

NLP Suite

Roberto Franzosi

Emory University

6/14/2020

Table of Content

Mission statement & target audience	3
GNU license agreement	3
Introduction to the suite of NLP tools.....	3
Suite architecture	3
_main files.....	3
_GUI files.....	3
_util files	4
Filename suffixes and prefixes	4
What the NLP GUI can do: An overview	4
TIPS files	6
? HELP and ReadMe messages	7
Reminder messages.....	8
Specialized GUIs	8
Language tools	8
Stanford CoreNLP	8
CoNLL Table search.....	9
CoNLL table & clausal analysis	10
CoNLL table & noun/verb analysis	10
CoNLL table & function words	11
N-grams and co-occurrence viewer	11
Nominalization.....	12
Extracting and visualizing 4 of the 5 Ws of journalism: Who, What, When, Where.....	12
The “shape of stories”	13
Sentiment and concreteness analysis	13
Topic modeling (Gensim & Mallet)	14
Word similarity	15

WordNet.....	15
Style measures: A summary GUI	16
Visualization tools	16
Word clouds	16
Excel charts	17
Geocoding and visualizing spatial data in Google Earth Pro	17
Statistical tools	18
A set of comprehensive tools for social science research.....	18
File and data management tools.....	19
Filename checker	19
File content checker & converter.....	20
File merger & splitter.....	20
File classifier (dumb classifier).....	21
File manager.....	21
Data manager	22
Knowledge bases and annotator tools (DBpedia & YAGO, dictionaries).....	22
Work ahead: What needs to be done.....	23
How to run the NLP suite	24
Download and install NLP Suite.zip.....	24
Download and install Python and Java	26
Install Python libraries	27
Download and install Stanford CoreNLP: A language parser and more	27
Download and install Gephi: A network software.....	27
Download and install Google Earth Pro: A Geographic Information System (GIS) software..	27
Run NLP.py	27

Mission statement & target audience

In an age of BIG DATA, the purpose of the suite is rather to provide humanists and social scientists a wide range of **computational tools for the analysis and visualization of smaller datasets**, the more typical datasets humanists and social scientists use (e.g., the works of one Nobel Prize winner, a handful of in-depth interviews, a few thousand newspaper articles).

Furthermore, the NLP suite is designed for non-specialists, for scholars with **no knowledge or little knowledge of Natural Language Processing**.

GNU license agreement

The suite of Java and Python NLP (Natural Language Processing) tools made available here have been developed by Roberto Franzosi and a team of collaborators at Emory University. The Suite is distributed freely under a **GNU License Agreement** (<https://www.gnu.org/licenses/gpl-3.0.en.html>).

Introduction to the suite of NLP tools

The Suite is expanding in functions and help files on a weekly basis. The current release consists of

1. some 200 Python 3 files with several functions in each,
2. 10 Java files,
3. some 50 TIPS (help) files.

The TIPS complement, with far more extensive explanations, the online HELP, ReadMe, and reminder messages available for each script.

Suite architecture

The Python scripts in the NLP Suite have filenames that clearly identify the Suite architecture. Filename suffixes designate three different types of files: main, GUI, and util.

_main files

1. _main files are the only ones that you can run in command line independently of others; they will call both GUI and util files.

_GUI files

The NLP suite provides a set of user-friendly GUIs (Graphical User Interface) scripts to make complex tasks available to non-specialists. Each GUI comes with a set of ?HELP buttons that provide minimal help for a specific line in the GUI, a ReadMe button for overall help on the tools available for the GUI, a set of pdf TIPS files for detailed help, and reminders for the novice. A total of some 50 (and growing) TIPS files accompany the NLP suite.

GUI files lay out on the screen the widgets of a specific script for easy Graphical User Interface (GUI).

_util files

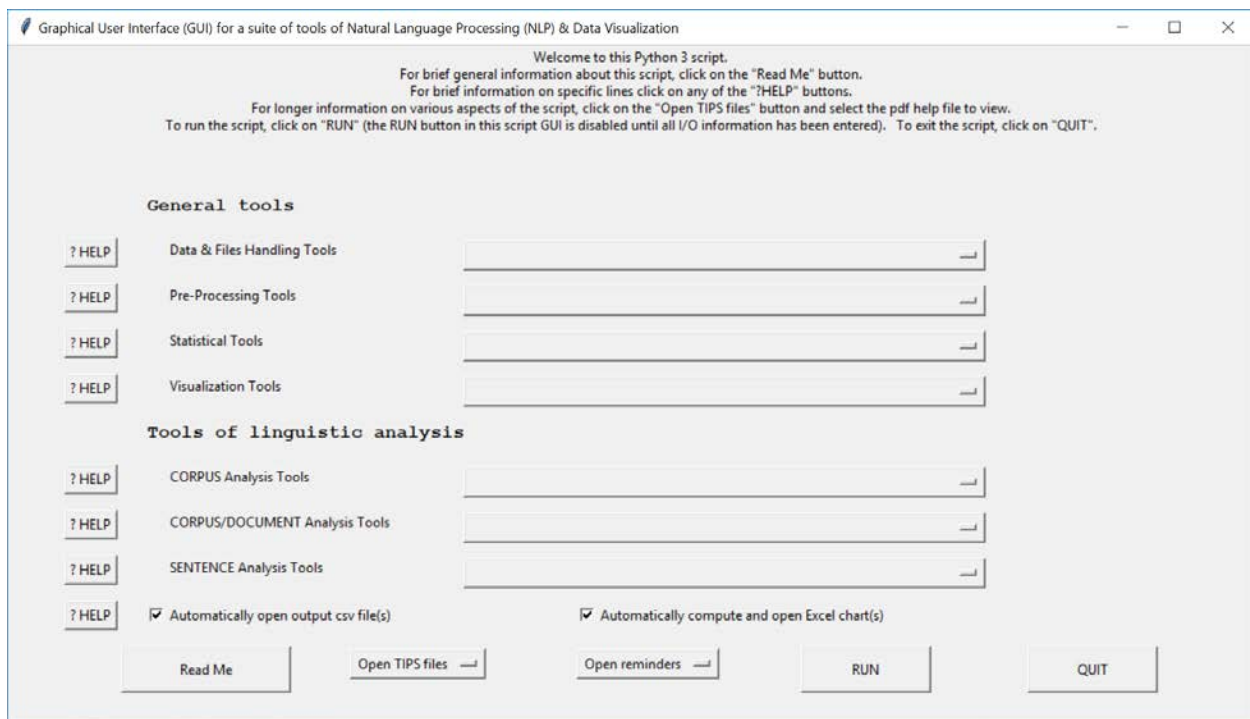
The _util files do all the computational heavy lifting behind each _main and _GUI script.

Filename suffixes and prefixes

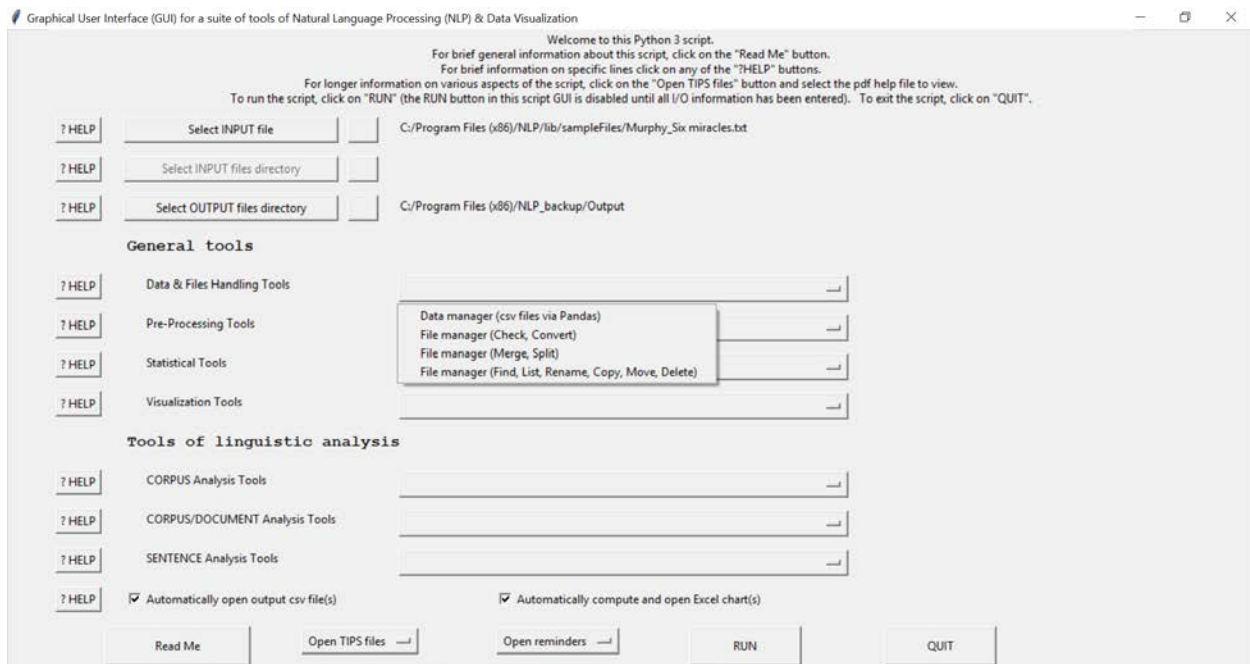
ALL SCRIPTS SUFFIXED BY main CAN BE RUN INDIPENDENTLY OF THE NLP_main.py. Thus on command line you can type, for instance, Python annotation_main.py and it will fire up the annotator_GUI.py independently of NLP_main.py. Filename prefixes cluster together scripts used for the same purpose. Thus annotator identifies all files dealing with html annotation.

