

МІНІСТЕРСТВО ОСВІТИ І НАУКИ  
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ “ЛЬВІВСЬКА ПОЛІТЕХНІКА”

**ТЕХНОЛОГІЇ ІНФОРМАЦІЙНОГО ПОШУКУ**

**МЕТОДИЧНІ ВКАЗІВКИ**

до виконання лабораторної роботи № 4  
Тематичний пошук в мережі Internet

**Мета роботи:** удосконалити навички тематичного пошуку, навчитися використовувати можливості розширеного тематичного пошуку для підвищення релевантності пошуку за встановленим завданням.

## ТЕОРЕТИЧНІ ВІДОМОСТІ

### 1. Тематичний пошук

Дослідження можливостей пошуку в Google за тематикою дає великий досвід роботи з пошуковою системою. Крім загального пошуку текстової інформації, часто необхідно знайти зображення, музичні та відеоконтенти, вузькоспеціалізовану інформацію, відомості про погоду або якийсь товар. У цьому випадку необхідно враховувати особливості шуканих даних.

Для цього Google пропонує широкий спектр служб і програм, представлених на сторінці [www.google.com](http://www.google.com) (Рис. 1). Переглянути повніший перелік можна, клацнувши на посиланні More (Більше).

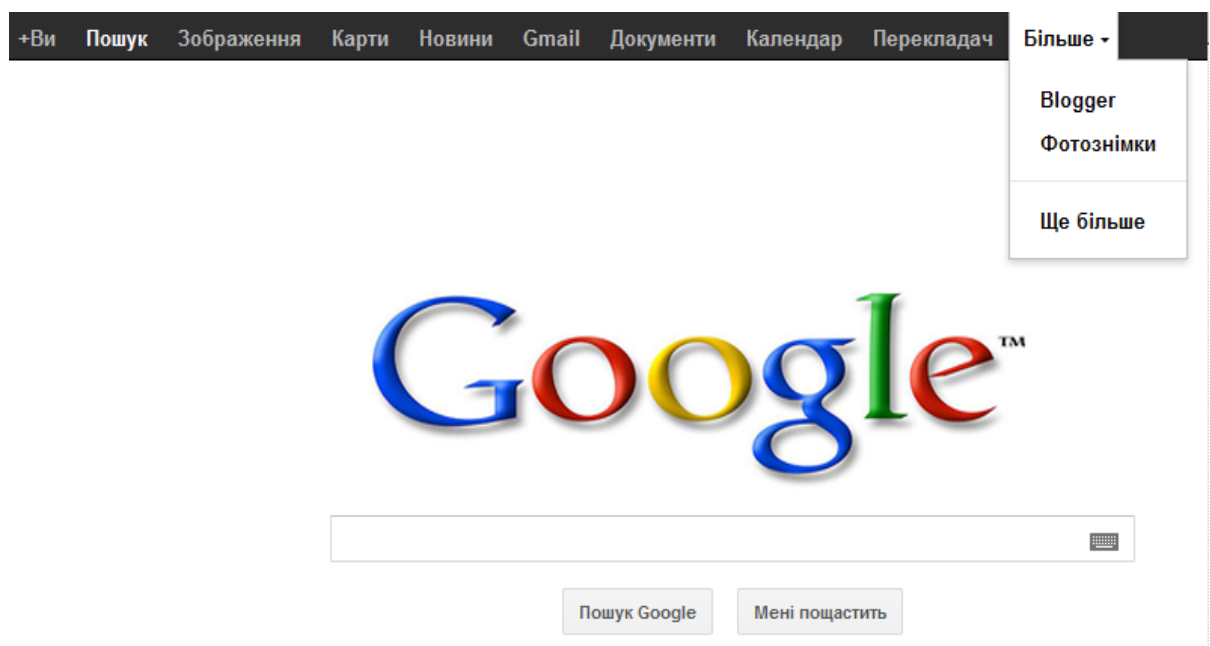


Рис. 1. Ресурси для тематичного пошуку

#### 1.1. Пошук зображень

Для пошуку зображень вибираємо відповідний пункт меню **Google – Зображення** (Рис. 2).

