

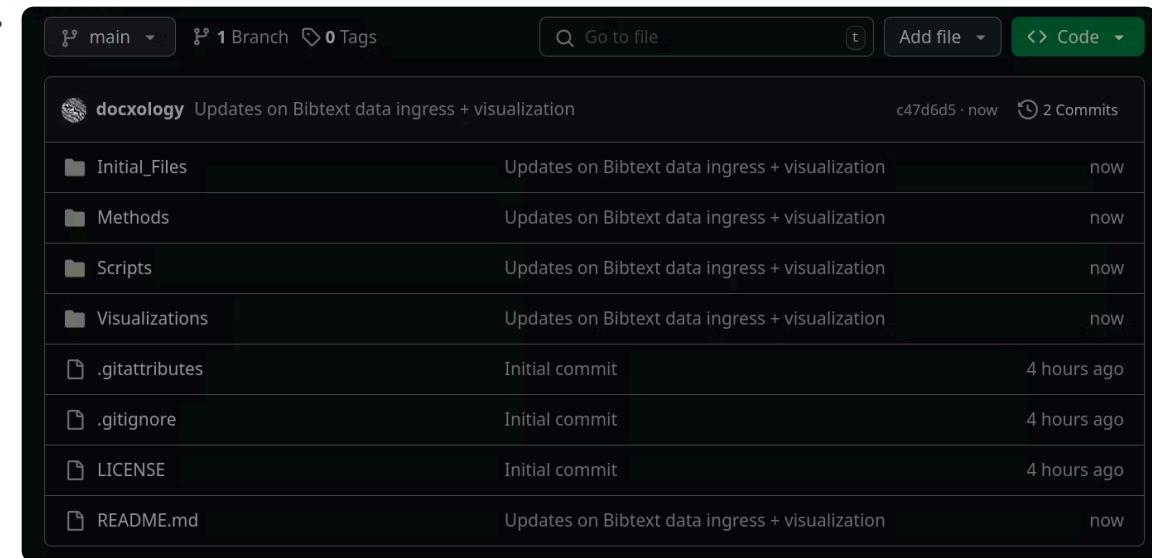
FORMINDEX_10-08-2024

October 8, 2024

Daniel Friedman

Situation & Overview

- Sanford Porter sent via email "5 different export files of the same 300 or so citations from "Porter or Wojcik"."
- The goal of this preliminary milestone, was to explore which bibliographic format I can use for further analyses.
- I decided to initially work with the [Bibtex](#) format because it is clear and plain text.
- All code and files are in a Github repo (currently private, request access via email if you want it):
<https://github.com/docxology/FORMINDEX> .



A screenshot of a GitHub repository page for 'docxology'. The repository has 1 branch and 0 tags. The commit history shows 2 commits. The commits are:

File	Message	Time
Initial_Files	Updates on Bibtext data ingress + visualization	now
Methods	Updates on Bibtext data ingress + visualization	now
Scripts	Updates on Bibtext data ingress + visualization	now
Visualizations	Updates on Bibtext data ingress + visualization	now
.gitattributes	Initial commit	4 hours ago
.gitignore	Initial commit	4 hours ago
LICENSE	Initial commit	4 hours ago
README.md	Updates on Bibtext data ingress + visualization	now

Data ingress

- Here are snippets of terminal logs, showing that the Bibtex format is legible.

```
$ python3 Visualize_FORMIS.py
Python executable: /home/trim/Documents/GitHub/FORMINDEX/Scripts/venv/bin/python3
Python version: 3.11.2 (main, Aug 26 2024, 07:20:54) [GCC 12.2.0]
sys.path: ['/home/trim/.local/lib/python3.11/site-packages', '/home/trim/Documents/GitHub/FORMINDEX/Scripts', '/home/trim/Documents/GitHub/FORMINDEX/Scripts/venv/lib/python3.11/site-packages/setuptools/_vendor', '/home/trim/Documents/GitHub/FORMINDEX', '/home/trim/Documents/GitHub/FORMINDEX/Scripts/venv/lib/python3.11/site-packages/bibparser']
bibparser successfully imported
Loaded 316 BibTeX records.
BibTeX resource counts:
phdthesis: 4
techreport: 85
incollection: 28
article: 197
book: 2
```

