Terms

- A query is broken up into terms and operators. There are two types of terms:
  - Single: a single word such as "test" or "hello".
  - Phrase: a group of words surrounded by double quotes such as "hello dolly".
- Multiple terms can be combined together with Boolean operators to form a more complex query (see below).
- Terms are NOT case-sensitive.

Fields

- When performing a search, you can either specify a field or use the default field. The field names and default field used in HeinOnline are library specific and are listed on each library's search page.
- You can search any field by typing the field name followed by a colon ":" and then the term you are looking for.
- As an example, let's assume an index contains two fields-title and text-and text is the default field. If you want to find the document entitled "The Right Way" which contains the text "go this way", you can enter:

> title: "The Right Way" AND text: "go this way"
> **or**
> title: "The Right Way" AND "go this way"
> (*Since, in this example, text is the default field, the field indicator is not required*)

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

Boolean Operators

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

OR

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

> To search for documents that contain either the phrase "watershed planning" or the word "watershed" use the query: **"watershed planning" OR watershed**

AND

The AND operator matches documents where both terms exist anywhere in the text of a single document. (The symbol "&&" can be used in place of the word AND.)

> To search for documents whose title field contains "real property" and creator field contains "rheinstein" use the query: **title: "real property" AND creator: rheinstein**

Boolean Operators, Cont.

+

The "+" or required operator requires that the term after the "+" symbol exist somewhere in a field of 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 that contain the term after NOT. (The symbol "!" can be used in place of the word NOT.)

> To search for documents that contain "watershed planning" but not "watershed system" use the query: **"watershed planning" NOT "watershed system"**

The NOT operator cannot be used with just one term. For example, the following search will return no   results: NOT "watershed planning"

-

The "-" or prohibit operator excludes documents that contain the term after the "-" symbol.

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

Range Searches

- Range Queries allow one to match documents whose field(s) values are between the lower and upper bound specified by the Range Query. Range Queries can be inclusive or exclusive of the upper and lower bounds. Sorting is done lexicographically.

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

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

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

Advanced Search Syntax
G u i d e l i n e s

Term Modifiers

- Support for modifying query terms to provide a wide range of searching options.
- Wildcard searches: Support for single and multiple character wildcard searches.
  - To perform a single-character wildcard search use the "?" symbol.
  - To perform a multiple-character wildcard search use the "*" symbol.
- The single-character wildcard search looks for terms that match with the single character placed.

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

- Multiple-character wildcard searches looks for zero or more characters.

> To search for "test", "tests" or "tester" you can 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

- Support for finding words which are a within a specific distance away. To do a proximity search, use the tilde symbol, "~", at the end of a phrase.

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

Grouping

- Support for using parentheses to group clauses to form sub-queries. This can be very useful if you want to control the Boolean logic for a query.

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

This eliminates any confusion and makes sure that "planning" must exist and either term "watershed" or "water rights" may exist.

Field Grouping

- Support for using parentheses to group multiple clauses to a single field.

> To search for a title that contains both the word "return" and the phrase "pink panther" use the query: **title: (+ return + "pink panther")**

Fuzzy Searches

- Support for fuzzy searches based on the Levenshtein Distance, or Edit Distance, algorithm. To do a fuzzy search, use the tilde symbol, "~", at the end of a single-word 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.

- Optionally, you can also specify a similarity parameter. 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 the terms found is provided. To boost a term, use the caret symbol, "^", with a boost factor (i.e., a number) at the end of the term you are searching. The higher the boost factor, the more relevant the term will be.
- Boosting allows you to control the relevance of a document by boosting its term.

> For example, if you are searching for "jakarta apache" and you want the term "jakarta" to be more relevant than "apache", boost it using the caret symbol along with the boost factor next to the term.   **jakarta^4 apache**

- This will make documents with the term "jakarta" appear more relevant.
- You can also boost phrase terms as in the example: **"jakarta apache"^4 "Apache Lucene"**
- By default, the boost factor is 1.
- Although the boost factor must be positive, it can be less than 1 (e.g., 0.2).