What the NLP GUI can do: An overview

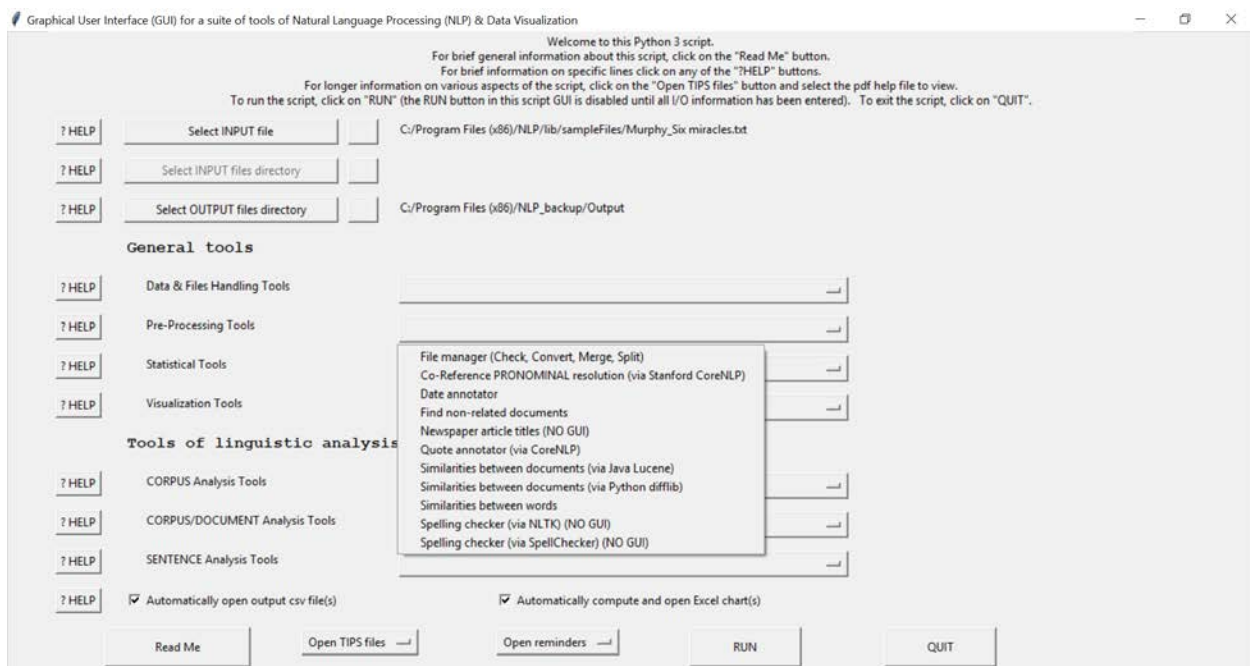
When you type Python NLP_main.py in command line, this is what you see. A GUI (Graphical User Interface) that provides access to all tools available in the NLP Suite.



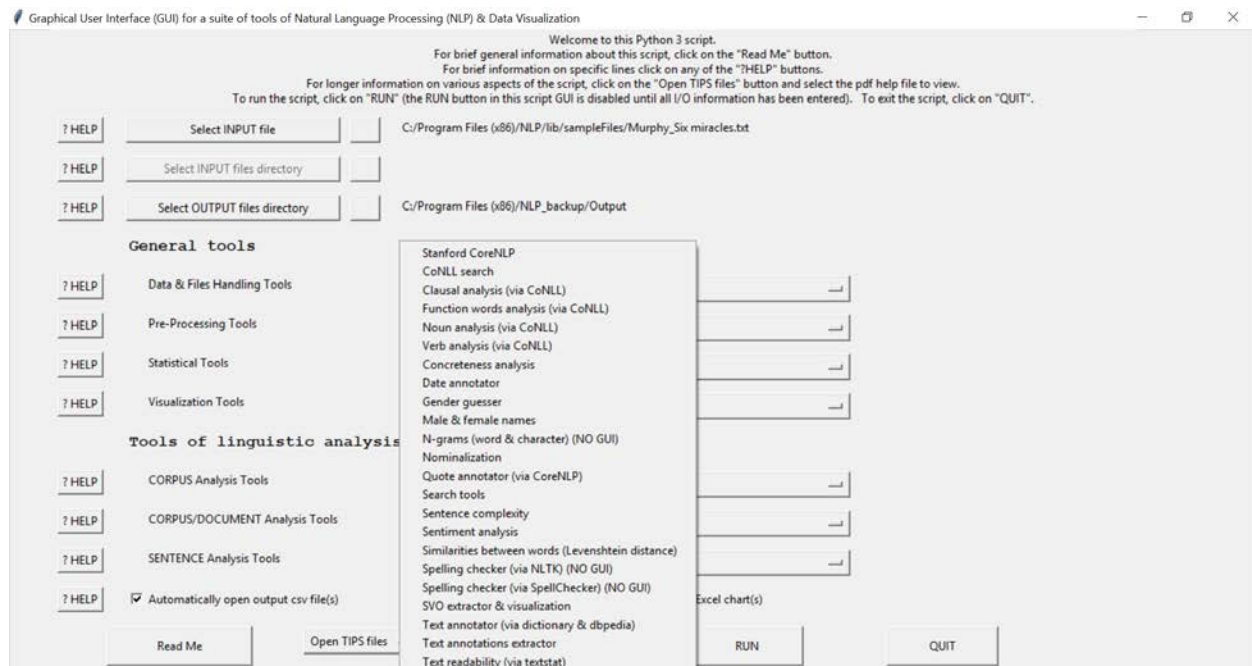
Each menu widget opens a set of more specific tools (for example, for data and file handling).



Or for pre-processing documents.

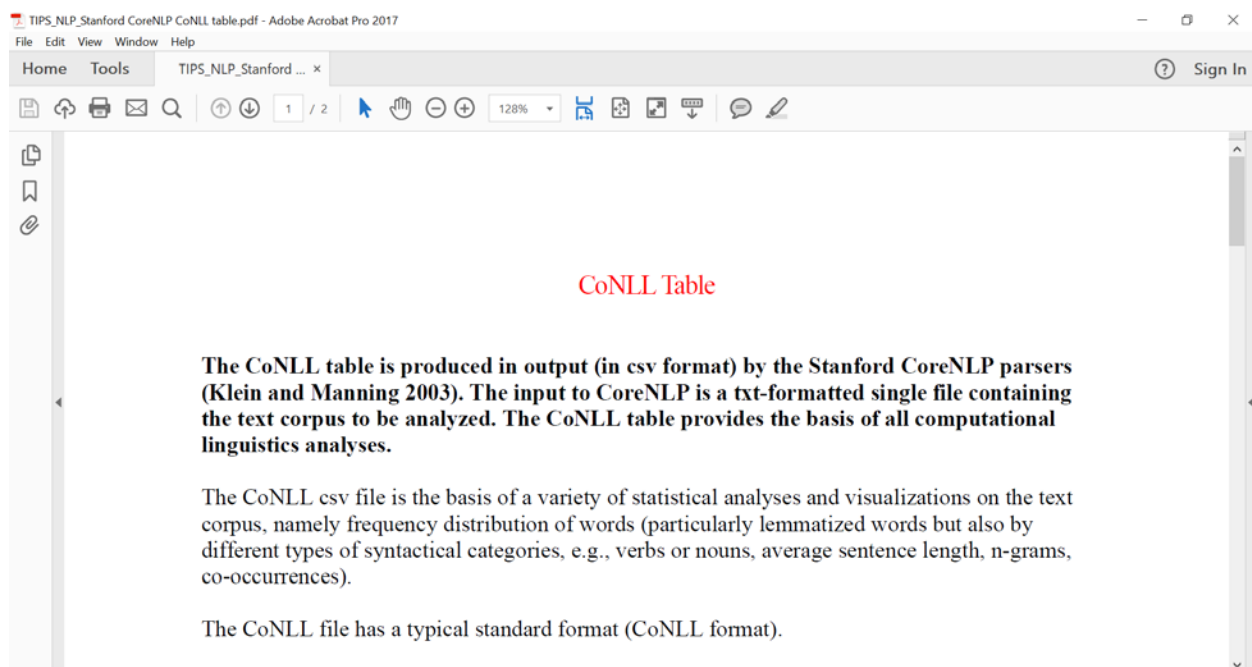


Or for analyzing a corpus of documents or a single document.

