



School of Law
UNIVERSITY OF GEORGIA

Prepare.
Connect.
Lead.

University of Georgia School of Law
Digital Commons @ University of
Georgia School of Law

Continuing Legal Education Presentations

January 27, 2017

Jan 27th, 8:15 AM - 9:00 AM

Found It! Getting the Most Out of Your Web Searches

Anne Burnett

University of Georgia School of Law Library, aburnett@uga.edu

Follow this and additional works at: <https://digitalcommons.law.uga.edu/cle>



Part of the [Legal Writing and Research Commons](#)

Burnett, Anne, "Found It! Getting the Most Out of Your Web Searches" (2017). *Continuing Legal Education Presentations*. 1.

<https://digitalcommons.law.uga.edu/cle/2017/2017/1>

This Event is brought to you for free and open access by the Alexander Campbell King Law Library at Digital Commons @ University of Georgia School of Law. It has been accepted for inclusion in Continuing Legal Education Presentations by an authorized administrator of Digital Commons @ University of Georgia School of Law. [Please share how you have benefited from this access](#) For more information, please contact tstriepe@uga.edu.

Found It! Getting the Most out of Your Web Searches

Anne Burnett, J.D., M.L.I.S.
University of Georgia Law Library
Athens, Georgia

Found It! Getting the Most out of Your Web Searches

Anne Burnett, J.D., M.L.I.S.
University of Georgia Law Library
Athens, Georgia

TABLE OF CONTENTS

Contents

I. Introduction.....	1
II. What's New?.....	2
III. A Sampling of Search Engines Other Than Google.....	3
III. Crafting Advanced Searched Queries	6
IV. Google Enhancements & Services	12
V. What is the Invisible Web?	15
VI. Conclusion.....	16

I. Introduction

With Google reporting that the number of searches it handled doubled from 2012 to 2016, its overall searches likely hit above the 2 trillion mark last year. (*Google Now Handles at Least 2 Trillion Searches per Year*, SEARCH ENGINE LAND, <http://searchengineland.com/google-now-handles-2-999-trillion-searches-per-year-250247> (last visited January 3, 2017)). Given those numbers, it is no surprise that Google retains the lion's share of the search engine market, with NetMarketShare reporting in December, 2016, that Google had 77.82% of the search market, followed by Bing (7.96%), Yahoo! (6.39%) and Baidu (6.36%).

While the five most popular searches in 2016 sought information about the following:

1. Pokémon Go
2. iPhone 7
3. Donald Trump
4. Prince, and
5. Powerball

(*Year in Search 2016*, GOOGLE, <https://www.google.com/trends/yis/2016/> (last visited January 3, 2017)), the internet continues to provide plenty of content for **legal researchers** as well. The American Bar Association reports that 91.6% of respondents to their tech survey use free online resources to conduct legal research, with 35.1% reporting that they actually **start** their research with a general search engine or

directory (AMERICAN BAR ASSOCIATION, 5 2016 LEGAL TECHNOLOGY SURVEY REPORT 34-35). Given the value of attorneys' time, maximizing the efficiency of web searching is paramount. This paper aims to provide advanced tips and strategies for crafting effective search queries as well as providing suggestions for resources useful for specific purposes. The emphasis herein is on Google's advanced searching tools, because of that search engine's popularity, but most other general search engines allow for similar search strategies to increase relevancy in retrieving results.

II. What's New?

Google and Facebook Change Policies in Response to Fake News Sites

Effective searchers constantly assess the reliability and authoritativeness of web content, but "fake news" sites have proven increasingly effective at fooling searchers, social media consumers and even mainstream news outlets. In late 2016, Google stated that "[w]e recently made improvements to our algorithm that will help surface more high quality, credible content on the web." Similarly, Facebook recently revised its Facebook Audience Network policy to include "fake news" sites in the types of sites in which it will not display ads.

Search Engine Queries Can Help Detect Cancer in Web Searchers

Microsoft computer scientists published research in the Journal of Oncology Practice detailing how they were able to analyze search queries to identify web searchers who would later be diagnosed with pancreatic cancer. The computer scientists first identified web searchers who had been recently diagnosed with pancreatic cancer and

then looked backwards at their search queries over the previous months. They used those search queries to then recognize searching trends which could help identify additional searchers currently entering the same search queries, successfully predicting who might also have pancreatic cancer with a low false-positive rate. John Paparrizos, Ryen W. White and Eric Horvitz, *Screening for Pancreatic Adenocarcinoma Using Signals from Web Search Logs: Feasibility Study and Results*, 12 JOURNAL OF ONCOLOGY PRACTICE 737-44 (2016).

