ReadMe File / Updating Instructions:

Contact: Sophie Kemprecos | skemprecos@gmail.com

Current Executions:

1. Extracted data from desktop version of Zotero in .csv format titled “FUS-IO PRIVATE”.

This was for the FUS-IO Annotated Collection, saving to a local folder titled “Tableau”.

1. Using Python wrote a script to help clean data and extract keywords into a readable and countable format for Tableau.

First section of CleanData.py deals with the modalities, second section handles the indications. Must change all file path’s to local directory of the extracted data from step 1. Possible to update keywords by simply entering new ones in the “keywords” list.

1. First section of CleanData.py creates FUS-IO\_expanded.csv which is used directly in the Tableau formatting. Second section builds on the FUS-IO\_expanded.csv by repeating the modality process with the indications to create a final copy with both changes as merged\_changes.csv
2. A third section creates the pub\_process.csv that identifies publication types, similar to the graphs that are available on the landscape analysis.
3. Uploaded the FUS-IO\_expanded.csv and merged\_data.csv to Tableau and began data visualization work.
   1. FUS-IO\_expanded.csv used modality column to create bubble map
   2. Merged\_data.csv and FUS-IO\_expanded.csv for other two features

For updating:

1. Download modified Zotero library and continue through same steps, just needing to change local directory
2. New list of modalities: in CleanData.py add new modality to “keywords” list, as needed add into the dictionary “normalization\_mappings”.
3. New list of indications: in CleanData.py add new modality to “tumor\_keywords” list

Troubleshooting:

If there are one or two unnormalized tags, you can manually change the cell in Excel, Numbers, etc.

Instructions are in CleanData.py on how the program works, has areas flagged for development

Tableau Dashboard Creation and Updating:

Merged\_data.csv was the main data used in the current dashboard, mainly due to the publication tagging being a final addition. Can use the pub\_process for every visualization for simplicities sake.

1. To update with new keywords/papers, simply download the library as a .csv file and run CleanData.py through to extract keyword data.
2. Using Tableau Public (free version), upload “pub\_process.csv” as a new data source.
3. Create three separate sheets, one for a bubble map, bar graph, and chart.
4. Load data in as desired, can just use same layout that is already part of the dashboard.