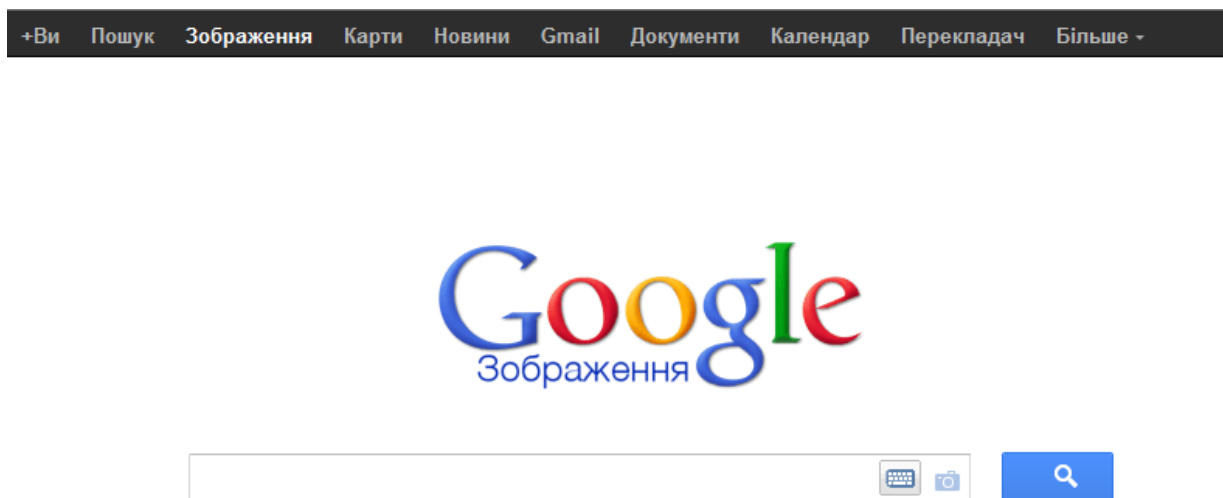


Рис. 2. Вікно пошуку зображень за допомогою Google

Щоб визначити, які саме зображення найбільше відповідають запиту користувача, Google аналізує текст на сторінках, де розміщені ці зображення, їх назви, замісник текст, і деякі інші ознаки

### **1.1.1 Простий пошук**

Щоб знайти, наприклад, зображення мікрофону, в стрічці пошуку потрібно ввести слово “*Мікрофон*” і клацнути на кнопці Пошук зображень. Пошукова система обробить запит і видасть сторінку з результатом пошуку (Рис. 3).

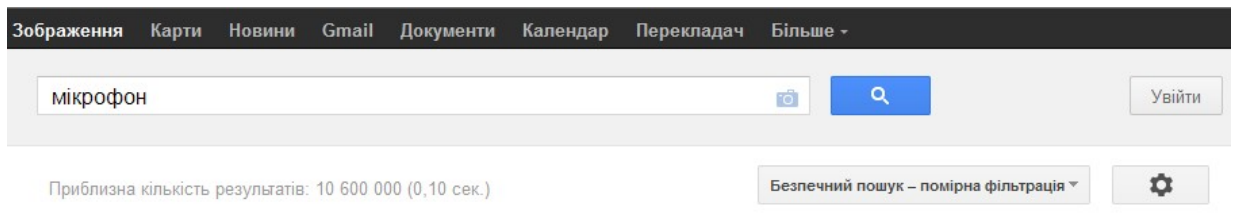


Рис. 3. Результат пошуку за запитом “мікрофон”

На цій сторінці кожна картинка має наступні параметри (Рис. 4):

- зменшене (якщо вона досить велика) зображення,
- ім'я зображення,
- посилання на веб-сторінку, де знаходиться оригінальне зображення (точніше, посилання на сторінку, збережену в кеші),
- розмір в вигляді «ширина \* висота» в пікселях,
- опис зображення,
- можливість пошуку схожих зображень,
- можливість пошуку ідентичного зображення інших розмірів.



[shure\\_sm58lce.jpg](#)

[musicalka.com.ua](#)

550 × 383 - Микрофон Shure SM58SE (min. 2 pcs.) Увеличить

[Схожі](#) - [Інші розміри](#)

Рис. 4. Параметри знайденого зображення

Іноді, при безлічі релевантних зображень на одному мережному ресурсі під зображенням може бути посилання “Додаткові результати з імя\_сайту”, що дозволяє подивитися ці зображення.

Зліва на сторінці результатів пошуку зображень розташований список, що дозволяє відфільтрувати зображення (Рис. 5).

### Усі результати

За темою

### Будь-якого розміру

Великі

Середні

Піктограми

Більші ніж...

Лише розміром...

### Будь-якого кольору

Різнокольорові

Чорно-білі



### Будь-якого типу

Обличчя

Фотографія

Ілюстрація

Графічний об'єкт

### Стандартний вигляд


Показати розміри

### Будь-коли

За минулий тиждень

Рис. 5. Можливості фільтрування зображень

#### 1.1.2 Розширений пошук

Розширений пошук зображень запускається зі сторінки пошуку картинок клацанням на посиланні Параметри під іконкою . Вікно розширеного пошуку зображень схоже на звичайний розширений пошук, проте відображає тематичну специфіку (рис. 6).