III. A Sampling of Search Engines Other Than Google

Section IV below includes strategies for using Google more efficiently and effectively. An additional strategy used by savvy researchers is employing alternative search engines, directories or other search tools when a Google search is not successful or when a specific type of tool is likely to be more effective.

This section lists some of these alternatives:

Bing
www.bing.com

LifeHacker recently compared Google and Bing, finding many similarities between the two search engines. *Search Engine Showdown: Google vs. Bing*, LIFEHACKER, <http://www.lifehacker.com.au/2016/08/search-engine-showdown-google-vs-bing/> (last visited January 4, 2017). Here are some of the takeaways:

- Bing's video search is better
- Bing provides more autocomplete suggestions

- Google's shopping suggestions are better than Bing's
- Bing includes a feature predicting whether flight prices will increase or decrease

Yahoo!

www.yahoo.com

Bing's search engine (Microsoft) actually powers Yahoo! searches, giving the two search tools a decent combined share of the overall search market. This veteran search tool still finds ways to tweak the search query, providing some tricks not currently available in Google. For example, searchers can specify the order of search keywords, meaning that the results will only include sites where the first listed search term appears before the other search term, even if the terms do not compose a phrase. Yahoo! is worth keeping in the web researcher's tool box, especially if a search in Google or Bing returns unsatisfactory results.

DuckDuckGo

duckduckgo.com

If privacy is a concern, try DuckDuckGo, which promises that it does not store searches, track searchers or transmit information to third parties. The search interface is simple and clean, reminiscent of the original Google, but DuckDuckGo provides plenty of bells and whistles, including its famous "bangs." An exclamation mark at the beginning of a search query triggers a popup box of available sites that can be searched directly from the search box. For example, a search of "!cfr wild horses" directly searches the electronic Code of Federal Regulations for the terms "wild horses." DuckDuckGo also allows for extensive customization with settings for location, language, results display and site appearance.

Omni Knowledge Discovery Engine

<https://www.omni.io/>

Omni allows the searcher to upload a document which is scanned and then Omni's "associative semantic search technology" matches the scanned document with the most similar documents in its library. Searchers can also type in text search queries or cut and paste snippets of text from other sources. The product is free for "Public" accounts but does require registration (name and email address). Public accounts allow for the submission of up to ten documents at a time. Results can be exported as PDFs. "Enterprise" (paying) accounts may load an infinite number of documents and export search results in a variety of formats, including PDF and CSV. As an example of Omni's potential usefulness, TechCrunch envisions "a law firm that must track millions of documents and court cases and have them ready for instant recall" using Omni to create their own database of documents and then searching within that database.

Wolfram|Alpha

www.wolframalpha.com/

Wolfram|Alpha is a Computational Knowledge Engine, or a question and answer site where the answers are computer-generated. Its stated goal is "to make all systematic knowledge computable and broadly accessible." Users submit a textual question in the input box which is invitingly labeled "enter what you want to calculate or know about." The results box includes a statement of how the system interpreted the question. The results also suggest additional queries that the system determines may be relevant.

Wolfram|Alpha is especially useful for statistics and calculations. A search for **eastern towhee** returns a clean graphical representation of information about the

bird, including its size, range, lifespan and an image. A query of **\$84.22, 15% tip, 3 people** rapidly calculates how much each person owes. It is, of course, limited by the data and algorithms of which it is comprised. A search for **lawsuits filed Cobb County** returned basic statistics about the county with no litigation statistics.

A Wolfram|Alpha app is available for both Android and iPhone.

Quora

www.quora.com

Quora is a question and answer site where the questions and answers are provided by the users. The answers are rated by readers. Examples of recent questions and answers on Quora include:

- Do psychopaths know they are psychopaths?
- Can you explain the plot of Game of Thrones in just four words?
- What are some of the most outrageous court decisions you have ever witnessed while working as a lawyer?
- What are the most fascinating and the worst aspects of working as a lawyer?

