Internet Search; Beyond the Basics

Brief Outline:

1: The **proper use of names** for search engines and browsers.

Google is NOT a browser. Chrome is Googles browser. Google is a search engine and a company. To Google something is synonymous with doing a web search, e.g. **"Google it".**

Google Chrome or just **Chrome**.
Mozilla Firefox or just **FireFox**.
Microsoft Edge or just **Edge**.

2: A **short history of searching**, **the old**.

The earliest search tools were **Gopher** (protocol) using **Veronica** as its search engine and **Archie** for searching **FTP** sites, these were used prior to HTTP (protocol) to search the World Wide Web. This dramatically reduced the amount of time required to find programs and documents. In the late 1980s, getting serious value from the Internet meant knowing how to use Gopher, Archie and Veronica.

3: **Searching, the new way**. All search engines use web spiders or crawlers to index websites.

The first thing you need to understand is what a Web Crawler or Spider is and how it works. A **Search Engine Spider** (also known as a crawler, Robot, SearchBot or simply a Bot) is a program that most search engines use to find what's new on the Internet. Google's web crawler is known as **GoogleBot**. There are many types of web spiders in use, but for now, we're only interested in the Bot that actually "crawls" the web and collects documents to build a **searchable index** for the different search engines. The program starts at a website and follows every hyperlink on each page.

4: **Types of search engines**, general and specialized.

General Search Engines:
https://www.google.com/webhp?gws_rd=ssl   Google
https://search.yahoo.com/   Yahoo
https://www.bing.com/   Bing (Microsoft)
https://duckduckgo.com/   DuckDuckGo (no tracking)
http://www.thesearchenginelist.com/   The Search Engine List

Specialized Search Engines:
https://scholar.google.com/   Google Scholar
https://censys.io/   Computer Science
https://www.shodan.io/   Computer, network and internet security
https://hulbee.com/   Semantic information recognition
https://images.google.com/   Image Search
https://www.youtube.com/   Video search
http://www.imdb.com/   Movie search

Fun Searches:
http://elgoog.im/tilt/
google atari breakout site:images.google.com
google games site:images.google.com

Internet Country Codes for searching:
http://www.web-l.com/country-codes/

These are just some of the many types of search engines available.

We will be using Google search for this presentation but most of these commands will work in other search engines. You will need to set your default browser search setting to **Google Search** for this exercise.

We will be using FireFox and the built in page search (**Edit>Find**) this is also available in other browsers such as Opera or Chrome.

Search all of googles database, images, youtube, etc.

First some basics:

https://www.google.com/webhp?gws_rd=ssl   Google Search Page (Home)

Use the Tabs:

The first tip is to **use the tabs in Google search**. **On the top of every search** are a number of tabs. Usually you will see Web, Image, News, and More. Using these tabs, you can help define what kind of search you need to do. If you need images, use the Image tab. If you are looking for a recent news article, use the News tab. It is rudimentary and most people use the tabs already. If you are not, then it is highly recommended to get associated with them. They can cut search times dramatically if utilized properly.

Use Google search to do math

8 * 5 + 5   Answer: (45)

Keep it simple

Now we're getting into the general tips. Google search knows how to search for a lot of things. What this means is you don't need to be too specific. If you need a pizza place nearby, use this to search.

Pizza places nearby

Google search will grab your location and deliver a variety of results about pizza places that are near you.

Common Google Search Operators

"Search Term"
This operator searches for the exact phrase within speech marks only. This is ideal when the phrase you are using to search is ambiguous and could be easily confused with something else, or when you're not quite getting relevant enough results back. For example:

"Tinned Sandwiches"

This will search for only the finer tinned variety of the [bread based snack](#), at the exclusion of all others.

OR
This self-explanatory operator searches for a given search term OR an equivalent term. For instance, if you have a fascination with the famous 'Sheens' you could search for:

"Martin Sheen" OR "Charlie Sheen"

– (and +)
The **–** operator removes pages that mention a given term from search results. For example, if you were searching for information about Manchester, but didn't want your results to be polluted by information about the city's [red clothed football team](#), you could search for the following:

Manchester -united

This would return results for "Manchester", while removing any that feature the word "united".

Using **+** forces Google to return common words that might ordinarily be discarded, for example:

Peanut Butter +and Jam

~
Adding a tilde to a search word tells Google that you want it to bring back [synonyms](#) for the term as well. For example, entering **"~set"** will bring back results that include words like "configure", "collection" and "change" which are all synonyms of "set". Fun fact: "set" has the most definitions of any word in the dictionary.

site:
This searches only within a given domain – delectable when you want to only search within the confines of a particular site. This really narrows your searching. Note the spaces between words when searching, the exception being the colon and URL.

site:https://spcugpcoutlaws.wordpress.com/ Steve Solberg
site:msn.com healthcare
site:newegg.com SSD 1TB
site:https://www.irs.gov/ 0.9 percent medicare funding
site:http://www.peninsuladailynews.com/ lost hikers

Search related sites:

Put "**related:**" in front of a web address you already know. For example, **related:time.com**.