На цій сторінці можна підвищити релевантність пошуку зображень, задаючи різні параметри.

В області **“Знайти зображення, що містять...”** розташовані наступні поля:

- усі ці слова - враховуються всі слова цього поля;
- точне слово або фразу - розглядається точнf відповідність введеної фрази;
- з будь-яке з цих слів - достатньо присутності хоча б одного з слів;
- жодне з цих слів - дані слова мають бути відсутні.

Область **“Потім виберіть категорію, щоб звузити результати...”** містить такі основні пункти:

- розмір зображення (великі, середні),
- співвідношення розмірів (високі, квадратні тощо),
- кольори зображення (різнокольорові, чорно-білі тощо),
- тип зображення (обличчя, ілюстрація тощо),
- тип файлу (JPG, GIF, PNG тощо),
- права на використання.

### **1.1.3 Пошук за зображенням**

Google дозволяє здійснювати пошук вмісту Інтернету абсолютно повному, здійснюючи пошук за зображенням, а не за допомогою тексту.

Пошук за зображенням можна здійснити двома способами (Рис. 7):

- через пункт "Завантажити зображення" – завантажити зображення, за допомогою якого потрібно здійснити пошук, з особистого комп'ютера;
- через пункт "Вставте URL-адресу зображення" – здійснювати пошук за зображенням, яке знаходиться в мережі інтернет.

Як копіювати URL-адресу зображення:

- на зображенні, адресу якого Ви хочете дізнатися, клацніть зображення правою кнопкою миші та скопіюйте його URL-адресу (рис. 8).
- Вставте скопійовану адресу у пункт "Вставте URL-адресу зображення".

## Розширений пошук зображень

Знайти зображення, що містять...

усі ці слова:

точне слово або фразу:

будь-яке з цих слів:

жодне з цих слів:

Потім виберіть категорію, щоб звузити результати...

розмір зображення:  Знаходьте зображення будь-якого потрібного розміру.

співвідношення розмірів:  Укажіть форму зображення.

кольори зображення:  будь-якого кольору  різнокольорові  чорно-білі  цього кольору: Знаходьте зображення потрібних кольорів.

тип зображення:  Обмежте тип зображень, які шукаєте.

регіон:  Знаходьте зображення, опубліковані в певному регіоні.

сайт чи домен:  Шукайте один сайт (як-от `example.org`) або обмежте результати доменом, як-от `.edu`, `.org` або `.gov`.

**Безпечний пошук**

без фільтрації **помірний** строгий  
Налаштуйте **Безпечний пошук**, щоб відфільтрувати вибрану кількість вмісту відвертого сексуального характеру.

тип файлу:  Знаходьте зображення потрібного вам формату.

права на використання:  Знаходьте зображення, які можна використовувати.