The site is free, but requires registration. Quora apps are available for both Android and iPhone. For a newer crowdsourced question & answer site, take a look at Jelly (<https://askjelly.com/>).

III. Crafting Advanced Searched Queries

Spend time up front in crafting a precise query and you'll spend far less time meandering through page after page of results. Following a few conventions and

learning some key symbols to include in the string of characters entered in the search box can greatly improve your results. Many of these strategies will work with numerous search engines, but look for a link to “search help” or “search tools” or “advanced searching tips” or something similar on any search site to get exact symbols and strategies:

- Google: <https://support.google.com/websearch/#topic=3378866> links to several search help pages within Google, including Search Operators and Search Tools & Filters.
- Bing: <http://onlinehelp.microsoft.com/en-us/bing/ff808415.aspx> provides some searching tips and techniques
- Yahoo:
<https://help.yahoo.com/kb/search/SLN2242.html?impressions=true>

The following tips and tricks focus on Google due to its ubiquitous use; however, a few different conventions for Bing and Yahoo are also noted.

Use quotation marks

Surrounding a phrase with quotation marks is the most common way to accomplish phrase searching. Google, Bing and Yahoo all use quotation marks to search for an exact phrase.

Exclude words from your results

Use the minus sign directly in front of words you do not want to appear in your search results. **Vikings –football –team** should produce results about the

seafaring raiders rather than the sports team and a search for **RICO –Puerto** should exclude many irrelevant hits from a search for sites about the racketeering statute. (Note: Yahoo also uses the minus sign. In Bing, use the word “NOT” (all capitals) in front of words you do not want included in your results.) Caveat: use the minus sign or NOT limiter carefully as you may unwittingly eliminate relevant sites from your results.

Find definitions

Type “define:” followed by the word you want defined, and Google will provide a list of web sites providing a definition. Try **define:infest**. Preceding a search term with “what is” often produces the same results.

Search for synonyms using ~ and alternate terms using OR

Google routinely searches for synonyms, but inserting the tilde symbol ~ directly in front of a word will search for even more synonyms. For example, ~**police** will also retrieve pages containing “law enforcement,” “cops” and “military.” To search for results containing related terms that might not be considered by Google to be synonyms, use a capitalized OR between the two terms. Google gives the example of **miniature OR standard** to have one search retrieve sites containing either word.

Use wildcard operator Use an asterisk (*) to formulate a search query asking Google to fill in the blank.

Christopher Columbus died in * retrieves numerous sites providing the year of Columbus' death (it was 1506, by the way).

Limit by language

Google's advanced search form allows researchers to limit search results by language. Simply choose the language in the drop down box.

Search (or exclude) a specific domain

Google allows searchers to limit their results to pages within a specific domain by typing "site:" before the desired domain abbreviation. For example, a researcher may desire only pages assigned a government, or .gov, domain. A search of **improving school lunch site:.gov** will only retrieve pages with the .gov domain. Placing a minus sign in front of the same search will retrieve pages from all domains but the typed one, so that **improving school lunch -site:.com** will find sites with .edu, .org, and .gov domains while excluding those with a ".com" domain.

When searching for foreign materials, specifying both a language and a country-specific domain often retrieves the best results. For example, a search for **bananas industry site:.sv** limited by English in the drop down box retrieves a relevant list of pages about El Salvador's banana industry in English. A search for **"climate change" site:.mx** limited to English might produce sites in English with the Mexican perspective on climate change. The Internet Assigned Numbers

Authority (IANA) provides a complete list of domain extensions at <http://www.iana.org/domains/root/db/>.

Searching a specific web site

The “site” limiter also allows the researcher to further limit the search by limiting the results to pages within a specific web site. Type “site:” followed by the web address. A search for “**constitutional law**” **site:www.law.uga.edu** will only retrieve sites containing the phrase “constitutional law” on the UGA School of Law web site.

Limit your results by a number range

Google allows you to specify a specific number range by placing two periods between the desired numbers. This is useful for searching for ranges of things like dates, prices and measurements. A search for **average temperature 70..80** retrieves an envy-inspiring list of places such as Canary Islands and Cyprus. The number range limiter works best when the search engine can ascertain from the context of the query exactly what is being limited.

