



Web Search Field Trip: Find it on the Web with Field Searching

By William H. Weare, Jr.

Wouldn't it be nice if you could do a simple search in Google and not retrieve millions of matches? Wouldn't it be nice if you could quickly and easily find Web documents that really match your query? Wouldn't it be nice if you could learn simple and effective Web search techniques that you could teach your students?

Many of us who search regularly—librarians, teachers, students, and others who do online research as part of their work—have learned to apply some relatively simple search techniques to Web searching, such as phrase searching and the application of Boolean logic. Most of these simple techniques mimic those available in library catalogs or library-provided electronic resources.

Field searching on the Web is similar to field searching in a library catalog or any library-provided electronic resource. However, unlike the records in a catalog or an article database, a Web document has only a few searchable fields. Web pages consist of a header (information about the document including the title), an address (Uniform Resource Locator or URL), and the body (which contains the content of the document). The title, URL, and body are searchable fields. A search for specific terms in the title field or URL will greatly reduce the number of hits and can significantly improve the likelihood of retrieving on-topic results. OneStat, a Web analytics provider, has found that most

URL. The most common type of Web syntax is the use of specific prefixes positioned before the term(s); this is the case with field searching. The remainder of the article describes specifically how to execute each of these techniques.

SEARCHING IN THE TITLE FIELD

A search for terms in the title field is best illustrated with an example. A basic search in Google for north american black bear returned more than 2,600,000 results (at the time of writing). Presumably, all of those results have those four terms somewhere in the document, although the terms may not be in close proximity to one another. The Advanced Search Page in Google includes the option to “Return results where my terms occur” and the user can select “in the title of the page.” A search for the same four terms—north american black bear—specifying that the terms be in the title of the page—yielded less than 100 results—all of which have those four terms in the title field. Because these terms appear in the title, these documents are very likely to be about North American Black Bears. Clearly, a search for specific terms in the title field results in an appreciably smaller set of focused, on-topic results than would an ordinary keyword search that looks for the terms anywhere in the document. Like Google, the Advanced Search Pages in Yahoo and Ask include the option to search for terms in the title of a particular Web document, but the Advanced Search Page in Live Search does not include this option (see screen shot 1).

Alternatively, it is possible to search for a term or terms in the title field from the search box on the main page. To locate just one term in the title field, input the syntax `intitle:term`, followed by the term to be located in the title of a document. For example, `intitle:Monet` (there is no space between the colon and the term) returns documents that have Monet in the title field. Google, Yahoo, Live Search, and Ask all support the `intitle:syntax`. It is important to understand that in Google, Yahoo, and Live Search, the use of `intitle:` will only retrieve documents that have the single term immediately following the colon in

“A field search that specifically targets the title or URL is not only an effective tool for your own use, but one that is well worth teaching to your students.”

Google, Yahoo, Live Search, and Ask—the four largest search entities that operate their own crawling and indexing technology—all support a variety of advanced search techniques. These options are accessible via the Advanced Search Page or by learning how to use the syntax that is largely unique to search engines.

FIELD SEARCHING

This article will address what is perhaps the most effective of these advanced search techniques: field searching. Very powerful and easy to learn, field searching is simply a method of locating specific terms in a specific field. We will review the field searching options available via each search engine's Advanced Search Page, and utilize a number of examples to show how to use syntax to do field searching from the main page. The latter approach provides more flexible and powerful options—and in some cases provides options that are not available via the Advanced Search Page.

searches consist of two or three terms (OneStat.com). If the location of those terms is not specified, the entire document is searched—the title, URL, and body. This type of keyword search can return millions of results. In contrast, a search confined to the title or URL would retrieve significantly fewer results than a basic search; the results are more likely to be specifically about the topic because the title or URL generally contains terms that describe the content of that document.

THE FIELD SEARCHING CHARTS

Figure 1, “Field Searching from the Advanced Search Page,” identifies which types of field searching are available from the Advanced Search Page of each of the four major search engines. Figure 2, “Field Search Syntax,” delineates the syntax used by each search engine for the two most effective field searches: searching in the title field and searching in the

the title field; the second and following terms may appear elsewhere in the document. For instance, in the search `intitle:north american black bear`, only the term `north` would be specifically searched for in the title field.

Note that the `intitle:syntax` will work with a phrase that is enclosed in quotation marks; for example, `intitle:"north american black bear"`—but keep in mind that the use of quotation marks dictates the order of the terms; if those particular terms are in any other order in the title field, that document would not be found.

