LLMs for Citation Management

Aerith Netzer

Digital Publishing and Repository Librarian

Northwestern University Libraries

aerith.netzer@northwestern.edu
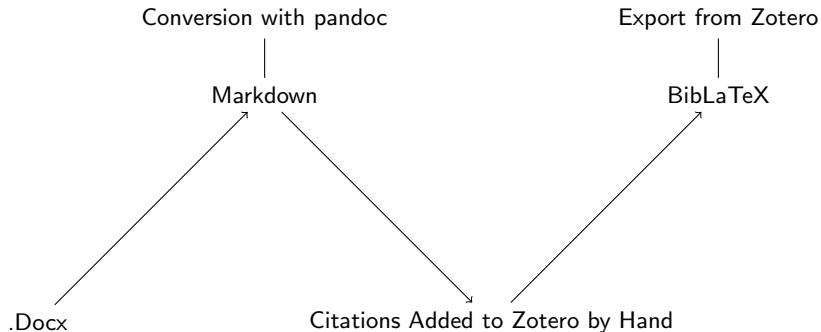
December 18, 2023

This Presentation's LaTeX File is Available on GitHub

- `https://www.github.com/aerithnetzer/OpenAI_Presentation`

The Code Demonstrated in this Presentation is Available on GitHub

- `https://github.com/aerithnetzer/py-bats`

Northwestern

Conversion with pandoc

Export from Zotero

Markdown

BibLaTeX

.Docx

Citations Added to Zotero by Hand

The Current State is Not Scalable

- The current process is time-consuming
- The current process is error-prone
- The current process is not scalable

The Problem of Plaintext Citations

- Plaintext citations are not machine-readable
- Many different citation styles cause confusion
- Citations are often not linked to the cited work

The Suboptimal Solution: OpenAI GPT-3.5

- Parse citations from article submission
- Create a .txt file with the citations in plaintext
- Call the OpenAI API to generate a biblatex file

The Optimal Solution: BibLaTeX Evangelising

- Create resources that help create .bib files
- Communicate the importance of BibLaTeX to researchers
- Create a .tex file with the citations in BibLaTeX