Розширений пошук

Рис. 6. Вікно розширеного пошуку зображень



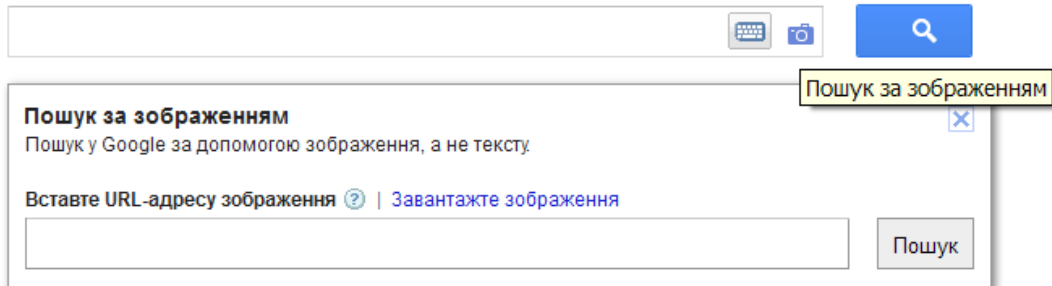


Рис. 7. Пошук за зображенням

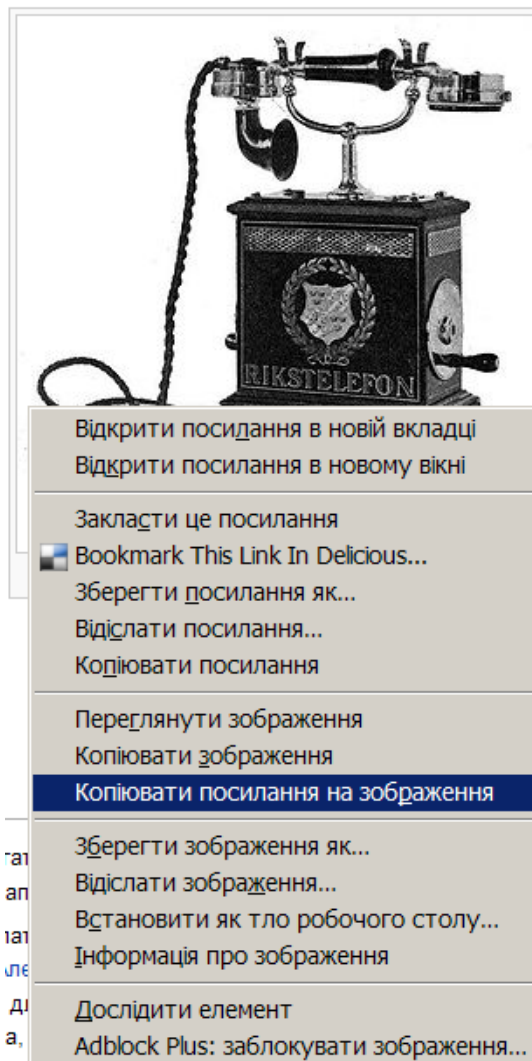


Рис. 8. Копіювання URL-адреси зображення

Результати пошуку за скопійованою URL-адресою зображення показані на рис. 9.

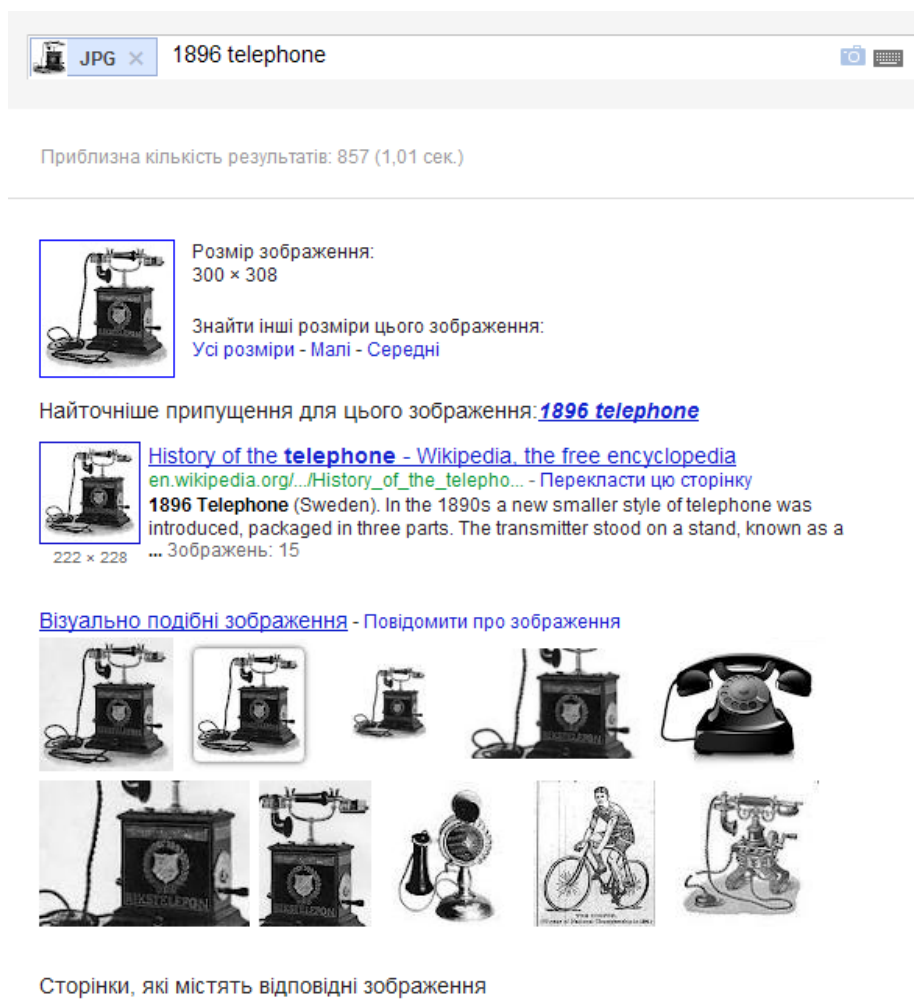


Рис. 9. Результати пошуку за URL-адресою зображення

## 1.2. Пошук відео матеріалів

Зі збільшенням кількості відеокамер, особливо у телефонах, в Інтернеті зростає кількість відео файлів.

Простий пошук відео файлів доступний зі сторінки пошуку Зображення або за адресою <http://www.google.com/videohp> (Рис. 10).