Generally, a Web search will include more than one term; unfortunately, the syntax needed to find more than one term in the title field varies by search engine. Google supports the syntax `allintitle:(all in title:)`—followed by the terms to be located in the title of a document. For example, a Google search for `allintitle:north american black bear` (no space between the colon and the first term) retrieves documents in which the four terms—`north`, `american`, `black`, and `bear`—all appear in the title (see screen shot 2). In Ask, (unlike Google, Yahoo, and Live Search), it is the `intitle:` (not `allintitle:`) syntax that will locate all of the terms following the colon in the title field. Yahoo and Live Search do not support a specific syntax for locating more than one term in the title field. However, it is possible to do so: simply use the `intitle:syntax` more than once; for example, `intitle:Monet intitle:biography` will retrieve Web documents with both `Monet` and `biography` in the title field.

SEARCHING IN THE URL

Every Web document has a Uniform Resource Locator (URL)—this is the “address” of the

document. Like a title search, a search for terms in the URL, rather than in the entire document, can return a significantly smaller set of focused, on-topic results. In Google, for example, there is an option to “Return results where my terms occur” and the user can select “in the URL of the document.” Again, a search for terms in the URL is best explained with an example. A basic search in Google for the terms `asthma` and `children`—without specifying that the terms be located in the URL of the document—yielded more than 3,500,000 results (at the time of writing); a search for the same terms in the URL yielded about 750 results. The content of the documents having both terms in the URL appears to be primarily about `asthma` and `children`. The Advanced Search Pages in Google, Yahoo, and Ask all include the option to search for terms in the URL; the Advanced Search Page in Live Search does not include this option.

As was the case with searching for terms in the title field, it is possible to search for terms in the URL from the main page. To locate just one term in the URL, input the syntax `inurl:` (in URL:), followed by the term to be located in the URL of a document. For example, `inurl:asthma` (there is no space between the colon and the term). Google, Yahoo, and Ask recognize the `inurl:syntax`. Note that in Google and Yahoo, the use of `inurl:` will retrieve only documents that have the single term immediately following the colon in the URL; the second and following terms may appear elsewhere in the document. For

example, in the search `inurl:north american black bear`, only the term `north` would be specifically searched for in the URL.

The syntax needed to find more than one term in the URL varies by search engine. From Google’s main page, use the syntax `allinurl:` (all in URL:) in the search box, followed by the terms to be found in the URL of the document; for example, `allinurl:asthma children`. In Ask, the `inurl:` syntax behaves like Google’s `allinurl:` and will locate all of the terms following the colon in the URL. For example, a search for `inurl:asthma children` will retrieve documents that have both `asthma` and `children` in the URL. Although Yahoo does not support specific syntax for locating for more than one term in the URL, it will allow for the repetition of the use of the `inurl:syntax` to find more than one term; for example, `inurl:asthma inurl:children`. The documentation on the Help Page at Live Search indicates that the `inurl:syntax` is supported; however, it does not currently work.

SCREEN SHOT 1

A search for north american black bear in the title field using the Advanced Search Page in Google.

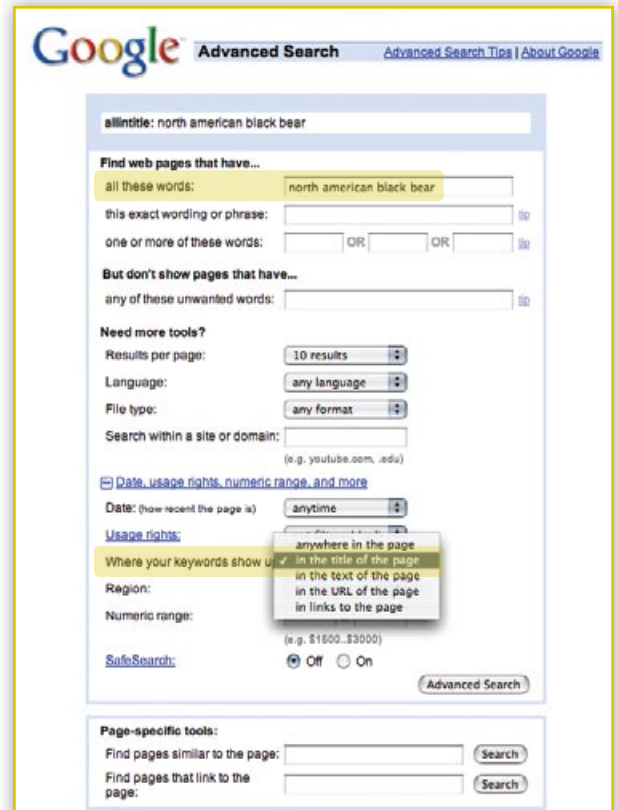


Figure 1: Field Searching from the Advanced Search Page

Is this search option available from the Advanced Search Page?				
	Google	Yahoo!	Live Search	Ask
Find term(s) in TITLE:	Yes	Yes	No	Yes
Find term(s) in URL:	Yes	Yes	No	Yes

“A search for specific terms in the title field or URL will greatly reduce the number of hits and can significantly improve the likelihood of retrieving on-topic results.”

