іменах. Наприклад, місто Київ, музей Лувр тощо.

**Використовуйте подвійні лапки (" ").** Якщо у запиті взяти ключові слова у подвійні лапки, то ви отримаєте посилання на сторінки, які містять даний рядок саме у такому вигляді, тобто слова будуть йти поруч одне за одним. Якщо ключові слова вводять без лапок, то отримують посилання на сторінки, які містять ключові слова, можливо, і окремо у будь-якому місці документа. Наприклад, якщо запит сформульовано таким чином: "Слово про похід Ігорів", то серед результатів пошуку ви отримаєте посилання на сторінки, які містять інформацію про давньоруську пам'ятку, перлину українського ліро-епосу — поему «Слово про похід Ігорів». Якщо ж у запиті ви використали ключові слова: Слово про похід Ігорів (без подвійних лапок), то ви отримаєте посилання на сторінки, які будуть містити інформацію про туристичні або військові походи, про людей з іменем Ігор тощо.

**Використовуйте знаки «+» та «-».** У запиті знак «+» перед ключовим словом означає обов'язкову його присутність у шуканому документі. І навпаки, щоб вилучити документи, що містять дане слово, потрібно поставити перед ним знак «-». Наприклад, якщо потрібно знайти інформацію про довільні паралелограми, прямокутники та ромби, але виключити інформацію про квадрати, то запит варто сформулювати так: паралелограм-квадрат.

**Користуйтеся чіткими і ясними ключовими словами,** що найбільше повно характеризують предмет вашого пошуку. Провідні пошукові сервіси дозволяють поставити запитання природною мовою. Наприклад «Tell me about the GreenPeace in 1998». Але такою можливістю варто користуватися обережно. У такому випадку пошуковий сервер відкидає всі незначні на його думку слова, крім тих, які тлумачить як ключові слова. Для поданого прикладу «GreenPeace» і «1998». Але існує можливість того, що буде відкинуто й те слово, важливе для результатів пошуку.

**Використовуйте синоніми.** Це стосується перш за все нових областей знання з неусталеною термінологією. У тому числі — інформаційних технологій.

**Використовуйте розширений пошук (Advanced Search).** На багатьох серверах є режим «розширеного» або «складного» пошуку, де діють спеціальні синтаксичні правила для застосування булевих операцій над ключовими словами, передбачено розрізнення великих і малих літер, пропонують вибір способів

впорядкування знайдених документів і т.і. Як уже згадано вище, варто докладно вивчити можливості обраного вами пошукового сервера, правила уведення й формулювання термінів. Без цього навряд чи вдасться досягти значної повноти пошуку.

**Локалізуйте пошук.** У багатьох випадках, коли предмет пошуку можна локалізувати географічно, корисно виконати пошук на серверах місцевих організацій, наприклад, на серверах університетів. Там перебуває величезна кількість посилань на місцеву пресу, компанії й громадські організації, що працюють у цьому регіоні й т.і. З різних причин пошукові сервери не індексують значну частину таких документів, тому їх можна знайти лише в результаті кропіткого локального пошуку інформації.

### **Метапошукова система**

**Метапошукова система** (метакраулер, мультипоточна система) – це пошуковий інструмент, який надсилає ваш запит одночасно декільком пошуковим системам, каталогам і, інколи, в так звану невидиму (приховану) павутину – зібрання онлайнової інформації, не проіндексованої традиційними пошуковими системами. Зібравши результати, метапошукова система знищує посилання-дублікати і, у відповідності з власним алгоритмом, об'єднує / рангує результати в загальному списку. На відміну від окремих пошукових систем і директорій, метапошукові системи не мають власних баз даних і не реєструють URL-адреси сайтів.

Метапошуком варто користуватися не у всіх випадках. Якщо документів за запитом багато, то метапошук не потрібен і, можливо, навіть шкідливий, бо змішує різні логіки ранжування. Але якщо документів за запитом мало, то метапошук може бути корисним завдяки тому, що об'єднує велику кількість пошуковиків.

Аргументи за метапошук економлять час, бо позбавляють необхідності вводити запит в кожній окремій пошуковій системі; результати, переважно, більш релевантні; можна використати для визначення наявності певного веб-сайту в головних пошукових системах, визначення рейтингів сайту за популярністю посилань (важливо для розробників сайтів).

Аргументи проти метапошуку полягають в тому, що деякі пошукові сайти і каталоги не підтримують додаткові технології пошуку (такі як лапки, в які беруть фрази, чи булеві оператори). Тому при застосуванні таких технологій у результатах метапошуку не буде відображено результати з таких пошукових систем, або ці результати будуть невідповідними.

Розглянемо типи метапошукових систем (Рис. 2).

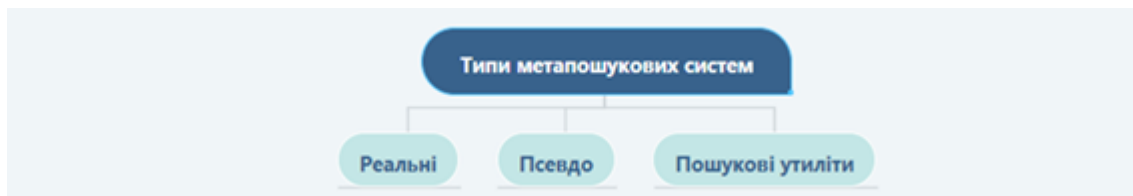


Рис. 2. Типи метапошукових систем

**«реальні»** метапошукові системи, які об'єднують/ранжують результати на одній сторінці;

**«псевдо»** метапошукові системи першого типу, які групують результати за пошуковими системами на одній довгій сторінці;

**«псевдо»** метапошукові системи другого типу, які відкривають для кожної пошукової системи, яку використовують, нове вікно;

**пошукові утиліти** – програмні пошукові засоби.

До метапошукових систем відносяться: [Yippy](#), [Dogpile](#), [Mamma](#), [Search](#) та інші.