WE1S "metadata" module

A component of the WE1S Workspace (see S-2), the "metadata" module allows users to manage and analyze the metadata of text collections in a project in the Workspace. Metadata include bibliographic information imported as part of the text data (e.g., author, publication source, date). But they can also include additional citational information as well as other fields that a researcher might add to address research questions (e.g., fields for the geographical region or media form of a publication source),

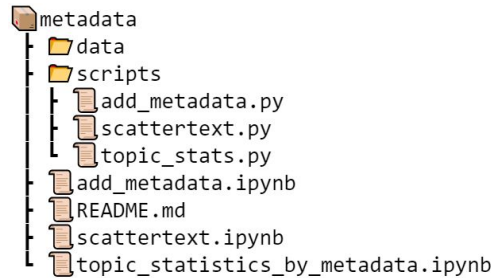The module includes three Jupyter notebooks (see M-15 for an explanation of data notebooks):

`add_metadata.ipynb`: Adds metadata to a project's JSON files from a CSV file after they have been imported into the project.

`scattertext.ipynb`: Uses Jason Kessler's Scattertext library to explore key terms in a text collection in relation to its metadata.[1]
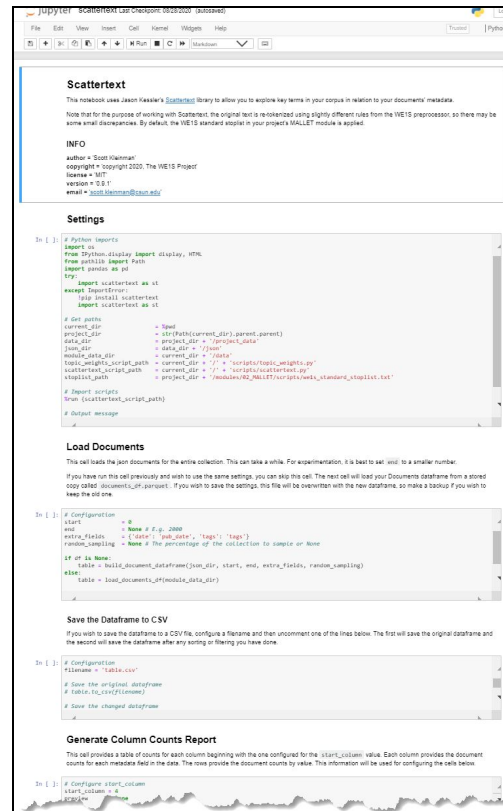
`topic_statistics_by_metadata.ipynb`: Extracts information from a topic model of a text collection and combines it with information in the collection's metadata to provide counts of the number of documents associated with specific metadata fields. (To run this notebook, a user must have previously run WE1S's "topic_modeling" module; see S-13).

---

[1] Because Scattertext is under active development and has many more functions than made available through this notebook, WE1S does not distribute it as part of its Workspace. It is downloaded and installed by the user when running the notebook's first cell.

Module structure

Screenshot from `scattertext.ipynb` (larger)

Further Information:

* M-15 (about Jupyter notebooks)
* Jason Kessler's Scattertext

Main Jupyter notebooks in this module:

* `add_metada.ipynb`
* `scattertext.ipynb`
* `topic_statistics_by_metadata.ipynb`

Code source: [TBD] (MIT License)