Рис. 10. Сторінка Google для пошуку відеофайлів

Підвищити релевантність результатів можна за допомогою розширеного пошуку, для чого слід натиснути на посилання Advanced Video Search (рис. 11) за адресою [https://www.google.com/advanced\\_video\\_search?hl=uk](https://www.google.com/advanced_video_search?hl=uk)

<b>Find Results</b>	with <b>all of the words</b>	<input type="text"/>	<input type="button" value="Search Videos"/>
	with the <b>exact phrase</b>	<input type="text"/>	
	with <b>at least one</b> of the words	<input type="text"/>	
	<b>without</b> the words	<input type="text"/>	
<b>Language</b>	Return videos in	<input type="text" value="Any Language"/>	
<b>Duration</b>	Return videos within the duration	<input type="text" value="All durations"/>	
<b>Domain</b>	<input type="text" value="Only"/> return videos from the site or domain	<input type="text" value="e.g. youtube.com"/>	
<b>Subtitles</b>	Search only closed captioned videos	<input type="checkbox"/>	
<b>Sort results by</b>		<input type="text" value="Relevance"/>	
<b>Results per page</b>		<input type="text" value="10"/>	

Рис. 11. Сторінка розширеного пошуку відеофайлів

Сторінка розширеного пошуку має звичний вигляд. В області Find results розташовані знайомі користувачеві поля:

- with all of the words - враховуються всі слова в цьому полі;
- with the exact phrase - розглядається точну відповідність введеної фрази;

- with at least one of the words - достатньо присутності хоча б одного з слів;
- without the words - дані слова мають бути відсутні.

У списку Return videos in можна вибрати мову, а за допомогою списку Return videos within the duration - задати тривалість відеофайлу. У списку Domain можна відібрати відео файли з певного ресурсу. У списку Subtitles можна вказати необхідність субтитрів у супроводі.

За допомогою розкривних списків Sort results by і Results per page можна задати відповідно метод сортування результатів і їх кількість, що виводиться на сторінці.

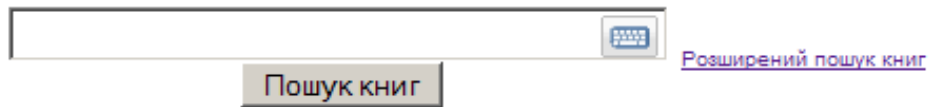
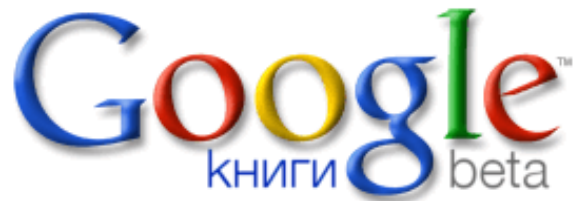
### **1.3. Пошук книг на наукових публікаціях**

Пошукова система Google дозволяє здійснювати пошук книг та наукових публікацій. Користувачі книжкового пошукача Google Книги можуть не тільки читати книги в Мережі, але і роздруковувати їх. Однак така можливість доступна лише щодо текстів, у яких закінчився термін дії авторських прав.

Робота над створенням книжкового пошукача Google, який повинен стати найбільшим бібліотечним каталогом у світі, почалася в кінці 2004 року. Зараз компанія перевела у цифровий формат вже десятки тисяч книг. Партнерами Google стали Мічиганський, Гарвардський, Стенфордський і Оксфордський університети та Нью-Йоркська громадська бібліотека. Очевидно, зо з цих бібліотек буде доступна лише невелика частина книг.

#### **1.3.1 Електронні книги**

Сервіс Google Книги, що дозволяє шукати книги, доступний користувачам за адресою <http://books.google.com.ua> (рис. 12)



**Шукайте повні версії текстів і відкривайте для себе нові книги.**


Рис. 12. Вікно Google Книги


На сторінці результатів відображаються наступні параметри (рис. 13):


- назва книги;
- зображення,
- посилання на детальну інформацію та перегляд,
- автор
- рік видання;
- кількість сторінок,
- вид обмеження на перегляд

технології

Приблизна кількість результатів: 1 080 000 (0,37 сек.)

[Банковское дело: Операции, технологии, управление:](#)  

[books.google.com.ua](#)  
 Александр Турбанов, Александр Тютюнник - 2010 - 681 стор. - [Попередній перегляд](#)  
 Обновленное издание, которое вы держите в руках, рассматривает вопросы управления ресурсами, риск- менеджмента, учета и отчетности, информационных **технологий** (ИТ) комплексно, интегрируя современные тенденции с традиционными элементами ...

