
4. Search Engine Analytics

— --- Sakshi Surve —

Contents

- Types of Search Engines
- Search Engine Analytics
- Search Engine Analytics Tools

Search Engines

- A **search engine** is a software program that helps people find the information they are looking for online using keywords or phrases.
- Search engines are tools that help users find information on the internet. They do this by searching for information that corresponds to a user's request (keywords).
- There are billions of websites on the internet, so search engines are crucial for helping users find the right information quickly.



Google Search



Microsoft Bing



Yahoo! Search



Ask.com



DuckDuckGo



Yandex



AOL Search



Ecosia



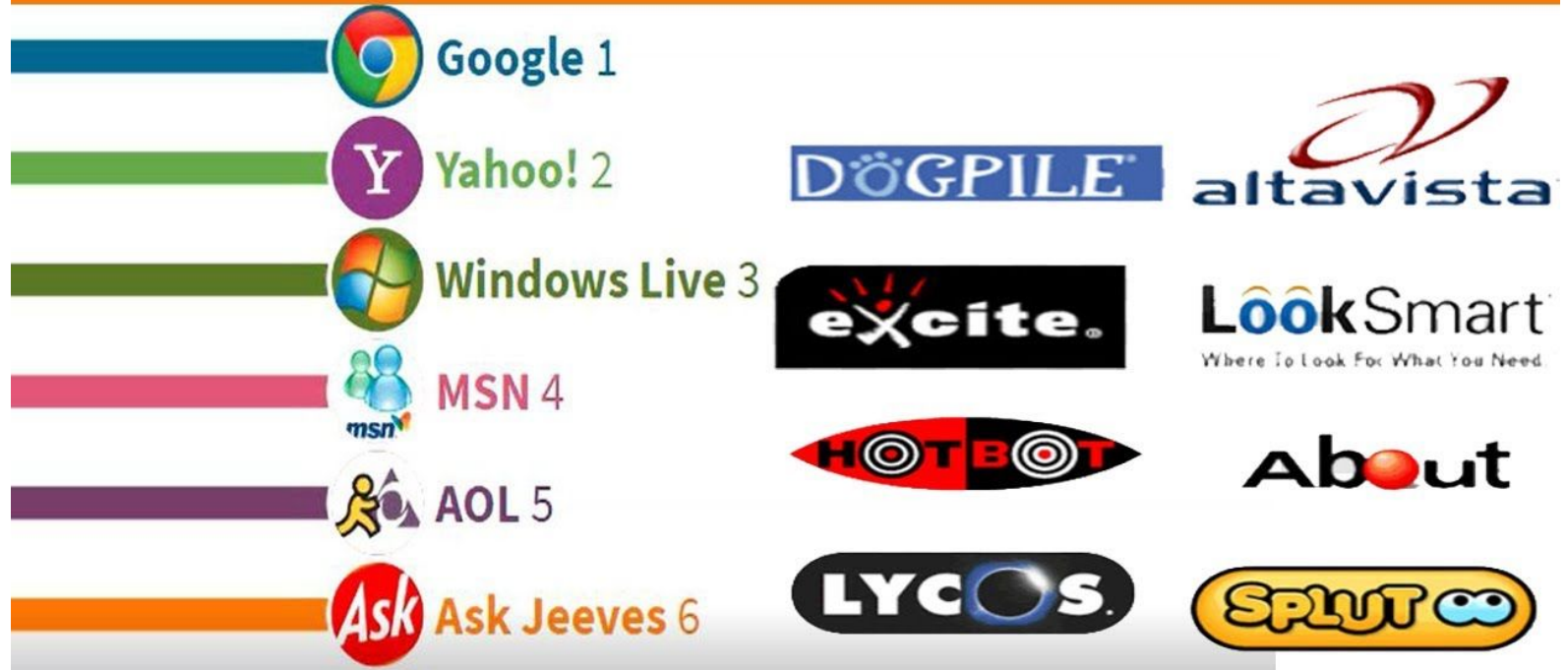
Brave Search



Baidu

Examples of Search Engines

Most Popular Search Engines



Images					
Music					
Videos					
Health					
Shopping					
Local					
Cooking					
Finance					
Jobs					

Two methods of collecting Web Data :

- Web Crawling
- Web Scraping

What is Web Scraping?