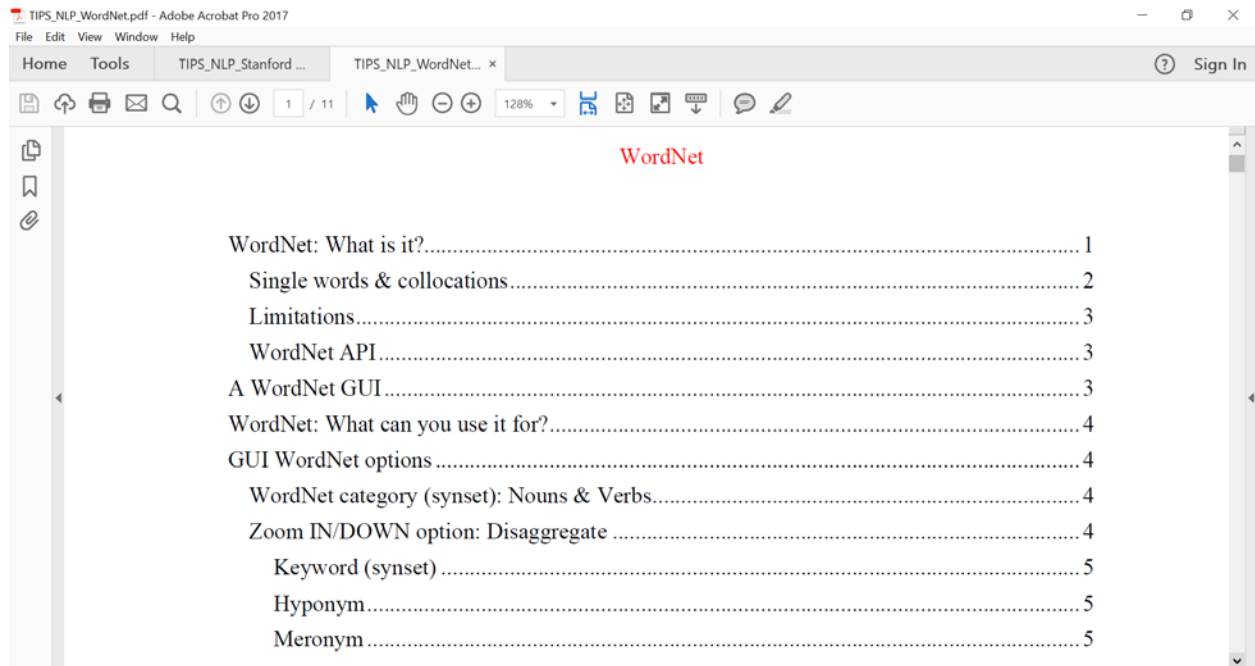


TIPS files

And here is a sample of a TIPS file, opened directly from the NLP Suite.

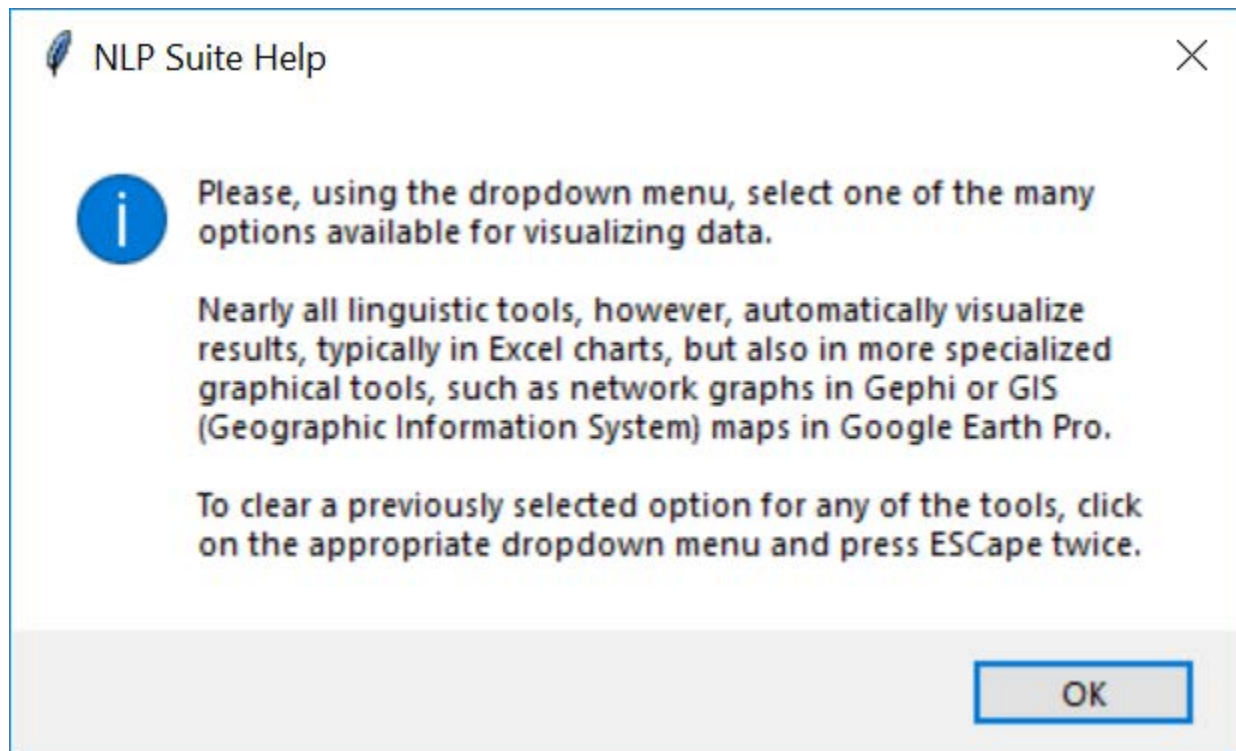


Some TIPS have more elaborate structures complete with Table of Contents.



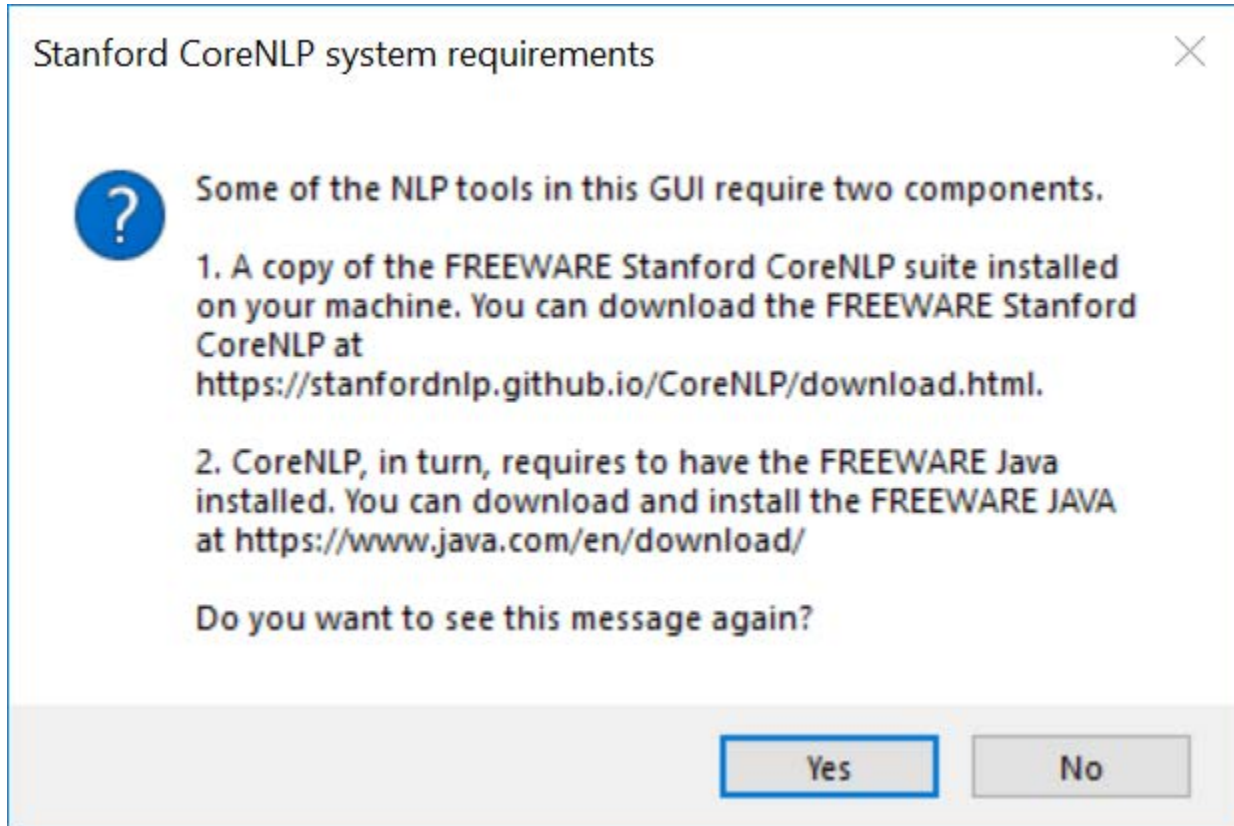
? *HELP* and *ReadMe* messages

This is a typical help message for a ? HELP button.