[Rfid-технологии на службе вашего бизнеса](#)  

[books.google.com.ua](#)  
 Бхуптани Маниш Морадпур Шахрам, Маниш Бхуптани, Шахрам Морадпур - [Попередній перегляд](#)  
 Предисловие к российскому изданию История внедрения **технологии** радиочастотной идентификации (RFID) в России насчитывает уже без малого десять лет. Бесконтактные карты на метрополитене или на проходных многих офисов и предприятий стали ...

[Технологии политических выборов:](#)  

[books.google.com.ua](#)  
 Мирослав Евгеньевич Кошелюк - 2004 - 238 стор. - [Попередній перегляд](#)  
 ... нового института демократических выборов в России. Необходимость данного издания диктуется, на наш взгляд, двумя причинами. Во-первых, понятие «избирательные **технологии**» сегодня стало буквально жупелом для широкой общественности.

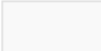
[Сумма технологий:](#)  

[books.google.com.ua](#)  
 Станислав Лем - 2004 - 668 стор. - [Попередній перегляд недоступний](#)

Рис. 13. Результаты поиска за ключевым словом “технології”

Клацання на обраній книзі відкриває для перегляду фрагмент де. як в бібліотечному каталозі міститься інформація про видання і кілька уривків з твору - пропозиції, що містять слово запиту Крім того, якщо на книгу не поширюється авторське право, то можна вибрати перегляд прикладів сторінок, якщо видавець або автор дав на це згоду, або задати повний перегляд видання Тут же знаходяться посилання на інтернет-магазини, в яких можна придбати даний продукт.

Однак, для ефективнішого пошуку використовується можливість розширеного пошуку книг (рис. 14).

Знайти результати	з усіма словами	<input type="text"/>	10 результатів
	що містять точну фразу	<input type="text"/>	<input type="button" value="Пошук Google"/>
	хоча б з одним зі слів	<input type="text"/>	
	що не містять слів	<input type="text"/>	
Пошук:	<input checked="" type="radio"/> Усі книги <input type="radio"/> Обмежений попередній перегляд і повний перегляд <input type="radio"/> Тільки повний перегляд <input type="radio"/> Лише електронні книги Google		
Вміст:	<input checked="" type="radio"/> Увесь вміст <input type="radio"/> Книги <input type="radio"/> Журнали		
Мова	Показувати сторінки такою мовою:	<input type="text" value="усі мови"/>	
Назва	Відобразити книги з назвою	<input type="text"/> наприклад, Books and Culture	
Автор	Відобразити книги, написані	<input type="text"/> наприклад, Hamilton Mabie або "Hamilton Wright Mabie"	
Видавець	Відобразити книги, видані	<input type="text"/> наприклад, O'Reilly	
Дата публікації	<input checked="" type="radio"/> Показати вміст, надрукований будь-коли	<input type="text"/> <input type="text"/> і <input type="text"/> <input type="text"/>	
	<input type="radio"/> Показати вміст, надрукований у період між	напр. 1999 і 2000 або січ. 1999 і груд. 2000	
ISBN	Відобразити книги з номерами ISBN	<input type="text"/> наприклад, 0060930314	
ISSN	Повернути журнали з ISSN	<input type="text"/> наприклад, 0161-7370	

Рис. 14. Вікно розширеного пошуку книг

Окрім вже звичних можливостей формулювання уточненого запиту, цей сервіс передбачає ще низку можливостей:

- вибір обмеження перегляду книг;
- вказання слів у назві книги;
- вказання автора книги;
- вказання видавництва;
- вказати дату публікації;

- пошук за ISBN та ISSN.

### 1.3.2. Пошук наукової інформації

Сервіс Google Академія (Scholar) надає простий спосіб пошуку академічної літератури. За допомогою цієї функції можна шукати інформацію з різних дисциплін з різних джерел: статті, книги, резюме і т. д.

Особливості пошуку за допомогою Google Академія:

- забезпечується пошук з різноманітних джерел;
- можливий пошук статей, резюме і цитат;
- визначається місцезнаходження книги, статті в бібліотеках, або Інтернеті;
- забезпечується інформація про статті з будь науково-дослідної області.

При рангуванні результатів Google Scholar враховує цитованість документа в іншій академічній літературі. Сервіс розташований за адресою <http://scholar.google.com.ua/> (рис. 15).

