SOLR Keyword Spotting | API

This search is only possible if the HTR has been post-processed (typically by UPVLC, contact info@readcoop.eu for questions)

Searching for keywords via the SOLR index can be done via GET request to

```
https://transkribus.eu/TrpServer/rest/keyword
```

with the following parameters:

- `query` string – the keyword to be searched

- `start` int (default: 0) – first result

- `rows` int (default: 10) – number of successive results to fetch
  - In order to process large amounts of hits, SOLR allows to define at a specific hit and show only the next N hits from there onward. This can be used to browse results page-wise (e.g. first page starts at 0 and shows 10 results, next page starts at 11 and shows next 10 etc.)

- `probL` float – lower limit for keyword probability (usually between 0.0 and 1.0)

- `probL` float – upper limit for keyword probability (usually 1.0)
  - Each keyword is stored with a probability value. It is possible to limit searches to results above or below a certain probability. (Note: Currently, the keyword probabilities are stored directly as provided. To transform these probabilities into true relevance probabilities, a calibration function is required in the user interface.)

- `filter` string – allows to specify certain fields and values to filter search results (can take multiple values as in …&filter=cId:1895&filter=id:4243_221_*…)
  - fields to filter by are
  - `id` : (string) index element id, consisting of document id, page number and a running number for word on the page, separated by underscores -> e.g. 4432_15_10 would be word 10 on page 15 of document 4432. Setting a filter string to 4432_15_* would limit searches to this document and page; *_20_* would limit searches to page 20 of any document.
  - `title` : (string) title of the document
  - `cId` : (int) collection id
  - `auth` : (string) name of the author

- `fuzzy` : int – takes all integer values, but SOLR currently only supports values between 0 and 2
  - SOLR allows to include results that differ in a certain amount of characters.

- `sorting` string – allows to sort by certain fields. (usually "rp desc" to show results with descending probability)

Example:

Searching for the keyword "london" in collection 1234 with any probability, displaying the first 100 results sorted by descending probability.

```
https://transkribus.eu/TrpServerTesting/rest/search/keyword?query=london&start=0&rows=100&probL=0.0&probH=1.0&filter=cId:1234&fuzzy=0&sorting=rp+desc
```