Reminder messages

And here is a typical reminder you get when you fire up the Stanford CoreNLP scripts.

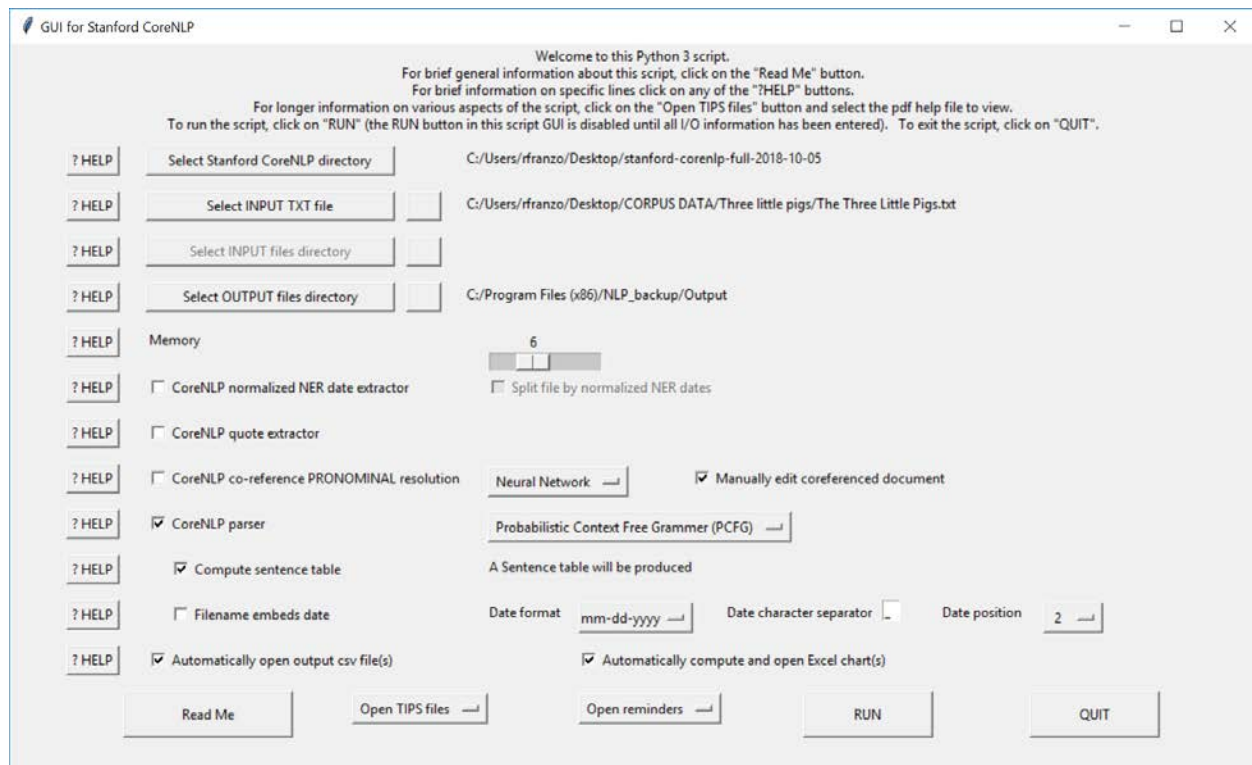


Specialized GUIs

You can run the tools in the NLP Suite via the `NLP_main.py`, or you can also run each separately, for as long as it is a `_main` type file. And of the core tools is the `Stanford_CoreNLP_main`. It is the engine upon whose output rely many of the other tools. Let's take a closer look at the Stanford CoreNLP GUI.

Language tools

Stanford CoreNLP



The Stanford CoreNLP (<https://stanfordnlp.github.io/CoreNLP/>) provides a wide range of open source, freeware natural language processing tools (originally in Java and more recently also in Python). The parser is at the core of their tools; it produces the CoNLL table used for many subsequent analyses.

CoNLL Table search

GUI for CoNLL Search

Welcome to this Python 3 script.
 For brief general information about this script, click on the "Read Me" button.
 For brief information on specific lines click on any of the "?HELP" buttons.
 For longer information on various aspects of the script, click on the "Open TIPS files" button and select the pdf help file to view.
 To run the script, click on "RUN" (the RUN button in this script GUI is disabled until all I/O information has been entered). To exit the script, click on "QUIT".

? HELP Select INPUT CoNLL table

? HELP Select OUTPUT files directory

? HELP Searched token

? HELP CoNLL search field

? HELP POSTAG of searched token

? HELP DEPREL of searched token

? HELP POSTAG of co-occurring tokens

? HELP DEPREL of co-occurring tokens

? HELP ☒ Automatically open output csv file(s) ☒ Automatically compute and open Excel chart(s)

Read Me Open TIPS files Open reminders RUN QUIT

CoNLL table & clausal analysis

GUI for Clausal Analysis

Welcome to this Python 3 script.
 For brief general information about this script, click on the "Read Me" button.
 For brief information on specific lines click on any of the "?HELP" buttons.
 For longer information on various aspects of the script, click on the "Open TIPS files" button and select the pdf help file to view.
 To run the script, click on "RUN" (the RUN button in this script GUI is disabled until all I/O information has been entered). To exit the script, click on "QUIT".

? HELP Select INPUT file

? HELP Select OUTPUT files directory

? HELP ☒ Clausal analysis

? HELP ☐ Visualize sentence structure (via dependency tree)

? HELP ☐ Sentence complexity

? HELP ☐ Text readability

? HELP ☒ Automatically open output csv file(s) ☒ Automatically compute and open Excel chart(s)

Read Me Open TIPS files Open reminders RUN QUIT

CoNLL table & noun/verb analysis

GUI for Noun and Verb Analysis

Welcome to this Python 3 script.
 For brief general information about this script, click on the "Read Me" button.
 For brief information on specific lines click on any of the "?HELP" buttons.
 For longer information on various aspects of the script, click on the "Open TIPS files" button and select the pdf help file to view.
 To run the script, click on "RUN" (the RUN button in this script GUI is disabled until all I/O information has been entered). To exit the script, click on "QUIT".

? HELP Select INPUT CoNLL table C:/Program Files (x86)/NLP_backup/Output/CoNLL/NLP_SCNLP_Murphy Miracles thicker than fog CORENLP_CoNLL.csv

? HELP Select OUTPUT files directory C:/Program Files (x86)/NLP_backup/Output

? HELP ☒ Noun Analysis

? HELP ☒ Verb Modality Analysis

? HELP ☒ Verb Tense Analysis

? HELP ☒ Verb Voice Analysis

? HELP ☒ Automatically open output csv file(s) ☒ Automatically compute and open Excel chart(s)

Read Me Open TIPS files Open reminders RUN QUIT

CoNLL table & function words

Function Words Analysis

Welcome to this Python 3 script.
 For brief general information about this script, click on the "Read Me" button.
 For brief information on specific lines click on any of the "?HELP" buttons.
 For longer information on various aspects of the script, click on the "Open TIPS files" button and select the pdf help file to view.
 To run the script, click on "RUN" (the RUN button in this script GUI is disabled until all I/O information has been entered). To exit the script, click on "QUIT".