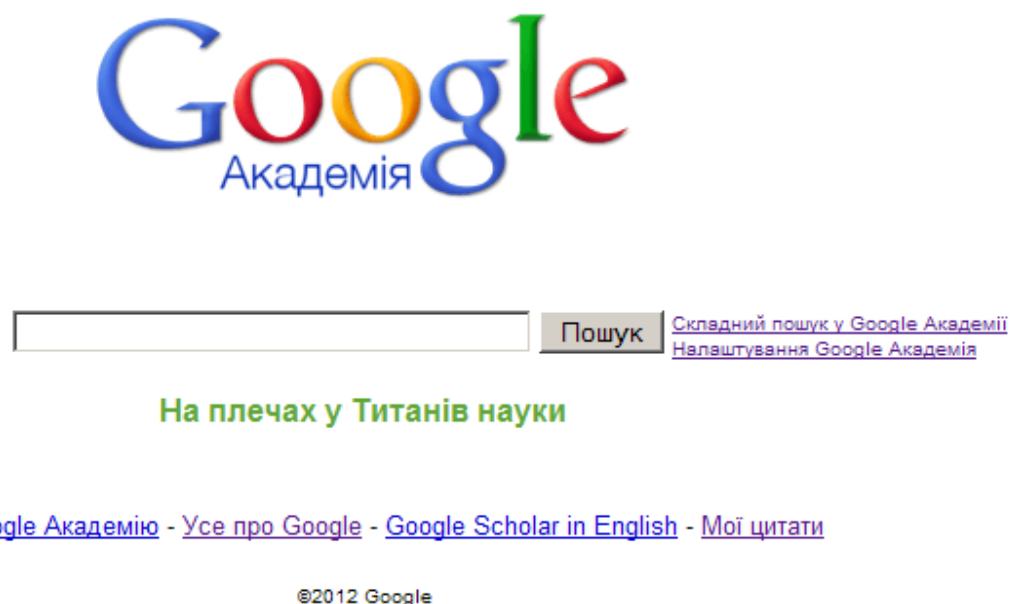


Рис. 15. Вікно сервісу Google Академія

Результат пошуку містить наступні посилання (рис. 16):

- заголовок - посилання на резюме статті або на саму статтю;
- автор та джерело статті;
- цитовано в - ідентифікує інші статті, що цитують поточну;



- пов'язані статті - шукає схожі документи;
- кількість версій - визначає кількість версій даної статті.

Google Академія   [Складний пошук у Google Академії](#) [Мої цитати](#)

Академія    Результати 1 – 10 з приблизно 615 000

Порада. Шукайте результати вибраною мовою: **українська**. Мову пошуку можна задати в меню [Налаштування Google Академія](#).

**Opinion mining and sentiment analysis** [\[PDF\] з домену iitb.ac.in](#)  
[B Pang...](#) - Foundations and Trends in Information Retrieval, 2008 - dl.acm.org  
 Abstract An important part of our information-gathering behavior has always been to find out what other people think. With the growing availability and popularity of **opinion**-rich resources such as online review sites and personal blogs, new opportunities and ...  
 Цитовано в 870 джерелах - Пов'язані статті - Кількість версій: 30

**Sentiwordnet: A publicly available lexical resource for opinion mining** [\[PDF\] з домену uib.no](#)  
[A Esuli...](#) - Proceedings of LREC, 2006 - gandalf.aksis.uib.no  
 Abstract **Opinion mining** (OM) is a recent subdiscipline at the crossroads of information retrieval and computational linguistics which is concerned not with the topic a document is about, but with the **opinion** it expresses. OM has a rich set of applications, ranging from ...  
 Цитовано в 524 джерелах - Пов'язані статті - Показати у форматі HTML - Кількість версій: 23

**A holistic lexicon-based approach to opinion mining** [\[PDF\] з домену psu.edu](#)  
[X Ding, B Liu...](#) - ... conference on Web search and web data mining, 2008 - dl.acm.org  
 Abstract One of the important types of information on the Web is the opinions expressed in the user generated content, eg, customer reviews of products, forum posts, and blogs. In this paper, we focus on customer reviews of products. In particular, we study the problem of ...  
 Цитовано в 154 джерелах - Пов'язані статті - Кількість версій: 13

**Determining term subjectivity and term orientation for opinion mining** [\[PDF\] з домену upenn.edu](#)  
[A Esuli...](#) - Proceedings the 11th Meeting of the ..., 2006 - acl.ldc.upenn.edu  
 Abstract **Opinion mining** is a recent subdiscipline of computational linguistics which is concerned not with the topic a document is about, but with the **opinion** it expresses. To aid the extraction of opinions from text, recent work has tackled the issue of determining the ...  
 Цитовано в 136 джерелах - Пов'язані статті - Показати у форматі HTML - Кількість версій: 22

Рис. 16. Результати пошуку у Google Академія

Цей сервіс також пропонує розширені можливості пошуку (рис. 17):

Google Академія **Складний пошук у Google Академії** [Поради щодо складного пошуку](#) | [Про Google Академію](#)

Знайти статті з усіма словами  Результатів на сторінку:

що містять **точну фразу**

з хоча б одним зі слів

без слів

де зустрічаються пошукові слова

**Автор** Видати статті таких авторів   
наприклад, "ЛВ Костенко" або Патон

**Публікація** Видати статті, опубліковані в   
наприклад, J Biol Chem або Nature

**Дата** Видати статті, опубліковані в період  —   
наприклад, 1996

Рис. 17. Налаштування розширеного пошуку в Google Академія

## ХІД РОБОТИ

1. Ознайомтеся з теоретичними відомостями.