- Web Scraping refers to the **extraction of data** from a website or webpage. Usually, this data is extracted on to a new file format. For example, data from a website can be extracted to an excel spreadsheet.
- Web Scraping can also be done manually, although in most cases automated tools will be used to extract the data.
- One key aspect of Web Scraping is that it is often done with a focused approach. This means that Web Scraping projects seek to extract specific data sets from a website for further analysis.
- For example, a company might extract product details from laptops listed on Amazon in order to figure out how to position their new product in the market.

What is Web Crawling?

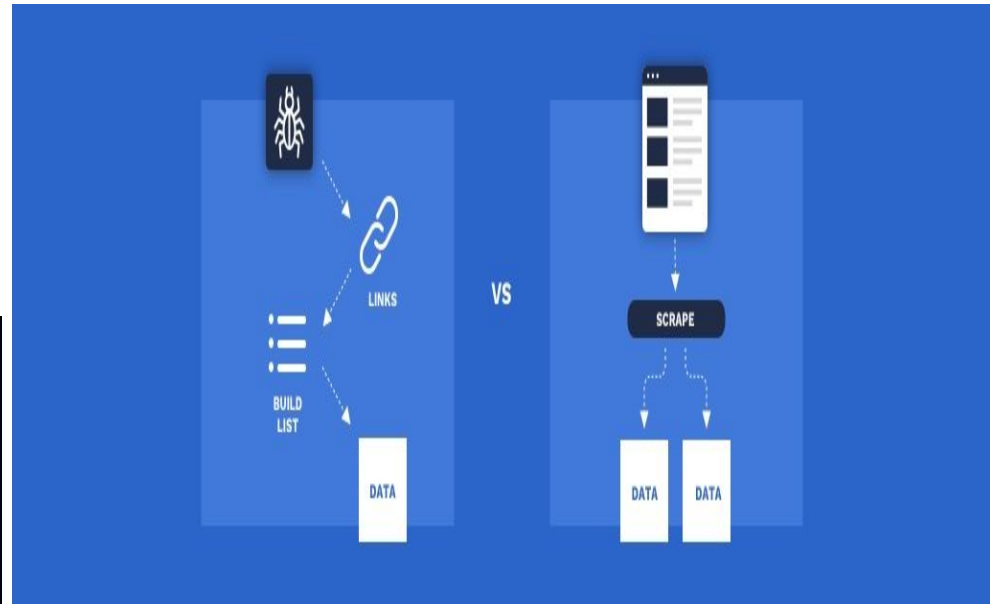
- Web Crawling refers to the process of using bots (or spiders) to **read and store all of the content** on a website for archiving or indexing purposes.
- Search engines (such as Bing or Google) use web crawling to extract all the information from a website and index it in their search engines. That's how Google can tell what pages will have the information you're looking for.

DATA CRAWLING

- INDEXING PAGES
BASED ON THE CONTENT
- CRAWLER BOTS

DATA SCRAPING

- EXTRACTING
INFORMATION FROM
THE CONTENTS OF
THE PAGE
- SCRAPER BOTS



WEB CRAWLING

- GENERATING **SEARCH ENGINE** RESULTS
- MONITORING **SEO** ANALYTICS
- PERFORMING **WEBSITE** ANALYSIS

WEB SCRAPING

- GENERATING **LEADS**
- COMPARING **PRICES**
 - STOCK MARKET **ANALYSIS**
- MANAGING BRAND **REPUTATION**
- MARKET **RESEARCH** FOR NEW **PRODUCTS**
 - ACADEMIC AND SCIENTIFIC **RESEARCH**
- COLLECTING **DATA SETS** FOR **MACHINE LEARNING**



4:38 / 6:10 • Use cases >



Types of Search Engines :

- Based on the way data is stored and extracted , there are three types of Search Engines :
 - Crawler-based
 - Directories
 - Metasearch engines.

Crawler based Search Engine

- Crawler-based search engines create their databases or lists automatically, without any human intervention.

E.g. google, bing

Crawler-based search engines are widely used to find and access content over the Internet.

They operate in three steps: 1) web crawling, 2) indexing, and 3) searching.

