

UNLOCKING GENEALOGY DOORS: THE KEY MAY BE IN SEARCHING/BROWSING TECHNIQUES

Jean Wilcox Hibben, PhD
Riverside County, California
circlemending@gmail.com
<https://circlemending.org>
909-994-6114

© 2022 – J.W.Hibben

GOALS FOR THIS INSTRUCTION

- ✓ To learn the basic steps of filling out search forms/templates
- ✓ To discover the variety of uses for “keyword” searches
- ✓ To become more comfortable with browsing and learn its value
- ✓ To understand some of the many “buzz words” related to Internet use, in general, and achieving successful searching/browsing results in particular

WHAT THIS LECTURE WILL INCLUDE (time permitting)

- How keywords are written/ordered can alter results
 - Misspellings can affect searches
 - Try alternate spellings or common spelling errors to change results
- Algorithms are developed through past searches & affect future discoveries
 - One of the reasons for cookie collection
 - You can “clear the cookies,” but that will affect future searches (aided by algorithms) where you might want the computer to remember what you had visited previously (including saved passwords)
- Think like a marketer to find the right terms for a search
- Use “advanced” options in Search Engines, when available
- Use a Boolean search to clarify what you are looking for
 - Quotation marks
 - AND, OR, and NOT
- Be flexible – try fuzzy searches for spelling, locations, and dates
- Searching OCR-identified keywords can be hit and miss (if all else fails, browse!)
- More = Less, Less = More
 - To narrow the search and get fewer results, add descriptors
 - To widen the search and get more results, remove descriptors
- Use a keyword search for anything involving
 - Finding other people at the same address
 - Locating a business when you have just an address
 - Discovering your ancestor by searching for someone else you know lived nearby
 - Determining the exact terms in title or subject of a collection/database is unknown
- Read the “about the collection” section on the target website to be sure that your search entity is even likely to be there
- Read the guidelines or lessons (or view videos) about successful searching provided by the website you are visiting

Etymology trivia: The term “database” has its origin in early computer use. Coined, as far as anyone knows, by computer scientist and IBM employee Edgar Frank “Ted” Codd (1923-2003). The exact date of development of the term is not clear, but sometime in the early- to mid-1960s (“Edgar F. Codd,” *Wikipedia* online at https://en.wikipedia.org/wiki/Edgar_F._Codd, 30 March 2021). From the combination of *data* and *base*.

GLOSSARY

Algorithm: The process of reaching a specific result because of a step-by-step action, moving from one step to the next in a prescribed manner; the instructions for completing an operation.

Artificial Intelligence (AI): Computer-controlled tasks, simulating “human intelligence.” (see OCR)

Boolean Search: A means by which the search engine user can optimize a search, giving the parameters of AND, NOT, and OR; these terms in the search can increase applicable results and decrease irrelevant findings; some SEs use the Boolean method and automatically apply AND; term emerged from the work of logician George Boole (see the references for more involved explanations) (<http://www.boole.org.uk/gb.htm>).

Collection: See “Database” and “Record Collection.”

Cookies: Memory tools for the computer so it knows what your interests are; often used to target products that you might be interested in buying. Accepting them can pose a minimal risk to the computer; they usually can speed up your computer searches and results.

Database vs (Record) Collection: Databases are divisions of information (on a given website); on Ancestry®, an example would be “1940 United States Federal Census,” consisting of smaller databases for the different divisions (geographic regions, further divided by counties, villages, etc.), into separate collections.

Descriptor: The modifiers that further define a search (e.g., if trying to find a firefighter who worked in Chicago in the 1920s, those words would be the descriptors in that same search: *firefighting Chicago 1920s*, HOWEVER, keep in mind that in pre-about 1990s sources, the masculine form of the word would be more likely to realize helpful results: “fireman” instead of “firefighter”).

Filter: Another term for descriptor or qualifier – it “strains out” or “filters” those things you are *not* seeking by listing what you are.

Fuzzy search: The opposite of “exact.” An example would be using the Soundex system to search for a name where the spelling can be similar, or “sounding like” what you have entered; when doing a search for a year of an event, giving the parameters of +/-5, it allows for the event to have been recorded for any year within 5 (plus or minus) of what you think (or know) the year to be . . . it’s “fuzzy.”

GDPR: General Data Protection Regulation is the set of rules that govern what is and is not legally acceptable re: collecting and processing the personal information of people living in the EU (European Union); regardless of location of the website and the user, site operators (webmasters, sponsoring organizations, etc.) must comply with the regulations or risk facing legal ramifications; its purpose is related to the privacy of European citizens, including how data collected will be used. (Keep in mind that, even if you and the website/company are located in the United States, not all users may be in a similar situation, so websites must be compliant with a larger range of regulations than those in the immediate, geographic region . . . that's why it's called the WORLD Wide Web.)

Keyword: The ideas/topics that one is searching for (e.g., if researching an ancestor who was a firefighter, one might include “firefighting” in the search).

OCR: Optical Character Recognition – using a visual graphic to find matches for the terms used in the search; by using the image, deciphered as words/numbers (scanned documents are often “searchable” through this medium), the computer determines matches, which can easily be misinterpreted due to misspellings, ink smears, and other elements that make the reading difficult. (A form of AI)

(Boolean) Operators: AND, OR, and NOT; quotation marks, and parentheses.

PPC: Pay-Per-Click – keyword search development with the consumer in mind; there is a financial component for website, consumer, and/or host.