### Вільне завдання для пошуку

2. Надайте власні приклади ВДАЛОГО! використання розширеного пошуку зображень, відео, книг, наукових матеріалів. Розробіть запити для різних варіантів використання розширених можливостей. Додаючи елементи запиту у різні поля, проаналізуйте, порівняйте і поясніть результати. Надайте відповідні копії з екрану.

**Підберіть таку предметну область і такі запити, щоб використання розширеного пошуку було виправданим, а результати очевидними!!**

### Завдання для пошуку щодо країни

1. Згідно з номером у підгрупі, виконайте кожне завдання для кожної країни. Поясніть, як Ви використовували розширений пошук. Для кожного з пунктів дайте окреме пояснення. При необхідності, надайте копії з екрану.

Порядковий номер студента у списку підгрупи	Назва країни
1.	Хорватія
2.	Франція
3.	Італія
4.	Польща
5.	Болгарія
6.	Чехія
7.	Німеччина
8.	Англія
9.	Іспанія
10.	Голландія
11.	Бельгія
12.	Австрія
13.	Чорногорія
14.	Нідерланди
15.	Швейцарія
16.	Греція
17.	Португалія

2. Знайдіть зображення прапора та герба відповідної держави.
3. Знайдіть зображення герба та прапора у різних розмірах. Виберіть найменше та найбільше зображення. Вкажіть розміри. Надайте зображення та відповідні копії з екрану.

4. Завантажте зображення прапора на персональний комп'ютер. Переіменуйте його, наприклад, 1.jpg.
5. Завантажте його у стрічку пошуку за зображенням. Проаналізуйте отримані результати.
6. Знайдіть зображення двох-трьох найвідоміших туристичних об'єктів країни (пам'ятник, будівля тощо). Підшукайте зображення різних розмірів. Виконайте для цих зображень пункти 3-4 цього завдання. Проаналізуйте результати.
7. Знайдіть чорно-білі зображення столиці країни за 50-60pp. 19 століття.
8. Знайдіть зображення президента країни. Знайдіть портрет дружини (чоловіка).
9. Знайдіть зображення, які стосуються країни в українському розділі Вікіпедія <http://uk.wikipedia.org>.
10. Знайдіть відео про столицю країни українською мовою тривалістю більше ніж 2 хвилини.

#### **Завдання для пошуку щодо наук та їх розділів**

1. Згідно з номером у підгрупі, виконайте кожне завдання для кожної науки або її розділу. Поясніть, як Ви використовували розширений пошук. Для кожного з пунктів дайте окреме пояснення. При необхідності, надайте копії з екрану.

Порядковий номер студента у списку підгрупи	Назва науки
1.	Менджмент
2.	Математичне моделювання
3.	Ділова комунікація
4.	Фізика
5.	Комерційна діяльність
6.	Трудове право
7.	Макроекономіка
8.	Логістика
9.	Логопедія
10.	Психологія
11.	Екологія
12.	Політологія
13.	Криміналістика
14.	Фінанси і кредит
15.	Мікробіологія
16.	Культурологія
17.	Педагогіка

2. Для кожної наукової сфери знайдіть підручники. Знайдіть окремо україномовні та російськомовні підручники.

3. Знайдіть провідних вчених або авторів публікацій у кожній з галузей. Знайдіть наукові публікації цих людей. Надайте відповідні зображення і опис.

## **СТРУКТУРА ТА ОФОРМЛЕННЯ ЗВІТУ ДО ЛАБОРАТОРНИХ РОБІТ**

Звіт з лабораторної роботи оформлюється на аркушах формату А4, які заповнюються з однієї сторони. Текст повинен бути рукописним або друкованим на принтері. Використання кольорових чорнил дозволяється лише для ілюстративних матеріалів.

Звіт до лабораторної роботи формується відповідно до змісту і повинен містити такі розділи:

- титульна сторінка;
- мета роботи;
- короткі теоретичні відомості;
- опис усіх етапів виконання роботи;
- опис отриманих результатів;
- висновки за результатами роботи.

Звіт зшивається з лівої сторони листів формату А4. Титульна сторінка звіту обов'язково друкується на принтері.