- **Web crawling** involves collecting and storing information about web pages
- **Indexing** involves categorizing and storing this data in a database for quick access using an index for a web page
- **Searching** involves using keywords to query the index and provide a list of relevant web pages ranked according to various factors

- **All search engines need to have crawlers**, some examples are:
 - Amazonbot is an Amazon web crawler for web content identification and backlink discovery.
 - Baiduspider for Baidu
 - Bingbot for Bing search engine by Microsoft
 - DuckDuckBot for DuckDuckGo
 - Exabot for French search engine Exalead
 - Googlebot for Google
 - Yahoo! Slurp for Yahoo
 - Yandex Bot for Yandex

Directories

- Directory listings are compiled and **created by human editors**
- Website owners can submit their website for inclusion in the directory
- The submission is reviewed by the editor before being added to the directory
- Some examples of human-created directories include **Yahoo Directory, Open Directory, and LookSmart.**

MetasearchEngines

- Metasearch engines compile and display results from other search engines.
- When a user enters a query, the metasearch engine submits the query to several individual search engines, and results returned from all the search engines are integrated, ranked, and displayed to the user.
- Examples of meta-search engines include Metacrawler, Mamma, and Dogpile.
- By integrating results from several search engines, metasearch engines are capable of handling large amounts of data and can help us save time by focusing on one search engine.

Types of Search Engines:

- Based on their scope, search engines can be divided into two types: 1) local and 2) global.
- **Local search engines** is embedded within a website and only indexes and searches the content of that website. Amazon's CloudSearch or any other search engine embedded within a website is an example of local search engine.
- **Global search engines** are used to search for content on the web. Examples of global search engines include Google and Bing

Search Engine Analytics

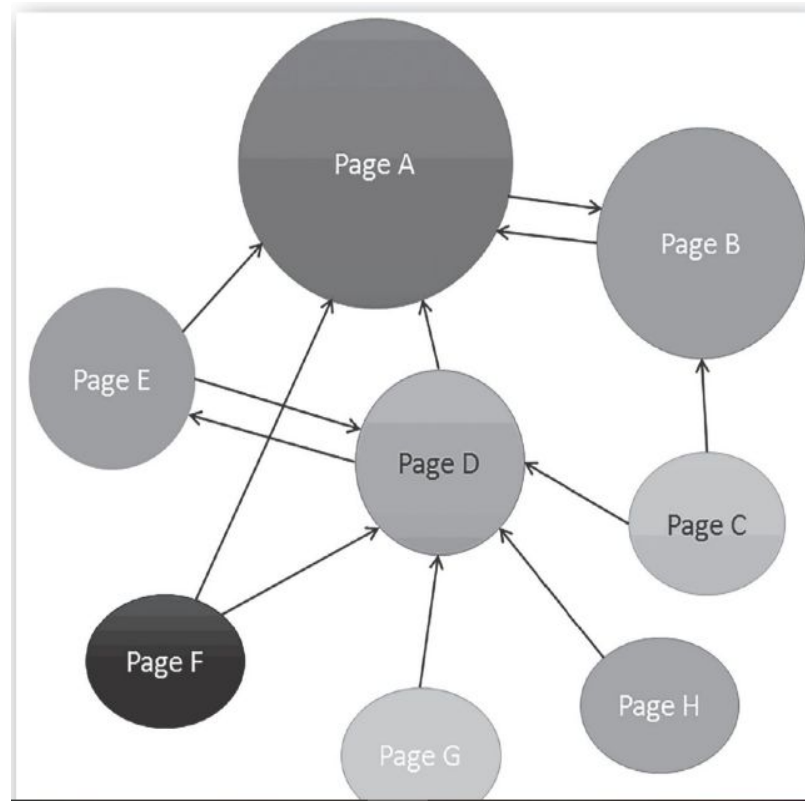
- Search engine analytics involves analyzing and interpreting data from search engines
- It helps website owners understand and improve their website's performance in search results
- It can involve analyzing various data points such as the number of visitors, keywords used, pages visited, and website position in search results
- Search engine analytics can be used to track the effectiveness of SEO efforts and guide future SEO strategies
- There are two types of Search Engine Analytics
 - 1) Search engine optimization
 - 2) Search engine trend analysis.

Search Engine Optimization

- **Search engine optimization (SEO)** is the process of **improving a website's ranking** on a search engine results page (SERP)
- A SERP is the list of the results returned by a search engine in response to a user's query.
- SERPs have both **organic and nonorganic results**,
 - Organic results being based on relevance to the user's query and
 - Nonorganic results being paid advertisements
- It is important for websites, especially commercial ones, to have a high ranking on SERPs because it can lead to more traffic and potential customers
- The main factor that determines a website's ranking on a SERP is its **PageRank**, which is determined by the quality and number of incoming links to the website
- Tools like **Open SEO Stats** can be used to check a website's PageRank and other relevant information such as traffic, hyperlink status, and speed of the page.