Comparisons:
VS.

apple vs. pc
mac vs. pc pros and cons
aca vs. ahca
Google vs. Bing 2017

link:
Use this operator to find links to a domain. Bonus note: Google only provides a sample of backlinks, meaning that this operator isn't very useful for uncovering the complete selection of links to a site, but it is good for quickly identifying a sample of sites that link to a specific domain. For example:

link:bynd.com

For a more complete selection of backlinks, use the Yahoo! operator –**linkdomain:**

inurl:view/view.shtml

Will reveal a list of webcams

*
Putting an asterisk in a search tells Google 'I don't know what goes here'. Basically, it's really good for finding half remembered song lyrics or names of things. If you put the asterisk in a search like:

I've got a brand new pair of *

Google will fill in the blank and tell you that you've got a brand new pair of Belgians, hopefully. Though it'll more than likely be rollerskates.

+ (immediately before query)
Google is now providing a wide range of synonym results in response to relevant search queries. For example, if I search for "California", Google knows that this is the same as "CA" and will also return results for California but if I want the full word I need to use *+California*

related:
Simple: it returns searches for sites that are related to a given domain. This one is interesting for testing Google's semantic perception of a given domain, for example:

related:msn.com

..
Use two full stops to search in a range of numbers, for example:

I own 1..100 cats

Will bring back results that encompass searches on "I own 1 cat" to "I own 100 cats". Totally useless.

Search social media

Put **@** in front of a word to search social media. For example: **@twitter** or **@facebook.**

Search for a price

Put **$** in front of a number. For example: **laptop $400**.

Search hashtags

Put **#** in front of a word. For example: **#throwbackthursday**

Exclude words from your search

Put - in front of a word you want to leave out. For example, **corvette speed -car**

info:website
Using this operator will tell Google to bring back information about a certain domain. It reveals:

- Google's cache of the site
- Pages that are similar to the one you searched for
- Pages that link to the domain you searched for
- Other pages on the same domain
- Pages that contain the domain text on their page

loc:placename

This operator brings back results from pages in a given place. Even better, it can be used to search for specific types of places within that location, for example:

loc:Brighton pub   or loc:Dynasty Restaurant sequim

Will mostly return pages for pubs that are in Brighton or Chinese restaurants in Sequim.

define:phrase
Google can do anything. If you search:

define:ululate   or define:spcug

It will bring back six definitions from different websites, from Wiktionary to encyclopedia.com.

daterange:
This query will search within a given date and time range.

daterange:genealogy

Search within a range of numbers

Put **..** between two numbers. For example, **camera $50..$100**.

source:
This is one that only works in [Google News](#) search. If we look for:

Gordon Brown source:the_guardian

Google will show all the mentions of Gordon Brown in articles where [The Guardian](#) is identified as the news source.

location:
This is another news operator that allows you to search for articles by location:

location:London

filetype:
This lets you search for a certain filetype. For instance:

filetype:music mp3

Will bring back only MP3 results. Useless note: you can also use the extension "ext:" to do exactly the same thing.

movie:
If you search:

movie:Iron Man 2

And then enter a location, Google will tell you where you can see the film and at what time. This operator can also be used in conjunction with the aforementioned "loc:"

phonebook:
This seems to only work in the US, but if you search:

phonebook:john smith

You'll be given a worrying list of phone numbers for people called John Smith.

weather:
This is a great and simple one:

weather:sequim

Will bring back results both for Sequim pages on weather websites, as well as a little weather widget at the top of the results page.

stocks:
I use this query to track the stock price of my investment portfolio – AND NOW YOU CAN TOO. Just use the operator followed by the company ticker symbol that you wish to receive information on, for example:

stocks:BAC

Will show stock information for [Bank of America](#).

cache:

cache:spcug.net
Shows Google's most recent cache of a webpage.

map:

map:sequim
Adding the word map after a locational search forces Google to produce map-based results.

in

Google can be used as a calculator. As part of this functionality, "**In**" is a superb function that can be used (among many other Google calculation operators) to work out the number of units of something in something else. For example:

mph in speed of light

Other Operators not commonly used:

allintitle: (and also intitle:)

Searches only for sites with the given word(s) in the page title. **Intitle:** does the same thing but for single words and can be used with more flexibility. For instance, if I searched:

intitle:hammer nails

The results would show pages with just "hammer" in the page title, and with "nails" elsewhere. Note: in [blog search](#) this same function is performed by **inblogtitle:** and **inposttitle:**

allintext: (and also intext:)

This operator searches only for sites where the given word(s) are in the text of the page.

allinanchor: (and also inanchor:)

This shows sites which have the keyterms in links pointing to them, in order of the most links. For instance, if I searched for *allinanchor: helicopters*, Google would show me the top sites which are linked to, where the anchor text for the link is "helicopters".

allinurl (and also inurl:)

Similar to the last few, but fetches results where the key words are in the URL. This is useful if you've forgotten the exact URL of a website, but can still remember bits of it. Note: in blog search this same function is **blogurl:**, making it handy for searching for topics on specific platforms. For example:

blogger blogurl:wordpress

Would find WordPress blogs that are – paradoxically – talking about Blogger.

References:

https://en.wikipedia.org/wiki/Gopher_%28protocol%29

https://en.wikipedia.org/wiki/Veronica_(search_engine)

https://en.wikipedia.org/wiki/Archie_search_engine

http://www.wpthemesplanet.com/2009/09/how-does-web-crawler-spider-work/

http://www.lifehack.org/articles/technology/20-tips-use-google-search-efficiently.html

https://bynd.com/news-ideas/google-advanced-search-comprehensive-list-google-search-operators/

https://www.gcflearnfree.org/searchbetter/google-search-tips/1/

https://www.gcflearnfree.org/searchbetter/all-about-online-search/1/

Steve Solberg
SPCUG
May 20, 2017