

Extract Document Citation Network (Core Only)

Menu path

Data Preparation > Database > ISI > Extract Document Citation Network (Core Only)

Description

Extracts the Document citation network from an ISI database.

Each Document and each referenced Document in the input database is represented by a node. An edge is drawn between the nodes for two Documents if and only if one of the Documents cited the other Document.

Core Document vs. Non-Core Document

There is a distinction drawn between Documents contained in your dataset and Documents in general. A Document in your dataset is called a "Core Document". Your Documents may (and probably do) reference Non-Core Documents.

The output network of this algorithm will contain nodes representing only Core Documents. For an algorithm that will represent Non-Core Documents too, see [Extract Document Citation Network](#).

Analyses

The output network will include the following data and metadata:

- Node (Document)
 - All actual data from the [Documents Table.DOCUMENT](#)
 - A generated prettified label for identifying this Document.
 - A generated prettified string giving the Source of this Document (called 'SOURCE').
- Edge (Citation)
 - Currently, no metadata is provided on edges.

Usage Hints

Load an ISI file into the tool, then create a database from it using [the ISI database loader](#).

It is strongly recommended that the database be cleaned before extracting any co-citation networks from it.

For a quick analysis of a small dataset you may wish to [merge together author entities with identical names](#). For a scientifically sound analysis of a larger dataset, you can [find author entity merging suggestions](#) (or [manually set your own merging orders from scratch](#)) and [perform the merge](#).

Then, you will probably want to [merge together journal entities according to recognized variants](#).

Finally, you must [match references up to documents in your dataset](#) (there are no citations to analyze, otherwise).

Implementation Details

The specific query run by the tool can be found in the [source code](#).