Google's PageRank algorithm

- Google's PageRank algorithm predominantly relies on the quality of incoming hyperlinks (or in-links) to rank websites.
- A website, for example, with in-links from a famous websites (e.g., cnn.com) will appear on the top of the SERP if compared with a website with no quality inlinks or many low-quality in-links.
- To understand the in-link quality and number argument, consider following figure, where nodes represents web pages and lines represent in-links (arrowhead pointing to a page) and out-links (arrowhead pointing away from a page).
- The PageRank algorithm will place page B higher on the SERP, even though there are fewer in-links to B when compared to D.
- The reason for this ranking is that in-links to website B are from an important website; that is, A.
- Bottom line, your objective is to increase the number of quality in-links to your website



Search Trends Analysis

- Search engine trend analytics involves **analyzing and understanding the keywords** used in search engines.
- **Google Trends** is a comprehensive tool for analyzing search engine trends and predicting future trends.
- It can be used to understand **how people search for a particular brand**, when interest in products or services spikes, which keywords drive the most traffic, and **how a brand's competitors are performing**.
- Google Trends data has also been used in various industries to **detect early warning signs and track flu epidemics**.
- From a business perspective, Google Trends can help **answer questions about search trends and patterns related to a brand**.
- <http://www.google.com/trends/>

Search Trends Analysis of your brand

- How people search for your brand?
- When does interest spike in your products or services?
- Which keywords drive more traffic?
- Which regions are interested in your brand?
- What are trending topics over the Internet?
- How are your competitors performing?

