

DATABASE SEARCHING BASICS

How databases store information and the rules they use to retrieve that information vary from database to database but share some very basic premises. Some represent organizational techniques common to all databases (e.g. Boolean operators, stopwords) while others represent less common techniques (e.g. plurals, proximity). The following represents a general overview of search terminology and strategies that apply to most databases. To learn the exact search criteria for a given database, find a link labeled *help* or *tips* or some other variation on providing assistance with using the database (e.g. *how to...*, *searching in...*, etc.).

Restriction/stemming (e.g. “search term”)

► Placing double quotation marks (some databases use single quotes) around a phrase or single word will restrict a search to that exact word or phrase. For example: a search for the term *draw* will yield results containing that word as if it were a stem (e.g. *drawn*, *withdrawn*, *drawing*, etc.) whereas a search for “draw” will limit the results to just the word *draw*. Similarly, searching for *riparian corridor* will yield results for papers that include just *riparian* and just *corridor*, as well as papers that include the full phrase *riparian corridor*. This would likely produce a very long list of results that you would have to read and evaluate, making your search inefficient. Binding the multiple words in quotes, “*riparian corridor*,” will restrict the results to records containing **that exact phrase** with *riparian* occurring immediately before *corridor* with no intervening words or letters. However, searching “*riparian corridor*” would NOT return papers in which this was treated as a plural, i.e., *riparian corridors*. This is a disadvantage of using quotation marks in a search phrase that may be a minor or major problem, depending on your search phrase.

<i>search term</i>	<i>search entry</i>	<i>results (all papers containing the following)</i>
draw	draw	drawn, drawing, withdrawn, drawl
draw	“draw”	draw
riparian corridor	riparian corridor	riparian, corridor, riparian corridor
riparian corridor	“riparian corridor”	riparian corridor

Note: Some databases automatically treat multi-word phrases as if they were in quotation marks and when you include quotes they search for items that include the phrase and the quotation marks. The usual result is that ZERO items are found. When you encounter such a result, try re-entering the phrase without the quotation marks OR dividing the search phrase into smaller phrases using the AND operator or placing them in different search fields.

Field codes (e.g. *au*, *ti*, *su*, etc.)

► More often than not, databases will have their field codes available in a pull-down menu to one side of the main search field box. However, using these field codes can often be used to shorten search time and yield more limited, relevant results.

Examples(from JSTOR):

ti:(cat dog) will find all items with the words *cat* and *dog* in the **title** in any order

au:"peter reiche" will find all items with the phrase *peter reiche* in the **author** field

su:(peacekeeping "united nations") will find all items with the word *peacekeeping* and the phrase *united nations* in the **subject** field.

Note: If you are only interested in primary research papers that deal with your topic, first try searching for your keywords in the **title** (ti:) field. This will return only those items in which your keyword was a major focus of the paper.

Note: If you are searching for papers written by a particular author, you might want to combine the authors name in the **au:** field with another keyword in the title or subject (**su:**) field to eliminate papers by other researchers who have the same last name and initials. This is a particularly good idea if the researchers name is something common, like J. Smith.

Boolean operators (e.g. *and, or, not*)

► Stringing search terms together using Boolean operators makes searching for various topics easier and more efficient. Also knowing that databases use Boolean logic as a default in their search methodologies can help users better refine their searches. AND searches yield only those results which contain **both** terms. OR searches yield results which contain *either* one *or* the other term. NOT searches *disregard* those results which contain the word following NOT. Using restriction means that each quote-bounded phrase is processed as a single search term thus limiting the results obtained and reducing the number of items that are not relevant to your search.

search

civil war OR stone river

"civil war" OR "stone river"

"civil war" NOT "stone river"

processed as

(civil AND war) OR (stone AND river)

(civil war) OR (stone river)

(civil war) NOT (stone river)

Note: In some databases, when a Boolean operator is used to connect two multi-word search phrases, the search automatically treats the multiple words in each phrase as if they were connected by an AND, as in the first example above.