? HELP Select INPUT CoNLL table C:/Program Files (x86)/NLP_backup/Output/CoNLL/NLP_SCNLP_Murphy Miracles thicker than fog CORENLP_CoNLL.csv

? HELP Select OUTPUT files directory C:/Program Files (x86)/NLP_backup/Output

? HELP ☒ Article Analysis

? HELP ☒ Auxiliary Analysis

? HELP ☒ Conjunction Analysis

? HELP ☒ Preposition Analysis

? HELP ☒ Pronoun Analysis

? HELP ☒ Automatically open output csv file(s) ☒ Automatically compute and open Excel chart(s)

Read Me Open TIPS files Open reminders RUN QUIT

N-grams and co-occurrence viewer

GUI for N-Grams and Word Co-Occurrences

Welcome to this Python 3 script.
 For brief general information about this script, click on the "Read Me" button.
 For brief information on specific lines click on any of the "?HELP" buttons.
 For longer information on various aspects of the script, click on the "Open TIPS files" button and select the pdf help file to view.
 To run the script, click on "RUN" (the RUN button in this script GUI is disabled until all I/O information has been entered). To exit the script, click on "QUIT".

? HELP Select INPUT files directory C:/Users/rfranzo/Desktop/CORPUS DATA/Lynching/1000 TXT BEFORE 1936 MATCHED WITH PDFs/TXT

? HELP Select OUTPUT files directory C:/Program Files (x86)/NLP_backup/Output

? HELP ☐ Date options Date option OFF Aggregate by year Format mm-dd-yyyy Character separator ☐ Position 2

? HELP N-grams search words

? HELP Co-Occurrences search words

? HELP ☒ Normalize results

? HELP ☐ Scale results

? HELP ☐ Lemmatize words

? HELP ☐ Display full information

? HELP ☒ Automatically open output csv file(s) ☒ Automatically compute and open Excel chart(s)

Read Me Open TIPS files Open reminders RUN QUIT

Nominalization

GUI for Nominalization

Welcome to this Python 3 script.
 For brief general information about this script, click on the "Read Me" button.
 For brief information on specific lines click on any of the "?HELP" buttons.
 For longer information on various aspects of the script, click on the "Open TIPS files" button and select the pdf help file to view.
 To run the script, click on "RUN" (the RUN button in this script GUI is disabled until all I/O information has been entered). To exit the script, click on "QUIT".

? HELP Select INPUT TXT file C:/Users/rfranzo/Desktop/CORPUS DATA/Three little pigs/The Three Little Pigs.txt

? HELP Select INPUT files directory

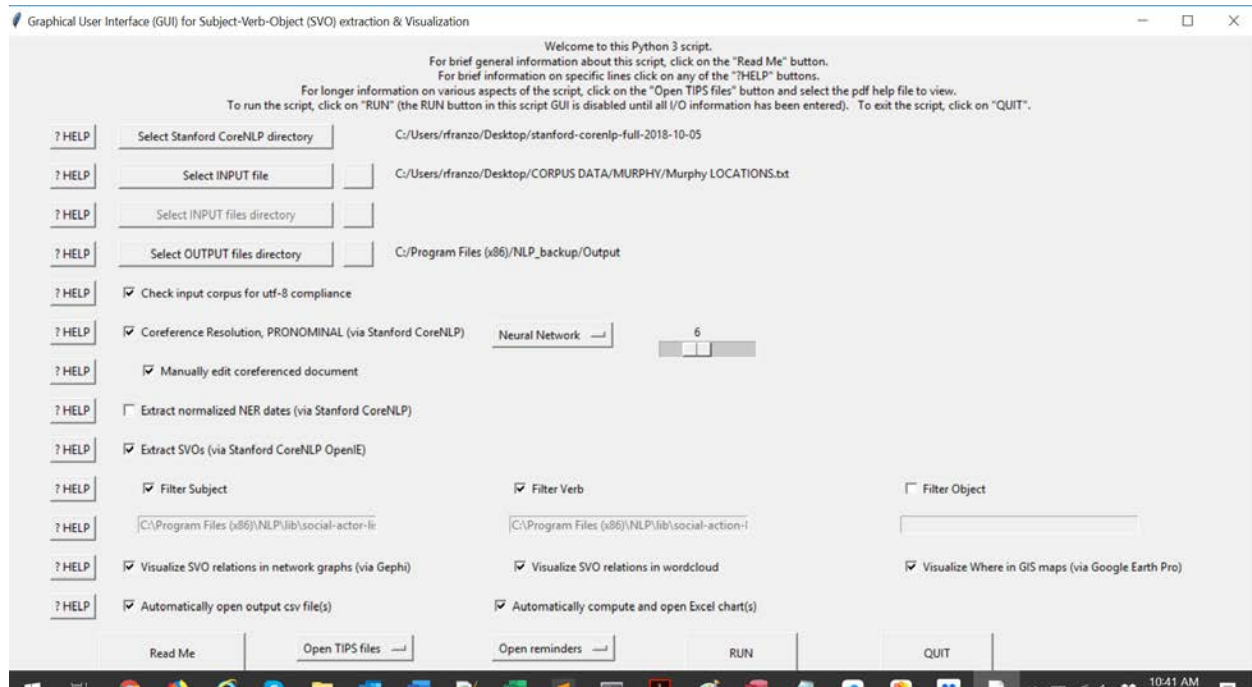
? HELP Select OUTPUT files directory C:/Program Files (x86)/NLP_backup/Output

? HELP ☒ Do NOT produce intermediate csv files when processing all txt files in a directory

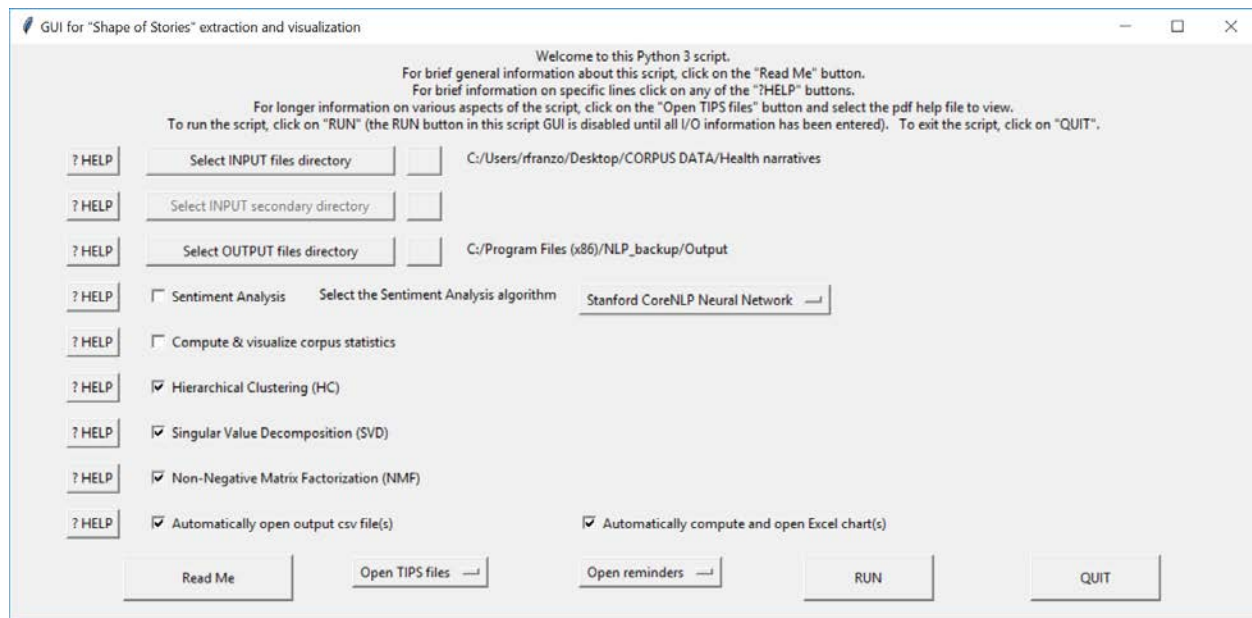
? HELP ☒ Automatically open output csv file(s) ☒ Automatically compute and open Excel chart(s)