Veganism
Topic

+ Compare

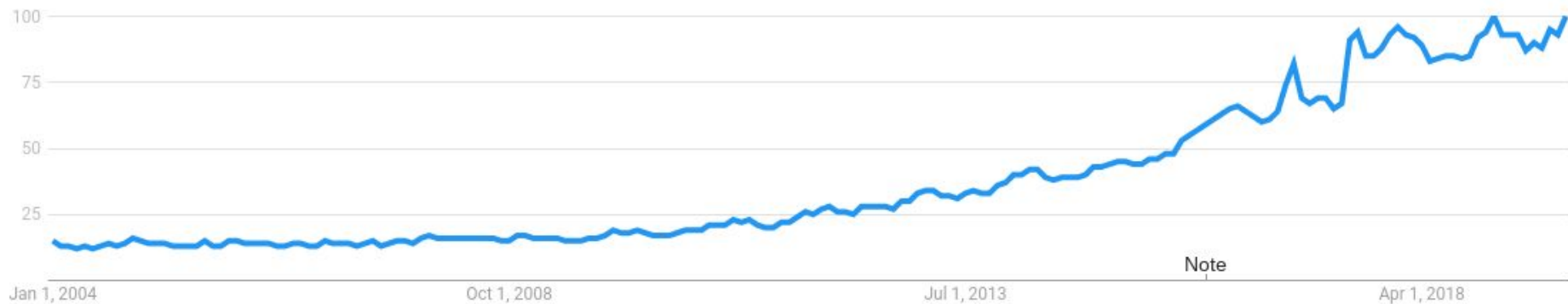
Worldwide ▾

2004 - present ▾

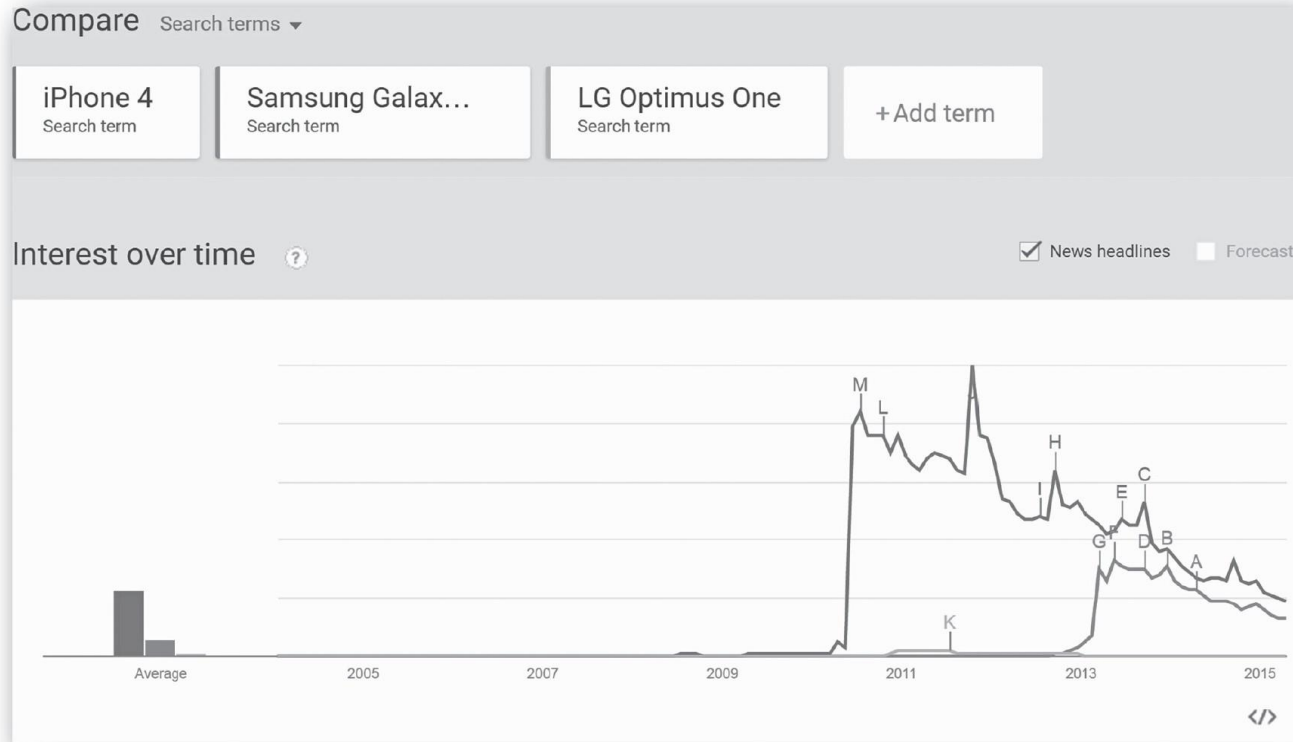
All categories ▾

Web Search ▾

Interest over time ?



Google Trends



Type of Analytics Provided by Google Trends

- Year in Search
- Trending Searches
- Trending on YouTube
- Top Charts
- Explore
- Subscription

Regional Interest

Regional interest ?

iphone 4 samsung galaxy s4 lg optimus one



Region City	
Ho Chi Minh City	100 <div></div>
Hanoi	90 <div></div>
Singapore	63 <div></div>
Bangkok	52 <div></div>
Paris	51 <div></div>
Dallas	47 <div></div>
London	44 <div></div>



Understanding the Research Function

Search Terms	The Results Displayed
Social media analytics	Results can include searches containing the terms “social,” “media,” and “analytics” in any order.
“Social media analytics”	Results will only include the exact search terms included inside of the quotation marks.
social + analytics	Results can include searches containing the words “social” OR “analytics.”
social – analytics	Results will include searches containing the word “social,” but will exclude searches containing the word “analytics.”

Grouping Search Term

Grouping Search Terms

Overall, you can search up to five groupings at one time, with up to twenty-five search terms in each grouping. Consider the following example.

- ✓ iPhone 2 + Galaxy S2 + LG G (Grouping 1)
- ✓ iPhone 3 + Galaxy S3 + LG Optimus F3 (Grouping 2)
- ✓ iPhone 4 + Galaxy S4 (Grouping 3)
- ✓ iPhone 5 + Galaxy S5 + LG Optimus F6 (Grouping 4)
- ✓ iPhone 6+ Galaxy S6 (Grouping 5)

Customizing the Search

- Google Trends allows users to customize search results by country, year, category, and type of web resource
- Users can also select custom data ranges from the date dropdown menu
- Trend data is only available from 2004 onwards.

Questions

- What is the function of a search engine?
- Explain different types of search engines.
- Differentiate between local and global search engines.
- What is search engine analytics?
- Explain the two main categories of search engine analytics.
- What is the purpose of search engine optimization?
- What is the purpose of search engine trend analysis?