Stopwords (e.g. *a, an, are, as, at, be, but, by, for, from, had, have, he, her, his, if, into, is, it, no, of, on, such, that, the, their, then, there, these, they, this, to, was, which, will, with, you*)

► These are words that most databases will not recognize because they appear with such frequency in text and will not affect a search. The words *and, or, not*, however, are recognized by most databases as Boolean operators and will affect your search if you include them. For example, if you do a search for the title *Fire and climatic change in temperate ecosystems of the western Americas* the "and" near the beginning of the title will be recognized as a Boolean operator and return results for items containing the individual words in the title: *fire, climatic,*

change, temperate, ecosystems, western, Americas. To avoid this, place the entire title in double quotation marks and specify a title search by using either a pull-down menu or a field code.

Truncation & Wildcards (e.g. *, ?, #)

► *Truncations* (usually designated with * or #) can be used to find various forms of a word. Searching for *cat** will yield *category, catalog, catatonic*, etc. If looking for general information on *birds* you might type in *bird**. This search will yield results which contain the words *birds, birding, birdman*, etc.

► *Wildcards* (usually the question mark ?) can be used in the middle of words to indicate a missing or non-existent letter. For example, placing a wildcard in *M?cDonald* will yield *McDonald* (non-existent letter) or *MacDonald* (missing letter). This can help get to results even if the word entered in the database has been misspelled or spelled alternatively. For example, *color* can also be spelled *colour*. A wildcard search for *colo?r* will yield results containing both spellings. Conversely, use wildcards if you are not sure of the spelling of a word: is it *dependant* or *dependent*? Searching with *depend?nt* will yield results containing both variations.

Plurals (e.g. &, +)

► The use of plurals is far from widespread among databases but they are increasing. Make sure to check the help or tips section of the database you are trying to use to see whether or not it allows plurals searching. Plurals represent a special form of truncation. By adding the ampersand or plus symbol to the end of a word, the database in question will automatically convert Ys to Is and F to V before adding *es*. For example, *sky&* will find *sky* and *skies*; *knife&* will find *knife* and *knives*.

Proximity (e.g. with, near)

► Not every database offers this kind of searching. In fact, many databases classify these words as stopwords. Check the help or tips section of the database you are using to make sure proximity terms can be used. Using proximity terms will yield results for terms which occur next to one another within several spaces or directly adjacent in either order. Some databases do use the search order to specify the results order. Thus, placing *bronchitis* before *pollution* in the search yields only results where *bronchitis* occurs before *pollution* in the text. Some databases even allow users to specify how many words in between the two in question are allowed.

Examples (from Biological Abstracts)

search entry

results

bronchitis with pollution

both *bronchitis* and *pollution* in the same field

bronchitis near pollution

both words in the same sentence

bronchitis w2 pollution

bronchitis then *pollution* with no more than two words between them

Parenthetical grouping

► Just like in mathematical equations, differential emphasis and ordering can be given to sets of search terms by using parentheses. Processing priority moves from the innermost set of

parentheses to the outermost. Parenthetical searches are most effective with Boolean and field code searches. Some examples of this appear under other terms in this list.

<i>search</i>	<i>results</i>
(car AND safety) AND (air AND bag)	references containing the words <i>car</i> and <i>safety</i> and the words <i>air</i> and <i>bag</i> first then compares the two lists and yields only those references that contain all words
(au:dawkins) AND (su:protostomes)	searches the author field for <i>dawkins</i> and the subject fields for <i>protostomes</i>

Embargo or “Moving Wall”

► A set period of time (from a few months to a few years) for which the issues of a given journal title are unavailable electronically. Moving walls often vary with individual publication. JSTOR employs an embargo which varies from 4 to 6 years depending upon the journal title. This occurs for a couple of reasons. Some publishers want their print versions to remain attractive and so place a moving wall on electronic access to their titles. It also takes time to convert print documents into online, full-text electronic articles- whether through scanning or through software conversion and then enter their full records into the database. Subscriptions through various databases or content providers may not be full subscriptions and so the publisher limits access to older issues. Try other databases/access points, publishers’s websites, or the print versions of the titles to avoid embargoes.