SCREEN SHOT 2

A search for north american black bear in the title field using the *allintitle:* syntax in Google.

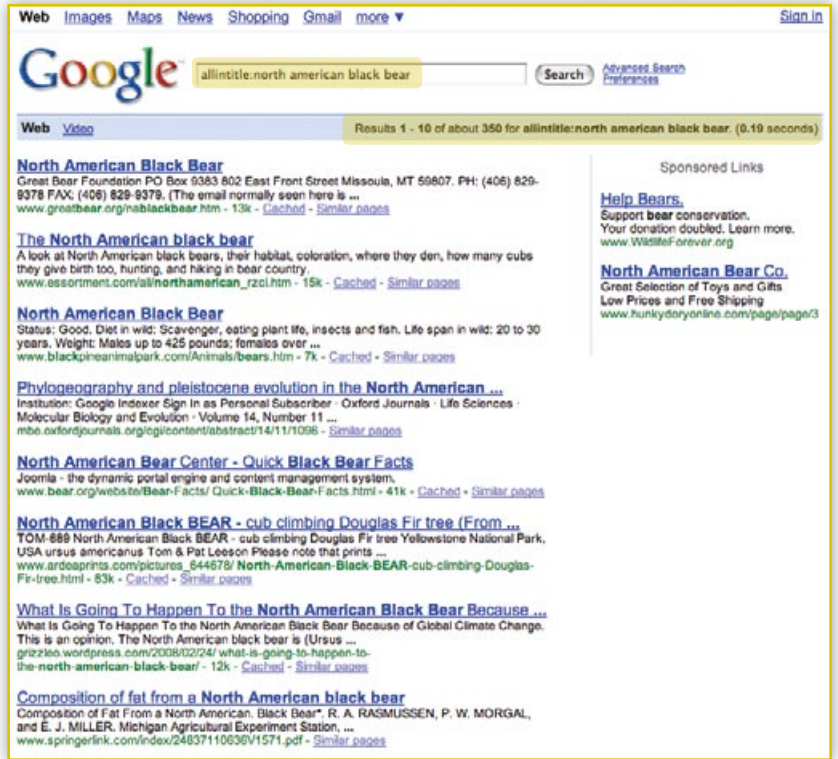


Figure 2: Field Search Syntax

SEARCHING IN THE BODY

It is also possible to limit a search for particular terms to the body of a document, and not search the title field or URL. However, a search limited to the body of the document is not particularly useful. For example, a basic search in Google for Monet and biography yielded “about 270,000” results; limiting the query to the text of the document via the advanced search page yielded “about 256,000” results. This is typical: for most searches there does not appear to be a significant difference in the number of results returned between a search for terms in the entire document (title, URL, and body), and a search limited to the body (or text) of the document.

SIMPLE, POWERFUL, EFFECTIVE, TEACHABLE

Simple, yet powerful, a field search that specifically targets the title or URL is not only an effective tool for your own use, but one that is well worth teaching to your students. Whether you choose to introduce them to the Advanced Search Page, or decide to teach the relatively simple syntax, your students will be pleased with this new tool. They might even be impressed with your search savvy. 🌈

REFERENCES

OneStat.com. “Less People Use 1 Word Phrase in Search Engines According to OneStat.com.” 24 July 2006. OneStat.com. 19 Nov. 2007. www.onestat.com/html/aboutus_pressbox45-search-phrases.html

William H. Weare, Jr., is the access services librarian at Valparaiso (Indiana) University, and can be reached at william.weare@valpo.edu.



What search syntax is needed to search within the title or URL of the document?

	Google	Yahoo!	Live Search	Ask
TITLE: Find One Term	intitle:term OR allintitle:term	intitle:term	intitle:term	intitle:term
TITLE: Find Two or More Terms	allintitle:term term term	intitle:term intitle:term intitle:term	intitle:term intitle:term intitle:term	intitle:term term term
URL: Find One Term	inurl:term OR allinurl:term	inurl:term	Does Not Work	inurl:term
URL: Find Two or More Terms	allinurl:term term term	inurl:term inurl:term inurl:term	Does Not Work	inurl:term term term

Notes: Do not put a space between the colon and the first term being searched for. It is not necessary to use the Boolean AND in a search, as AND is the default operator in the four major search engines.