



Advanced Search Syntax

User's Guide

HeinOnline uses a Lucene/SOLR search platform. While search relevancy algorithms take into account users who prefer to use more general search strategies, we recommend utilizing the search syntax and structure tips presented in this guide to achieve the best possible search results.

Terms

- Search queries are broken up into terms and operators. There are two types of terms:
 - Single: a single word such as "ethics" OR "discovery".
 - Phrase: a group of words surrounded by quotation marks such as "civil rights".
- Multiple terms can be combined together with Boolean operators to form a more complex query
- Terms are NOT case-sensitive.

Fields

- When performing a search, you can either specify a field or use fields provided in the advanced search tool. Field names and default fields are library specific.
- You can search any field by typing the field name, then a colon (:), then the search term.
- For example, let's assume an index contains two fields: title and text, with text being the default field. To search for a document titled "Civil Rights" which contains the phrase "racial discrimination", enter:

title: "Civil Rights" AND text: "racial discrimination"

or

title: "Civil Rights" AND "racial discrimination"
(text is the default field in this example)

- In the absence of quotation marks, the field is only valid for the term that it directly precedes, so the query title: Civil Rights will only find "Civil" in the title field. It will find "Rights" in the default field, which in this case is text.

Boolean Operators

- Boolean operators allow terms to be combined through logic operators. Boolean operators OR, AND, +, NOT, and - are supported. Boolean operators must be in all capital letters.

OR

The OR operator links two terms and finds a matching document if either of the terms exist within a document. This is equivalent to a union using sets.

To search for documents containing either the phrase "right to vote" or the word "suffrage" use the query: **"right to vote" OR suffrage**

Boolean Operators, continued

AND

The AND operator matches documents in which both terms exist anywhere in the text or metadata fields of a single document. The symbol && can be used in place of the word AND.

To search for documents with the title field containing the phrase "real property" and author/creator field containing "Rheinstein" use the query: **title: "real property" AND creator: Rheinstein**

+

The + (or required) operator dictates that the term after the + symbol MUST exist somewhere in a single document.

To search for documents that must contain "watershed" and may contain "planning" use the query: **+ watershed planning**

NOT

The NOT operator excludes documents which contain the term after NOT. The symbol ! can be used in place of the word NOT.

To search for documents which contain "real property" but not residential, use the query: **"real property" NOT residential**

-

The - (or prohibit) operator excludes documents which contain the term after the - symbol.

To search for documents which contain "watershed planning" but not "watershed system" use the query: **"watershed planning" - "watershed system"**

Range Searches

- Range queries allow you to match documents whose field(s) values are between the lower and upper boundaries specified by the range query. Range queries can be inclusive or exclusive of the upper and lower boundaries. Sorting is done lexicographically. Inclusive range queries are denoted by square brackets. Exclusive range queries are denoted by curly brackets.

date: [1934 TO 1944] will find documents whose date fields have values between 1938 and 1944, inclusive.

- Note that range queries are not reserved for date fields. You can also use range queries with non-date fields.

title: {Aida TO Carmen} will find all documents whose titles are between Aida and Carmen, but not including Aida and Carmen.

Term Modifiers

- These provide the ability to modify query terms to allow for a wide range of searching options.
- Wildcard searches: single and multiple-character wildcard searches are supported
 - Use the ? symbol to perform a single-character wildcard search
 - Use the * symbol to perform a multiple-character wildcard search
- The single-character wildcard search looks for terms matching the single character placed.

To search for "text" or "test" use the search: **te?t**

- The multiple-character wildcard search looks for zero or more characters.

To search for "test", "tests" or "tester" use the search: **test***

- You can also use the wildcard searches in the middle of a term: **te*t**
- You cannot use a * or ? symbol as the first character of a search term

NOTE: The current syntax does not support the use of a proximity search and a wildcard search in the same search string. That means you cannot do "**consumer product* safety standards**"~15.

Proximity Searches

- Proximity searches allow you to find words which are with a specific distance from one another. To conduct a proximity search, use the tilde symbol, ~, at the end of a phrase.

To search for "watershed" and "planning" within 10 words of each other in a document use the search: "**watershed planning**"~10

Grouping

- Use parentheses to group clauses to form sub-queries. This can be useful in controlling Boolean logic for a search query.

To search for either "watershed" or "water rights" and "planning" use the search:
(watershed OR "water rights") AND planning

- This eliminates confusion and makes sure that "planning" must exist, and either the term "watershed" OR "water rights" may exist.

Field Grouping

- Use parentheses to group multiple clauses to a single field

To search for a title that contains both the word "King" and the phrase "civil rights" use the search: **title: (+ King + "civil rights")**

Fuzzy Searches

- Use fuzzy searches, which are based on an “edit distance” algorithm, to search for terms similar in spelling to another term.

To search for a term similar in spelling to “roam” use the fuzzy search: **roam~**. This search will find terms like “foam” and “roams”.

- A similarity parameter can also be specified. The parameter value is between 0 and 1, and the closer the value is to 1, the higher the similarity will be. For example: **roam~0.8**. The default similarity parameter if not otherwise specified is 0.5.

Boosting a Term

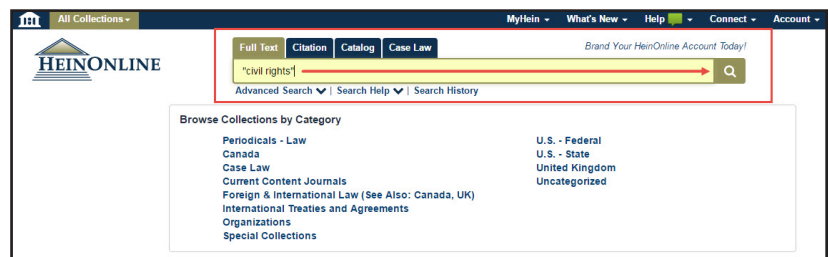
- The relevance level of matching documents based on terms found is built into the HeinOnline search engine. In order to boost a term, use the caret symbol, ^, with a boost factor at the end of the term you are searching. The higher the boost factor, the more relevant the boosted term will be.
- Boosting allows you to control the relevance of a document by boosting its term.

For example, if you are searching for “Iroquois Indian” and you want the term “Iroquois” to be more relevant than “Indian”, boost it using the caret symbol along with a boost factor number next to the term: **Iroquois^4 Indian**

- This search will make the term “Iroquois” more relevant.
- You can also boost phrase terms as in the example: **“Iroquois Indian”^4 “Indian Customs”**
- By default, the boost factor is 1.
- Although the boost factor must be positive, it can be less than 1 (for instance, 0.2).

One Box Searching vs. Advanced Search

- The search syntax discussed within this guide is best applied to the main search bar, located at the top of any page in HeinOnline.



- The advanced search tool provides specified fields and Boolean operator drop-down options to perform a simpler and/or more specific search.