Qualifier: Another term for “descriptor” that adds the “qualifications” for the search; usually an adjective (e.g., *1920s* qualifies that the search needs to be confined to that decade).

Record Collection: Often used as a synonym of “Database”; think in terms of hard copies of documents vs. electronic (digitized) versions. You wouldn't have a document database that is hard copy only (that would be a collection), but you might have an electronic database that was developed from a hard copy collection. My father's math library collection was not digitized (that was not an option when he was teaching), but now, it might be that his library holdings have been digitized, creating “Dr. L. Roy Wilcox's Math library database, IIT.”

SEO: Search Engine Optimization (a goal for the website to use the keywords for which people are searching); a keyword search tool using analysis of popular words sought (optimizing the results).

URL: “Uniform Resource Locator” – being a type of acronym (an *initialism*) for a website's address.

Wild Cards: The symbols used to replace unknown characters (e.g., to find Johnson, if the name spelling might be Johnsen, try searching for *Johns?n*; if a number of letters are unknown, but it's unclear but how many, the * replaces the unknown, *John**); when using wild cards, include as many descriptors as known (e.g., *H. P. John* Milwaukee Wisconsin Farwell avenue 1913*).

World Wide Web (www AKA “The Web”): An application; it is accessed via The Internet using HyperText Transfer Protocol (HTTP – or HTTPS, informing the user that the website is secure. Also, see text box just above the Glossary).

RESOURCES

- Ned Batchelder, “Database: What’s a Base?” *NedBatchelder*, online at https://nedbatchelder.com/blog/200509/database_whats_a_base.html, September 2005; accessed by Jean Wilcox Hibben, 9 May 2021.
- “A Boolean Search Could Help You Find Your Ancestor,” *FamilyTree.com*® blog, online at <https://www.familytree.com/blog/a-boolean-search-could-help-you-find-your-ancestor/>; accessed by Jean Wilcox Hibben, 10 May 2021.
- Tara Calishain, “Super Search Secrets,” *Familytree*, online at <https://www.familytreemagazine.com/uncategorized/super-search-secrets/>; accessed by Jean Wilcox Hibben, 10 May 2021.
- David Deady, “The Beginner’s Guide to Boolean Search Terms,” *Social Talent Blog*, online at <https://www.socialtalent.com/blog/recruitment/the-beginners-guide-to-boolean-search-terms>, 5 August 2020; accessed by Jean Wilcox Hibben, 10 May 2021.
- The Figleaf Team, “The Pros and Cons of Disabling Cookies,” *Figleaf: Perspectives*, online at <https://figleaf.com/blog/perspectives/pros-cons-disabling-cookies/>, 26 November 2019; accessed by Jean Wilcox Hibben, 9 May 2021.
- Jake Frankenfield, “Laws & Regulations: General Data Protection Regulation (GDPR),” *Investopedia*, online at <https://www.investopedia.com/terms/g/general-data-protection-regulation-gdpr.asp>, 11 November 2020; accessed by Jean Wilcox Hibben, 9 May 2021.
- Howcast, “How to Delete Cookies,” *YouTube*, online at https://www.youtube.com/watch?v=FVLR_ZhwYPQ, December 2019; accessed by Jean Wilcox Hibben, 9 May 2021.
- Dayna Jacobs, “My Cheatsheet for Genealogy Boolean Searches,” *On Granny’s Trail* blog, online at <https://ongrannystrail.com/2015/06/07/my-cheatsheet-for-genealogy-boolean-searches/>, 7 June 2015, accessed by Jean Wilcox Hibben, 10 May 2021.
- Richie Koch, “Cookies, the GDPR, and the ePrivacy Directive,” *GDPR.EU*, online at <https://gdpr.eu/cookies/>, 2021; accessed by Jean Wilcox Hibben, 9 May 2021.
- “Searching: Truncation, Wildcards and Proximity,” *JSTOR*, online at <https://support.jstor.org/hc/en-us/articles/115012261448-Searching-Truncation-Wildcards-and-Proximity>, 2019; accessed by Jean Wilcox Hibben, 9 May 2021.
- Mindy Young, “How to Use Keywords: A Recipe for Success,” *Online Writing Jobs*, online at <https://www.onlinewritingjobs.com/writing-tips/how-to-use-keywords-a-recipe-for-success/>, 1 June 2017; accessed by Jean Wilcox Hibben, 9 May 2021.

Etymology trivia: Why have I seen the word “Internet” written with a capital I? As with any names, conventional usage calls for the capitalization of the first letter (e.g., Utah Genealogical Association). By definition, an internet (not capitalized) refers to an inter-connected network (such as an internal system within a defined scope – e.g., the military or school – and so is used as a generic term – a common noun. When it entered the public sphere, with access to the World Wide Web (www), it was called THE Internet (as opposed to one of many “internets” throughout the world) and the capitalization shows that the word is being used as a proper noun. It’s a matter of context, and, as with so many names (e.g., “Kleenex,” “Xerox,” etc.), folks tend to use those words in place of the more descriptive term (i.e., “facial tissue,” “photocopy,” etc.) it is very likely that, over time, “Internet” will become just as common-place as those. (<https://proofed.com>, 2019) This is already becoming evident in URLs (dropping the “www” in many cases) and the use of The Internet as a synonym for the World Wide Web (it’s not). (<https://Differ.com>)