Force search for exact word

Google assists searches by automatically including results for certain synonyms and variant spellings and abbreviations. To bypass this automatic expansion of your search query and force Google to return results including only the exact word as typed, place quotation marks around the word. For example, a search for “**canine**” **diseases** will only return results including the word “canine” whereas

the same search without the quotation marks would also return results with the word “dog.” In Yahoo, place a plus sign directly in front of the word you want to appear exactly as entered; for example, **+childcare** will only retrieve results containing the word “childcare” while those sites using “child care” will be excluded.

Pay attention to word order

The order in which words are entered in a search query usually do not make a difference in the total overall results list; the difference may be in which sites make it to the *top* of the results list. Search engines like Google index words as *bigrams*, or words next to each other in a specific order, and the closer your query is to the order in which the words appear in relevant sites, the more relevant your top results will be.

Proximity searching

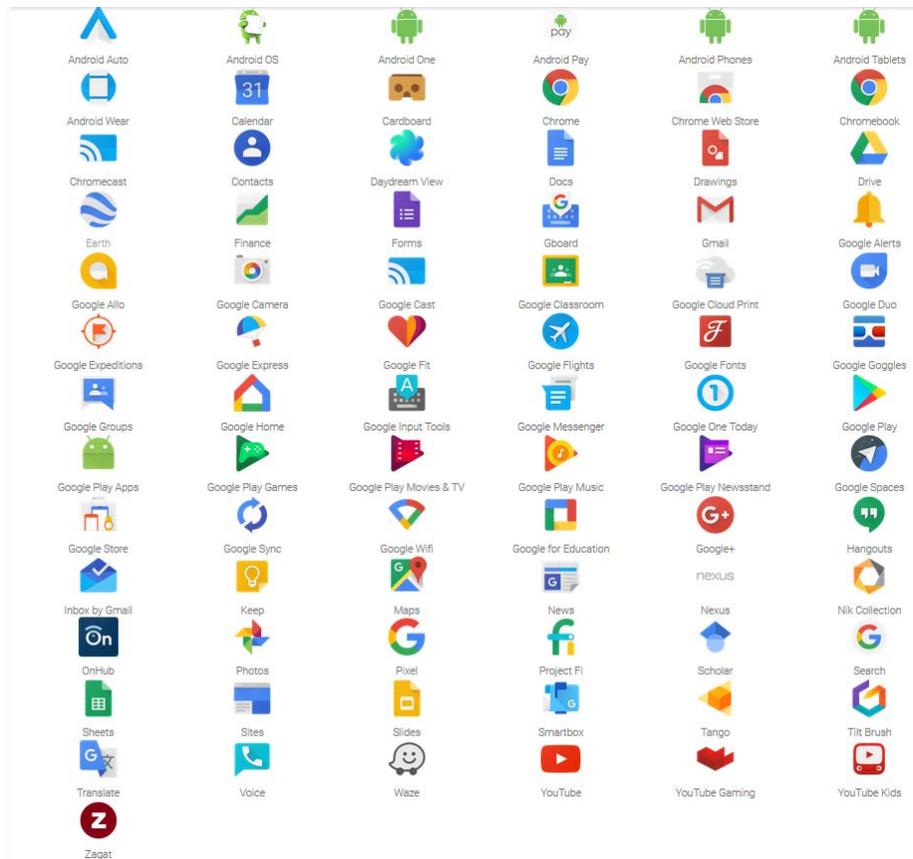
Google has long been thought to lack any kind of proximity searching, but numerous web searching blogs started reporting in 2011 that Google had quietly implemented proximity searching several years earlier. Use a capitalized **AROUND(n)** between keywords, replacing the “n” in parentheses with a number signifying the maximum number of terms between the two keywords. For example, a query of **John AROUND(2) Smith** searches for all the possible combinations that a proximity search such as ‘John within 2 of Smith’ would create, thereby allowing the searcher to simultaneously search for John Smith, John X. Smith, John Xxxx Smith or Smith, John.