Read Me Open TIPS files Open reminders RUN QUIT

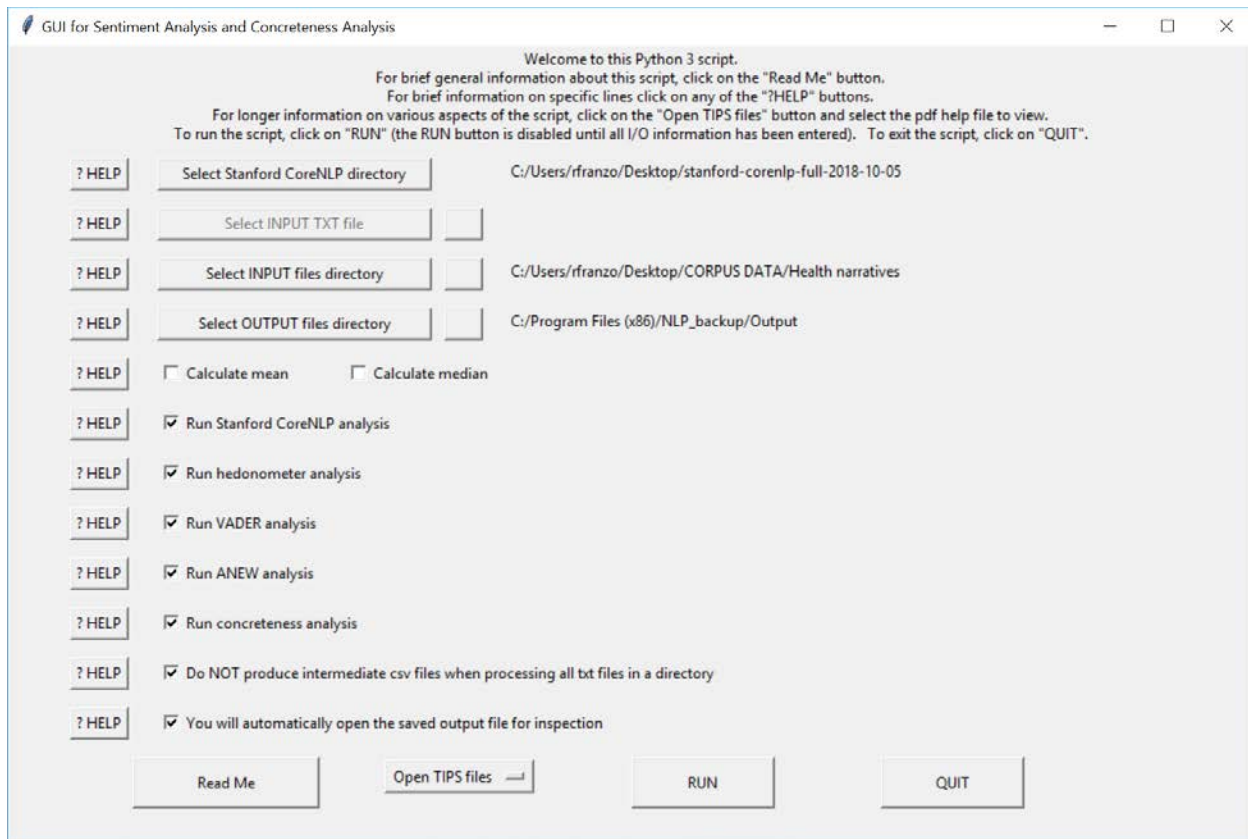
Extracting and visualizing 4 of the 5 Ws of journalism: Who, What, When, Where



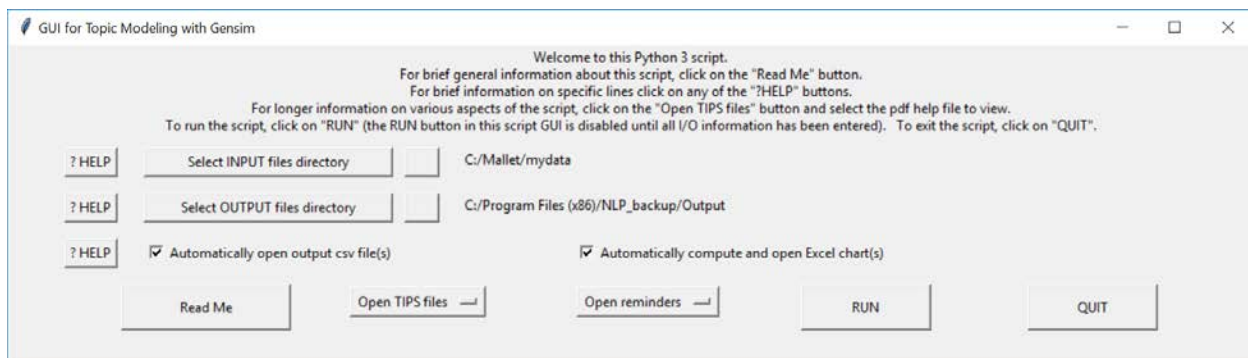
The "shape of stories"

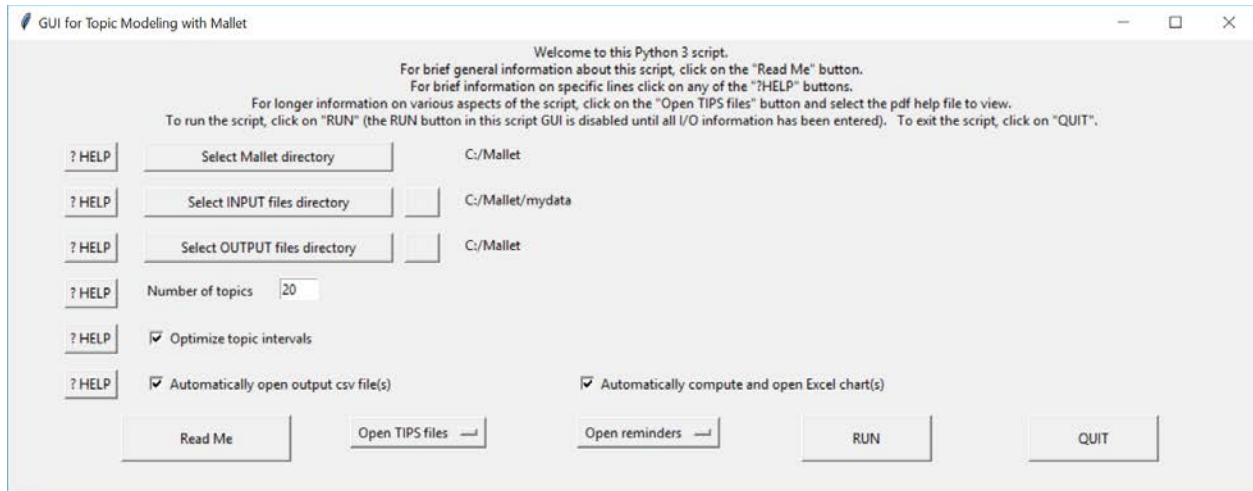


Sentiment and concreteness analysis

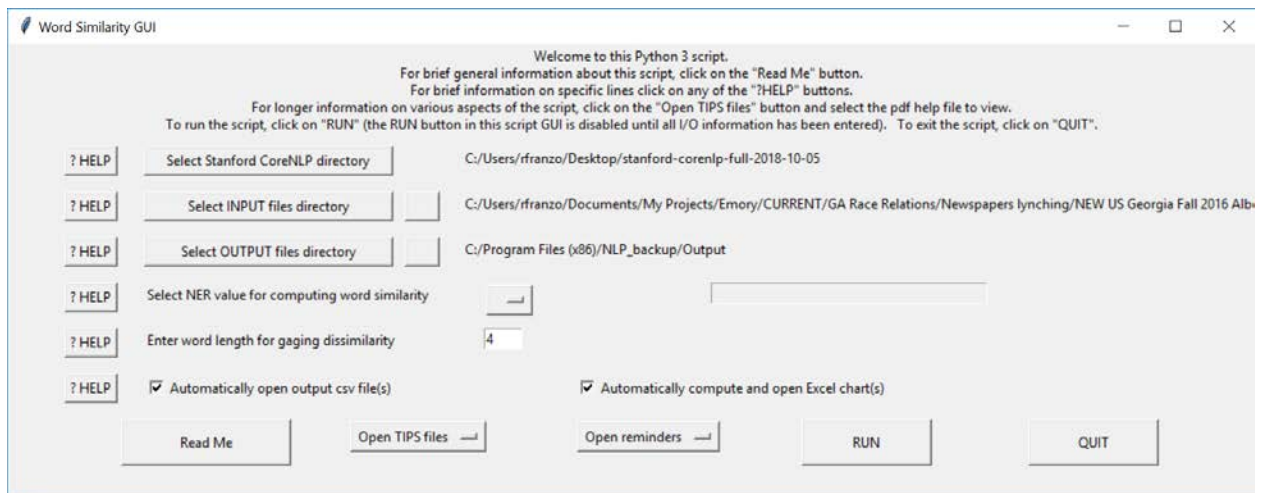


Topic modeling (Gensim & Mallet)

