Extract Documents by Year for Burst Detection

Menu Path

Data Preparation > Database > ISI > Extract Documents By Year For Burst Detection

Description

This algorithm extracts all of the document titles, abstract texts, full texts, and publication years included in the ISI dataset. It produces a table designed to be processed by the burst detection algorithm. This table consists of the following columns:

- *Title*: If provided, the title of the document.
- *Abstract*: If provided, the abstract text of the document.
- *Title and Abstract*: If provided, the title and the abstract text of the document combined into one field (in case you want to detect bursts over both).
- *Year*: The year this document was published.

Usage Hints

Load an ISI file into the tool, then create a database from it using the ISI database loader.

Given your ISI database, run this extraction on it. Normalize the text of the results on the column you are detecting bursts on (*Title*, *Abstract*, or *Title and Abstract*). Run burst detection on the results of the normalized text.

Implementation Details

The specific query run by the tool can be found in the source code.

Links

- Load ISI File into Database
- Normalize Text?
- Burst Detection

See Also

⊘ The license could not be verified: License Certificate has expired! Generate a Free license now.