

Hide and seek on the Internet

By Donna Moughty



The ease and accessibility of the Internet has attracted many new people to the field of genealogy, and provided experienced genealogists with new tools and faster communication. The amount of genealogical information is exploding and being able to find pertinent information is critical to effective research. But even though the tools have changed, the research basics we've learned over the years haven't. We need to remember to:

- Set an objective and develop a plan
- Work from what we know
- Evaluate each piece of information. Does it make sense? Does it fit with what we know?
- Document the source for all information
- Consider new evidence with an open (and inquiring) mind

Remember, you can't do it all on the Internet

There are two basic types of searching on the Internet. First is a basic Internet search using a favorite search engine. This type searches HTML or Web pages and returns a list of sites that fit the criteria we have given. This is where we will focus. The second type of search is a database search.

This search looks at data that has been structured and matches our criteria to specific fields, such as first name, last name, or location.

So how does a Web search engine work?

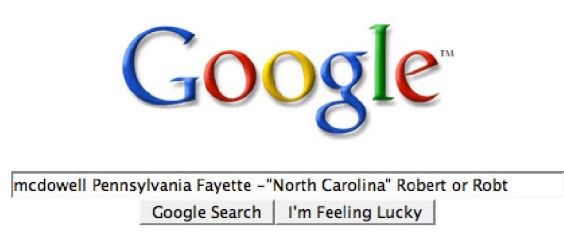
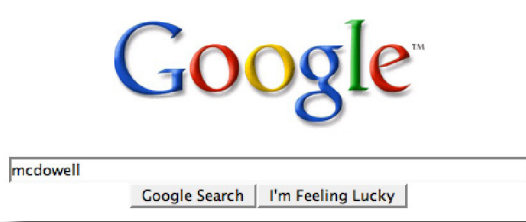
Search engines have robots (sometimes called "bots" or "spiders") that crawl the Internet and create an index of the words on the Web pages they visit. When you type a word or words into your search engine, the index allows a quick response to your query, returning all pages that match your search criteria. There are many different Web search engines and it's important to note that not all of them search the same Web sites.

This explains why you can get different results by using different search engines. Some of the most popular search engines are [Google](#), [Yahoo](#), [MSN](#), and [AltaVista](#).

There is another type of search engine called a "meta" search engine. These do not have their own robots, but search multiple indexes created by other search engines. They can be helpful in getting different results. [Dogpile](#) and [Clusty](#) (formerly known as Vivisimo) are meta engines.

Boolean searching

How do you get the results you want out of a search engine? Boolean searching. Don't panic... no need to break out in a sweat (like in ninth grade algebra). A Boolean search allows you to add qualifiers to your search to limit or expand the number of relevant hits. (See D. Joshua Taylor's column "[A quick refresher—search engines.](#)") For example, type a name into your favorite search engine, McDowell for example, and in .16 seconds you are treated to 12,500,000 hits. Now, I don't know about you, but there is a slim chance that I'm going to check out all 12 million hits. Let's stop and



think about what we actually want to find. I'm interested in genealogy so, logically I can add that word and limit the search. So now I type "McDowell" and "genealogy" into my search engine and in .14 seconds, I've reduced my number of hits to 387,000. That's a pretty good reduction from 12 million, but it's still unlikely I'm going to check them out. Where were my McDowells living? In Pennsylvania, so let me add that to my query...down to 139,000. Can I be more specific? They lived in Fayette County, Pennsylvania, so let's add Fayette to our query. Now we're down to 46,000. What we've been doing is working with the Boolean qualifier "and." Each time we add another word to our query, the search engine returns only pages where all of those words appear. This happens to be the default for Google and many other search engines (check the Frequently Asked Questions (FAQ) for your search engine), so if you just type a list of words, every word you type must appear somewhere on the page in order to provide a hit. That doesn't mean they need to be clustered together, they just have to appear somewhere on the page.

What if I want the words to appear exactly in the order I type them? By placing quotes around a specific phrase, such as "Robert S. McDowell," only pages with the exact phrase will be returned. In this case, 548 of which four referred to my Robert S. McDowell.

As I check out some of the results, I discover that I'm getting a number of hits that refer to McDowells in other states where there is a county named Fayette. How do I get rid of these results? By using a minus sign (-) immediately preceding the word. Here I'm going to use two Boolean operators: the quotes and the minus sign. To construct the query, I now type: McDowell genealogy Pennsylvania Fayette - "North Carolina" - "West Virginia". That brought me down to 647 hits. I'm getting close to something I might be able to look at.