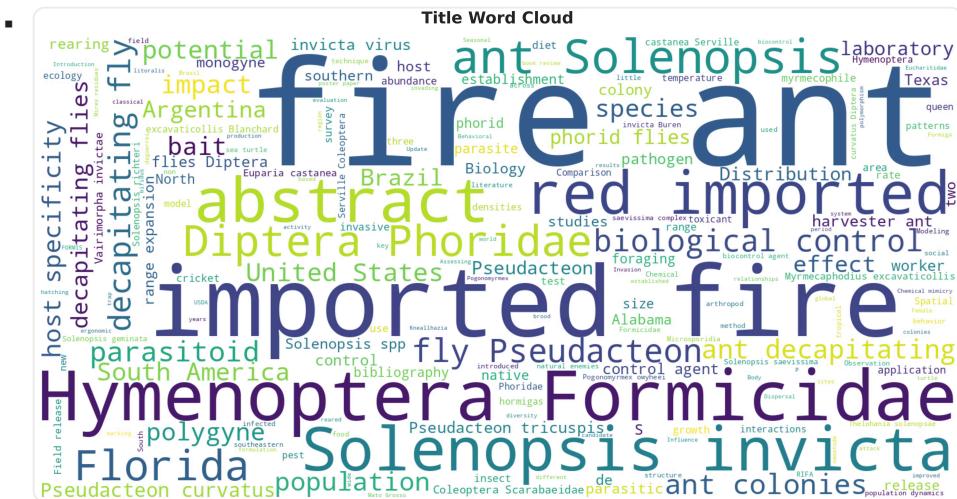
- This shows that we are successfully reading in the 316 records of several different types.

Preliminary Visualizations

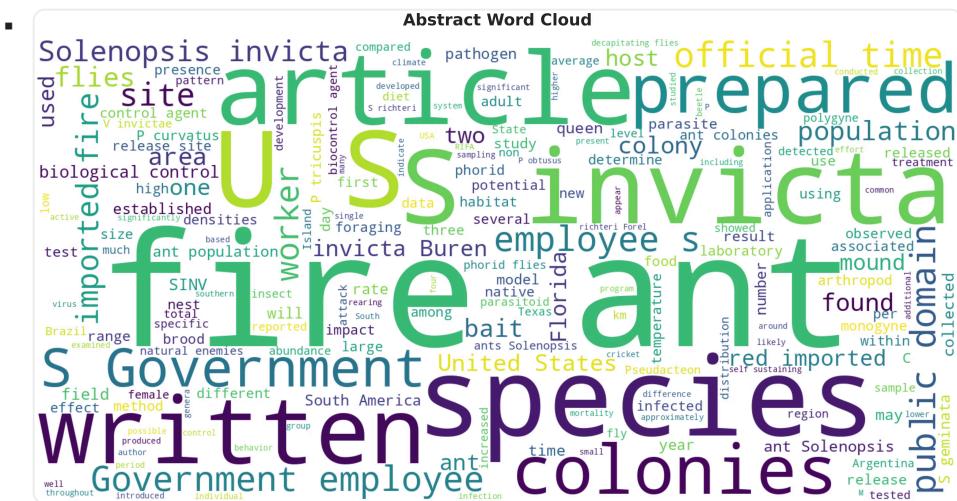
- These are just examples of simple visualizations which can be scaled and customized from here.