Full-text availability

► Most databases provide some kind of full-text availability but some do not. If full-text is provided as part of the subscription to a database then, most likely, the article will appear in either PDF or HTML format. One or two offer other formats (e.g. JSTOR offers TIFF files as well as PDFs). It’s useful to have some kind of storage medium (e.g. flash device, etc.) with you if downloading full-text documents at a public computer station. If you do not have one, save the file to the desktop then use a web-based email server to attach the file to an email that you send to yourself. Alternatively, you can obtain an iLocker account to have online storage on a University server.

Accessibility/owned by library

► Some databases will alert users to whether or not a journal title resides in the local library but most will not. To determine whether or not a local library has a journal title, check that library’s catalog by doing a periodical title search and then checking the specific holdings (i.e. which particular volumes and issues they have). Library catalogs at this time do not allow searching of individual journal articles. Library catalogs (like CardCat) maintain records about *publication* titles only. Users may type in *Nature* in CardCat and then look at the holdings list which appears at the bottom of the individual bibliographic record for that title to see which volumes and issues the library actually owns. Another solution is to use the E-Journals list of electronically accessible journal titles.

Multilink, Citation Linker, and OneSearch

► “**Multilink-enabled**” means that [Google Scholar](#), [Web of Science](#), or whatever search engine or database you might be using has been connected to “[link resolver](#)” software that will search all the databases in a specified list for full-text versions of articles you may find in sources that do not offer full-text or for items in databases for which we have no subscription. Currently, you can activate this Multilink link-resolver by clicking the blue  button or link beside or beneath article/book citations. Multilink-enabled databases will also be indicated on the [Articles & Databases](#) list with a small icon. On occasion, Multilink will not be able to find a full-text version of your article but it will give you the option for requesting it via our [ILL services](#) (see the [Interlibrary Loan subtab](#) under **the Research**/Library Resources tab).

► Another method for finding full-text versions of articles is to use the [Citation Linker](#) application powered by Multilink. You can find it in the upper right hand corner of the [E-Journals](#) tab. Once you have a citation for an article you want, simply fill in the fields in the Citation Linker and click GO. It will then do the same thing as Multilink by searching our journals and databases for a full-text version of your article. Multilink also powers the [E-Journals](#) list which is another source for finding full-text articles.

► [OneSearch](#) refers to a search engine that allows you to search more than one Libraries database at a time, including CardCat. Not all of the databases are part of OneSearch but most are and can provide a strong results list for your search term. As always, though, using only one source for all your information- even a meta-search tool such as this- can leave your knowledge of research topics lacking. Again, there is no quick and easy road to accurate and authoritative research. Real work is involved. Search results will overlap between sources but they will NOT be the same.

DOI: Digital Object Identifier

► DOIs are special strings of numbers and letters that form a persistent link to individual publications. They are issued at the time of publication, much like an ISBN or a serial number. DOIs can be attached to a number of different publications, including journal articles, books and book chapters, conference papers, reports, and so on. The DOI system provides a technical and social infrastructure for the registration and use of persistent interoperable identifiers, called DOIs, for use on digital networks.

You may have already seen a DOI and not realised what it was – all DOIs start with “10.” and are ‘built’ according to the publisher and journal involved – for example, the DOI for the 2004 article “Post-fire survival and reproduction of rehabilitated and unburnt koalas” from the Elsevier-published journal *Biological Conservation* is **10.1016/j.biocon.2004.03.029**.

You can think of DOIs as social security numbers for individual electronic files (ie journal articles or images) which means that the article can be tracked and found by this number as

though you had typed in the full citation. DOIs act as permanent links to individual papers, documents, articles, abstracts, photographs, models, etc.

Most indexing services (ie databases/aggregators) use the DOI standard now and some citation styles like APA are beginning to require it if available. Some will allow you to search by the DOI (eg [Web of Science](#), [SciFinder Scholar](#)) but you can also use [OneSearch](#), [dx.doi.org](#), [DataCite](#), and [CrossRef](#) to “resolve” the link (i.e. find/retrieve the object in fulltext or original format).

In the Ball State University Libraries system, you can use [Citation Linker](#) to find fulltext articles with nothing more than a DOI: just copy and paste the DOI into the DOI field on the Citation Linker page and click submit. MultiLink will then search our databases for a fulltext version of the article in question or connect you to our Interlibrary Loan service to obtain it from another institution.