

DSL Latin Text Search Engine: University of Oklahoma Libraries

Full-text Search Tips

Elasticsearch is a robust search engine with full-text searching capability. Along with standard keyword searches, this application allows exact phrase, proximity, partial word, and filter options.

Using the keywords *civitatis multos* as an example, here are the different ways to control the search

- **OR:** By default, when searching for two or more words, the application searches for all of the words but not necessarily together. Searches **civitatis OR multos** Search results = 898 texts.
- **AND:** To search both words together, click **Match (AND)** below the search box. Searches **civitatis AND multos** Search results = 28 texts.
- **Match Phrase:** Below the search box, click Match Phrase to search the exact phrase and order of the words. Searches the precise phrase “**civitatis multos**” Search results = 1 text. **Proximity Searches:** While a Match Phrase search expects all of the words in exactly the same order, a proximity search allows words to be further apart or in a different order and can specify the maximum distance of words in a phrase. Searching “sea level rising” = “sea level rising” only Searching “sea level rising” ~5 = “rising sea level” or “sea level is rising” or “rising water levels above sea” **Partial Word Searches:** The wildcard character (*) represents any number of characters to allow partial word searching. Searching *arm** = *armis*, *arma* **Filter Options:** Click **Query String** and **Filter Options** button to refine your search further.
- **Filter Options: Era, Author, Familiar Name** - By default all three are searched, this filter allows limiting the search to just one or two
- **Stop Words** - Eliminates words in the search query. For example: When searching *arma*, add the stop word *quae* to avoid getting *arma quae* in the results.
- **Search within Fields** - To search only within a field type, such as author or title, add the field name and search command directly to the search box.

TITLE: Sea Level = keywords in the Title field only DATA: "lisa murkowski" ~5 AND 2015 = keywords in DATA field only, exact phrase within 5 lines, in 2015

Search Results Ranking:

Elasticsearch sorts matching results by their relevance score, that is, by how well each document matches the query. A keyword search for *et voce* (et OR voce) will rank documents that have both words together at the top of the search results list, followed by documents that have one word or the other word in the document.

Example: Search *et voce* (et OR voce) Documents at the top of search results list will contain et and voce together, followed by documents with either et or voce in the text.

Conducting Statistical Analysis

With a user account it is possible to collect your search results for statistical analysis. Items in the results list can be selected individually or by page. Steps to export data:

Saving results in CSV format

1. Conduct search
2. Select items (individually or SELECT all on page)
3. Click Save to CSV

To Collect data via

1. Conduct search
2. Select items (individually or SELECT all on page)
3. Click Statistics button at top right of page
4. Click General Stats
5. Click on link that appears above search results list