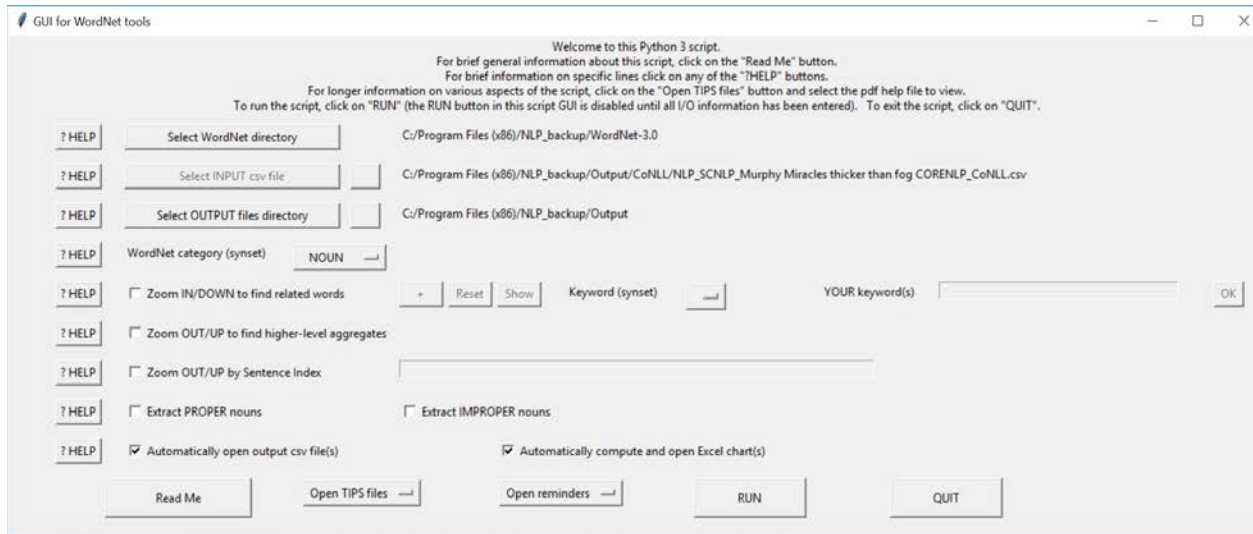




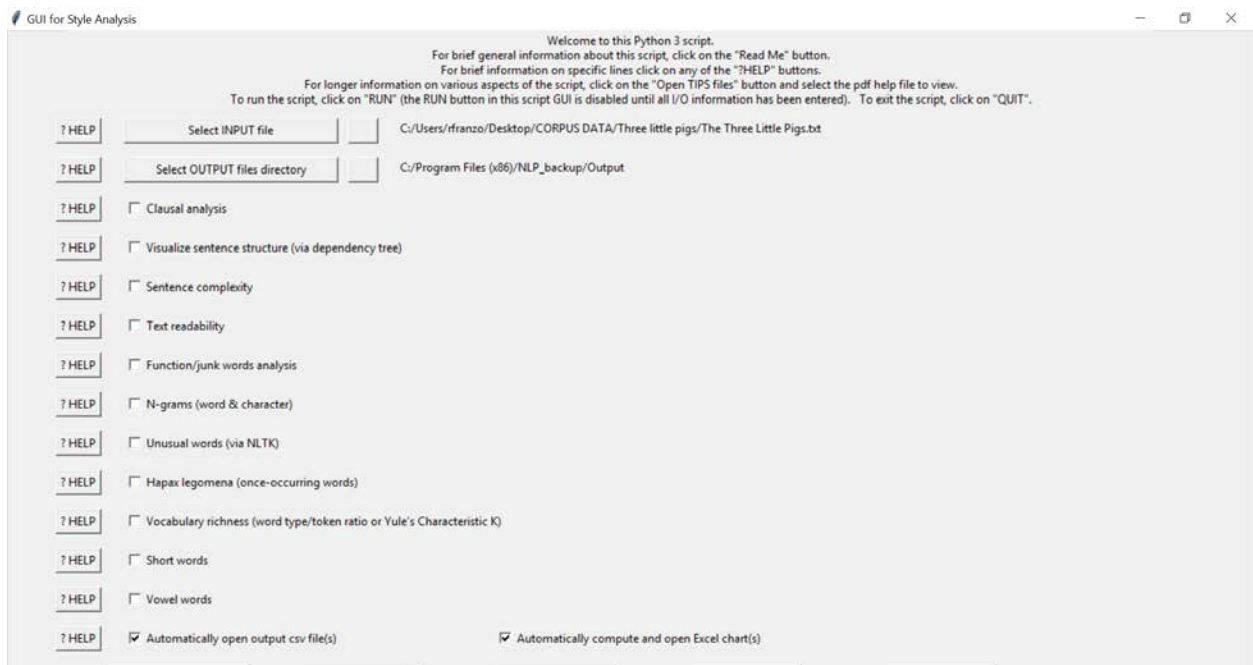
Word similarity



WordNet



Style measures: A summary GUI



Visualization tools

Word clouds

GUI for Word Clouds

Welcome to this Python 3 script.
 For brief general information about this script, click on the "Read Me" button.
 For brief information on specific lines click on any of the "?HELP" buttons.
 For longer information on various aspects of the script, click on the "Open TIPS files" button and select the pdf help file to view.
 To run the script, click on "RUN" (the RUN button in this script GUI is disabled until all I/O information has been entered). To exit the script, click on "QUIT".

? HELP Select INPUT file C:/Program Files (x86)/NLP_backup/Output/Murphy-svo.csv

? HELP Select INPUT files directory

? HELP Select OUTPUT files directory C:/Program Files (x86)/NLP_backup/Output

? HELP Select the word cloud service you wish to use

? HELP Select png image file

? HELP ☐ Use different colors for different words (csv file)

? HELP Select csv field ☐ Color RGB color code + Reset Show

? HELP ☒ Do NOT produce intermediate wordcloud files when processing all txt files in a directory

? HELP ☒ Automatically open output csv file(s) ☒ Automatically compute and open Excel chart(s)

Read Me Open TIPS files Open reminders RUN QUIT

Excel charts

GUI for Excel Charts

Welcome to this Python 3 script.
 For brief general information about this script, click on the "Read Me" button.
 For brief information on specific lines click on any of the "?HELP" buttons.
 For longer information on various aspects of the script, click on the "Open TIPS files" button and select the pdf help file to view.
 To run the script, click on "RUN" (the RUN button in this script GUI is disabled until all I/O information has been entered). To exit the script, click on "QUIT".

? HELP Select INPUT csv file C:/Program Files (x86)/NLP_backup/Output/CoNLL/NLP_SCNLP_Murphy Miracles thicker than fog CORENLP_CoNLL.csv