- Word clouds

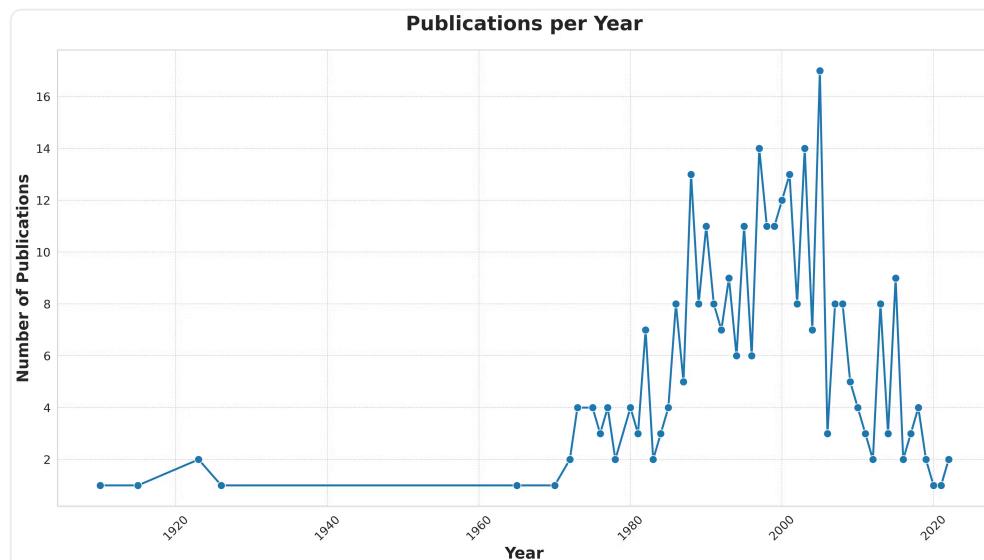
- Titles



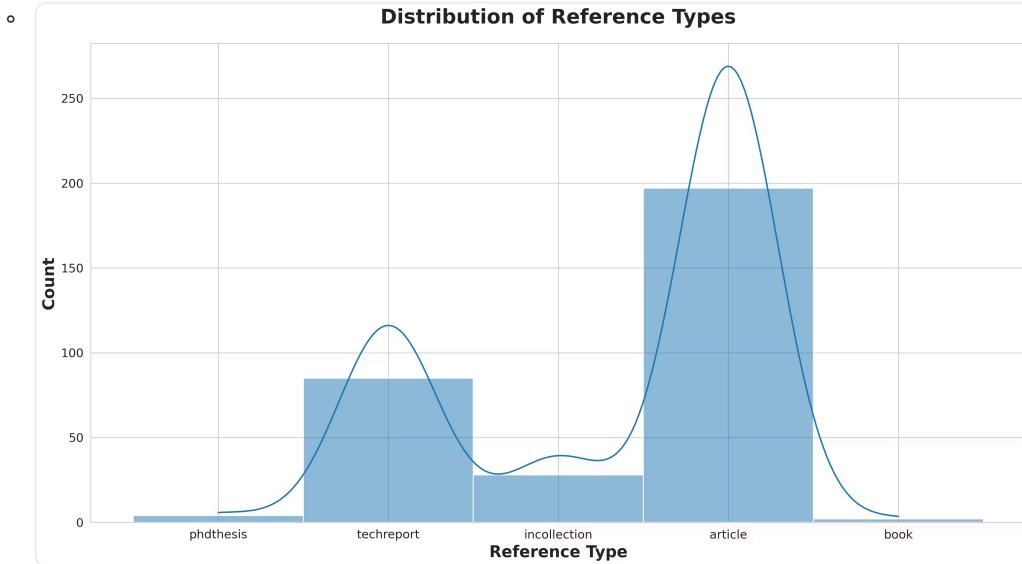
◦ Abstracts



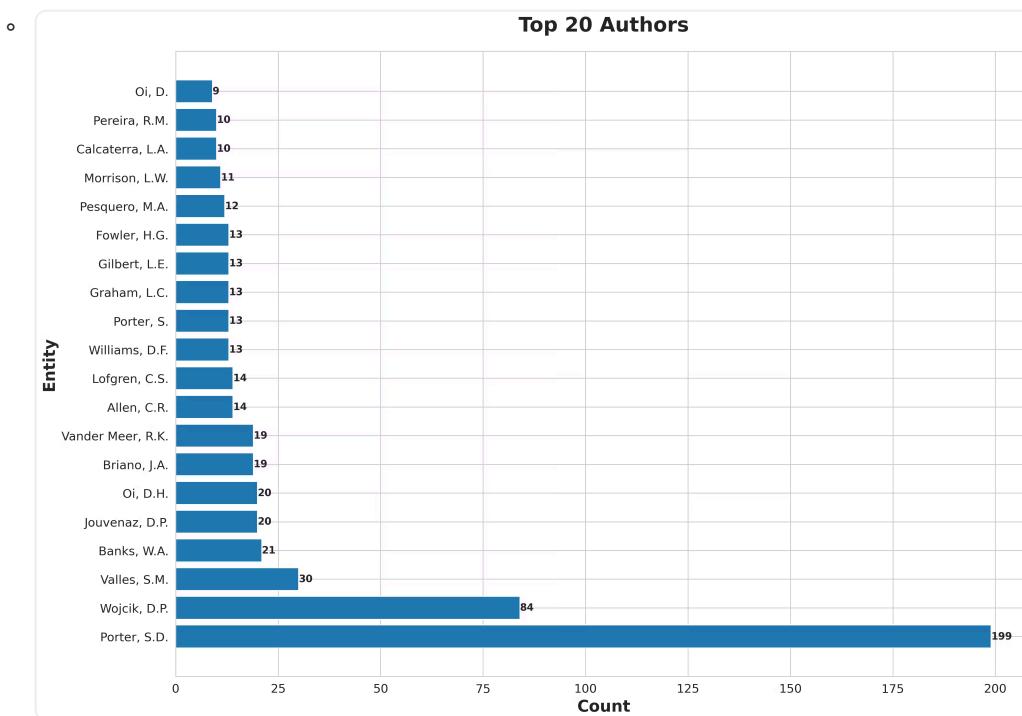
- Publications per year



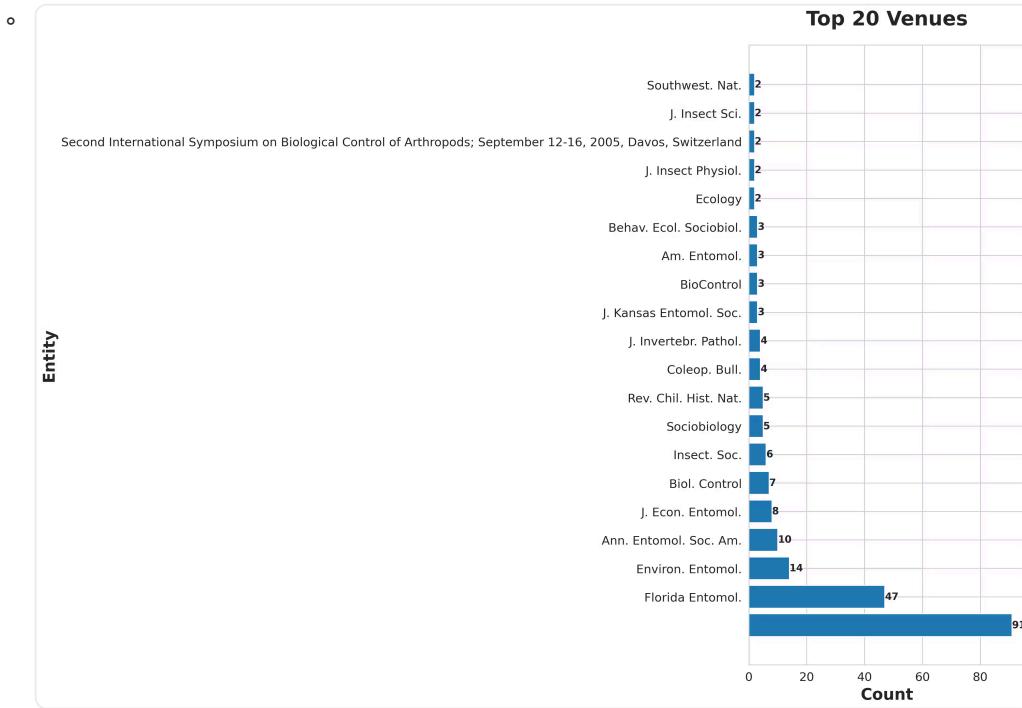
- Types of References



- Top Authors



- Top venues



Next steps

- Get full Bibtex export from Sanford.
- Ingress full FORMIS database from Bibtex and perform initial descriptive analysis.

..---.----.-.---..--.---..-