Use the Advanced Search Box

Searchers can use a search engine's "advanced search" function to build a search query incorporating many of the above tools without having to remember all of them. Google no longer links its advanced search page from its home page, but a quick search query of "Google Advanced Search" retrieves the page quickly. Once a search query is formulated using the advanced search boxes, the resulting URL can be altered to quickly revise the search. For example,

IV. Google Enhancements & Services

Google's "even more" page at <https://www.google.com/intl/en/about/products/> lists its many different products, from specialized search tools (alerts, Scholar) to social networking (Google+) to collaboration tools (Calendar, Drive, Hangouts) and much more. Prominently featured at present is Google Assistant with its related products, Google Home and Google Allo. Constantly in development, with ever-increasing apps and tools, Google's many projects are too numerous to list, but this screen capture shows some of the many features currently on offer, with a few of the most helpful tools and features also highlighted below:



Google Alerts

www.google.com/alerts)

Google Alerts sends email updates of latest results of a specified query. Searchers specify whether the email alert should be sent once a week, once a day, or as-it-happens. Setting up an alert is easy and quick, and searchers can modify existing alerts.

Google My Activity

<https://myactivity.google.com/myactivity>

My Activity (formerly Search History) allows searchers who are signed into Google to review their recent interactions with Google products, including conversations with Google Assistant, regular Google search queries, a record (and

map) of physical locations visited and a copy of the contacts, calendars, apps, and other data from your devices. Users can pause tracking of these activities as well as delete the record of past activities.

Google Translate

<http://translate.google.com/?hl=en>

Google Translate will automatically translate text, web pages, and files between over 100 languages. As with any automated translator, Google Translate is useful for getting the gist of a site but cannot be relied upon to give a sophisticated translation.

Google Scholar

scholar.google.com

Google Scholar searches for articles, theses, books, abstracts and court opinions, from academic publishers, professional societies, online repositories, universities and other web sites. If searching for court opinions in Google Scholar, check the box next to Legal Documents under the search box. The search form then allows the researcher to limit the results to federal court decisions or to those from specified states. Google Scholar also serves as a citator by listing in the sidebar the cases which subsequently cite the displayed case.

Searchers activating the My Library function in Google Scholar can easily save search results for easy retrieval in the future.

Google's inclusion of the following disclaimer

Legal opinions in Google Scholar are provided for informational purposes only and should not be relied on as a substitute for legal

advice from a licensed lawyer. Google does not warrant that the information is complete or accurate.

leads to a similar statement here: do not assume that the results include all relevant case law on a specific legal issue! This collection should serve solely to give the researcher an idea of the law or to help locate a specific case.

V. What is the Invisible Web?

Not all content on the web is accessible through search engines like Google or Bing. In fact, the vast majority of web content is missed by traditional search engine spiders and crawlers, which cannot “see” the content buried in subscription sites or generated on-the-fly by databases. Referred to as the “Deep Web,” “Invisible Web” and “Hidden Web,” this content is estimated by the Deep Web Source Repository to be at least 400 to 500 times the size of the “free web” (also referred to as the “surface” or “visible” web).

Portals and directories are useful for locating content that might not be findable by Google’s crawlers. Following are a few examples:

- FDsys (<https://www.gpo.gov/fdsys/>) is a portal for federal government information and searches across 50+ government collections. Note that FDsys will soon be replaced by the new Govinfo.gov site.
- AllAcademic (<http://research.allacademic.com/>) serves as a searchable archive for academic conference proceedings, paper abstracts, free academic journals and other material related to academia.

- The Internet Archive (<https://archive.org/>) includes over 279 billion pages of archived web sites from 1996. This is especially useful for finding a snapshot of content that has been removed from the web. To view an older version of a web site, simply enter the URL in the site's Wayback Machine text box. In addition to archiving regular web pages, the site also captures images, books, movies, music and software (one of the more popular collections is the Internet Arcade which is a web-based library of arcade (coin-operated) video games from the 1970s-90s, emulated in JSMESS software).

Efforts to make the contents of the invisible web more accessible are constant and ongoing. Many educational institutions and libraries endeavor to provide additional tools and resources for searching this content, and traditional search engines such as Google are also constantly working to improve their ability to index this content as well.

VI. Conclusion

Realistically, most searchers, including busy lawyers, are content to type a few keywords into the Google search box and see what they get. The author of this paper is no exception. However, when a quick Google search turns up poor results, remember these tips and tricks for refining a search, and also keep a few of the alternative tools in your back pocket when Google falls short.