My ancestor was Robert McDowell. Sometimes I find him listed as Robert and sometimes with the abbreviation Robt. What if I add the two given names, connected with another Boolean operator—OR—that tells the search that one of these names must appear, but not necessarily both? My new query is constructed: McDowell genealogy Pennsylvania Fayette - "North Carolina" - "West Virginia" Robert or Robt; it returns forty-five hits.

Advanced searching

Now that you understand the basics of Boolean searching, you don't have to memorize the syntax. Just use the advanced search feature of your favorite search engine.

Advanced searching allows you to limit your search in other ways. Are you researching an ancestor whose native language was not English? Try searching for the surname in the native language. On the Google Advanced Search page, choose the language from the drop-down menu. You can't read that language? When you're presented with the results, click on "Translate this page" to the right of the title. Although the translation may not be perfect, you'll get the general idea of whether or not it pertains to your research. In Alta Vista, you can use the "Babel Fish Translation" function. Just copy



and paste what you want to translate into the window and select the translation you want from the drop-down menu. You can also translate Web pages by typing in the URL and providing the translation request from the drop-down menu. [I would not recommend using this to translate letters to request information from a foreign country because, as I mentioned above, these translations aren't perfect. Instead, use the letter forms on FamilySearch.org which have been written by native speakers.]

Advanced searching also lets you select the format for your hits. Do you only want pictures? What about content that is new in the last two months (great for repetitive searches)? You can also limit your search to words occurring only in the title or the URL. If you're searching for medical information perhaps you only want information from a known, reliable Web site such as the Center for Disease Control. There is a great deal of flexibility when using advanced searching.

Meta engines

The search process doesn't change with meta search engines. The difference is that these search engines don't have their own robots, but use the indexes already created by other companies. Clusty (formerly known as Vivisimo) is one of my favorites, since it clusters its results into categories. A search on Moughty turns up the results of a young, very talented competitive swimmer named Kaitlin, as well as results for Sarah (a producer at Frontline) and Kelly (a youth minister in Washington, D.C.). Moughty also is used in literature as slang for "mighty." I can easily skip over those results and look more closely at the ones that refer to genealogy, family history, Ireland, or Argentina.

Wikis and blogs

Other results that might show up include "wikis" and "blogs." Wikis are open and editable Web pages. The idea is to let anyone add information about a topic (i.e., Wikipedia). The person may or may not be reliable; you don't know the background of the individual adding the information, so buyer beware! Blogs are written by individuals, so you should also treat them as any other non-primary source. They can be a wonderful place to find new information, and are being used more and more for research notes. An example of a blog used for research notes is at Steve's Genealogy Blog.

Individual Web site search engines

If you find a Web site that has its own search engine, use it! Robots may not index all of the pages on a site, or they might not index them regularly. I love USGenWeb and it's one of the first places I go to see what's available when I'm researching a new area. I knew the Keck family was originally from Onondaga County, New York; however, I had been told by my client that the marriage I was looking for had taken place in Long Island. Unable to find the marriage, I used the "Search this site" feature on the county Web site. A local individual had indexed the Presbyterian Marriage books and there was the marriage I had not been able to find!

In the past, not all USGenWeb sites had a search function; however [RootsWeb](#) has added a universal search. The page has some excellent examples of searches and uses some Boolean operators, including near, except, and priorities.

So instead of slogging through too many results, create your Internet queries in a way that will get you to the pertinent information quickly. That gem you've been looking for may simply be hiding. **DG**

[Donna M. Moughty](#), a professional genealogist and former regional manager for a national technology company, has been conducting family research for more than fifteen years. She teaches classes for beginners and lectures on a variety of subjects including Internet, Irish research, and computer topics. In addition, she provides consultations, research assistance, and training. Donna has attended the National Institute for Genealogical Research, the third and fourth Irish Genealogical Congresses in Ireland, and the Salt Lake Institute–Irish Research. She is a member of the Association of Professional Genealogists and coordinator for its professional management conferences, treasurer of the Genealogical Speakers Guild, and a delegate to the Federation of Genealogical Societies. Donna writes a weekly [blog](#).