? HELP Select OUTPUT files directory C:/Program Files (x86)/NLP_backup/Output

? HELP Series 1 + Reset

? HELP Select X-Axis Column Enter X-Axis label

? HELP Select Y-Axis Column Enter Y-Axis label

? HELP ☐ Add hover over info worksheet Select hover over info column

? HELP ☐ Count the Y-Axis values Enter Y-Axis Column specific value

? HELP ☐ Plot chart with 2 Y-Axes Enter second Y-Axis label

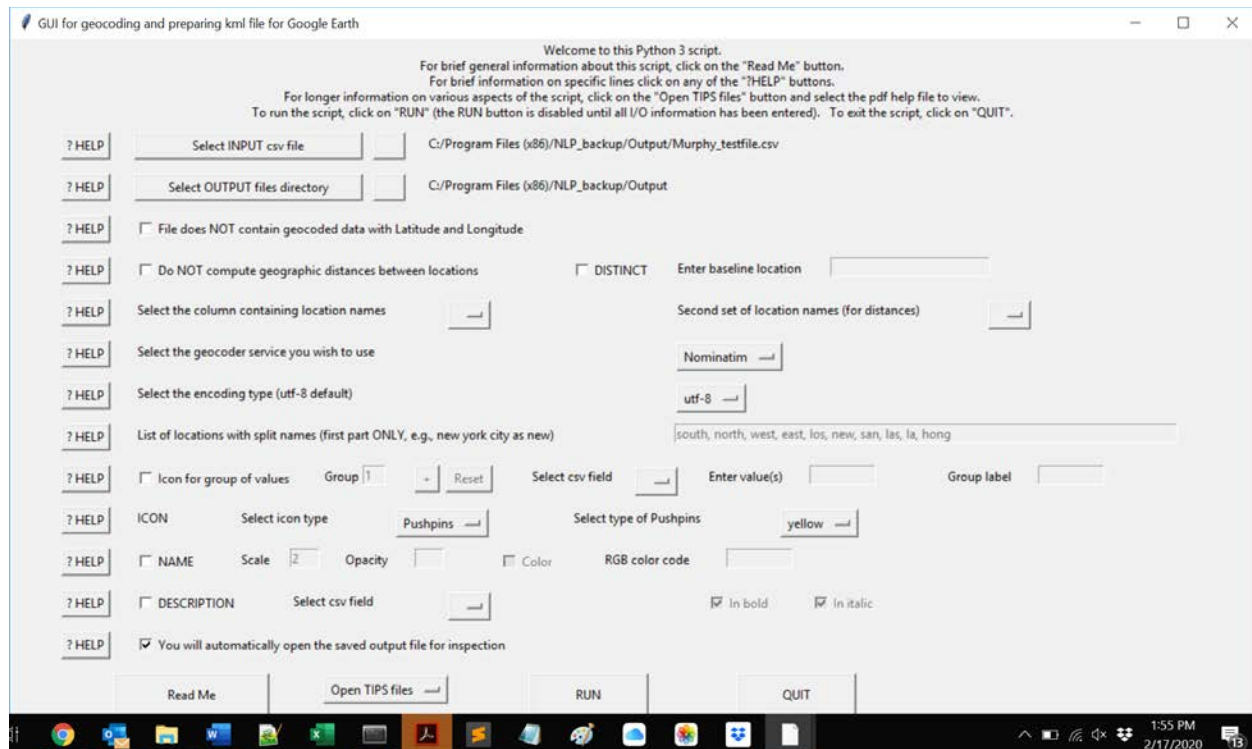
? HELP Select Excel chart type

? HELP Enter the chart title to be displayed

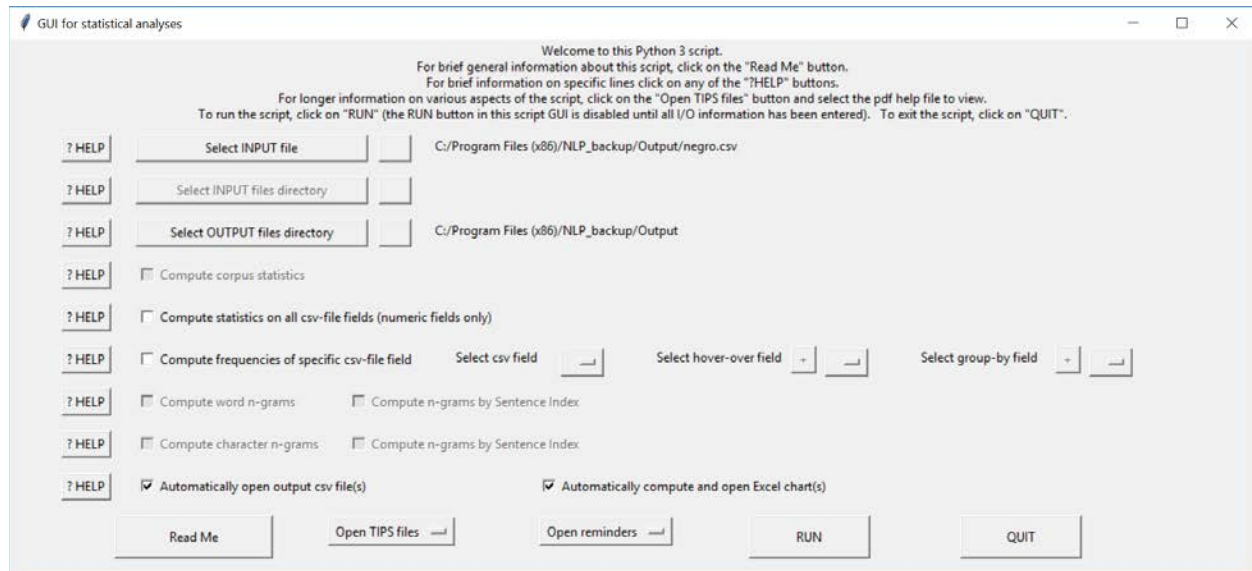
☒ Automatically open output csv file(s) ☒ Automatically compute and open Excel chart(s)

Read Me Open TIPS files Open reminders RUN QUIT

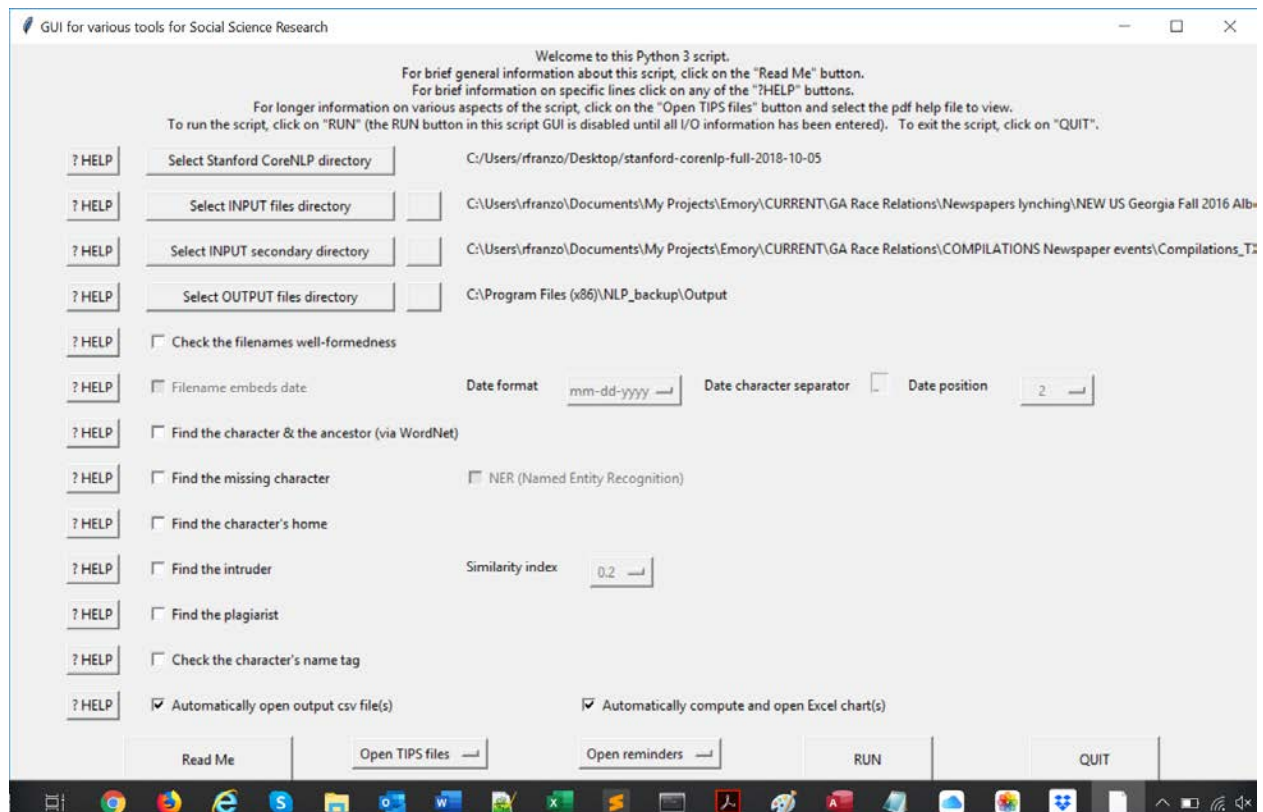
Geocoding and visualizing spatial data in Google Earth Pro



Statistical tools



A set of comprehensive tools for social science research



File and data management tools

Filename checker

Graphical User Interface (GUI) for Filename Checker

Welcome to this Python 3 script.
 For brief general information about this script, click on the "Read Me" button.
 For brief information on specific lines click on any of the "?HELP" buttons.
 For longer information on various aspects of the script, click on the "Open TIPS files" button and select the pdf help file to view.
 To run the script, click on "RUN" (the RUN button in this script GUI is disabled until all I/O information has been entered). To exit the script, click on "QUIT".

? HELP Select INPUT files directory

? HELP Select OUTPUT files directory

? HELP ☐ By dictionary value

? HELP Item character separator ☐ Item position

? HELP csv dictionary file

? HELP ☐ By suffix value

? HELP ☐ By sub-string value

? HELP Enter value

? HELP ☐ By number of embedded items Separator character(s) ☐ Number of items ☐ ☒ Include first # items only

? HELP ☐ Filename embeds date Date format Date character separator ☐ Date position

? HELP ☐ Include subdirectories

? HELP ☒ Automatically open output csv file(s) ☒ Automatically compute and open Excel chart(s)

Read Me Open TIPS files Open reminders RUN QUIT

File content checker & converter

Graphical User Interface (GUI) for File Manager Checker & Converter

Welcome to this Python 3 script.
 For brief general information about this script, click on the "Read Me" button.
 For brief information on specific lines click on any of the "?HELP" buttons.
 For longer information on various aspects of the script, click on the "Open TIPS files" button and select the pdf help file to view.
 To run the script, click on "RUN" (the RUN button in this script GUI is disabled until all I/O information has been entered). To exit the script, click on "QUIT".

? HELP Select INPUT file

? HELP Select INPUT files directory C:/Users/rfranzo/Desktop/CORPUS DATA/Sample text

? HELP Select OUTPUT files directory C:/Program Files (x86)/NLP_backup/Output

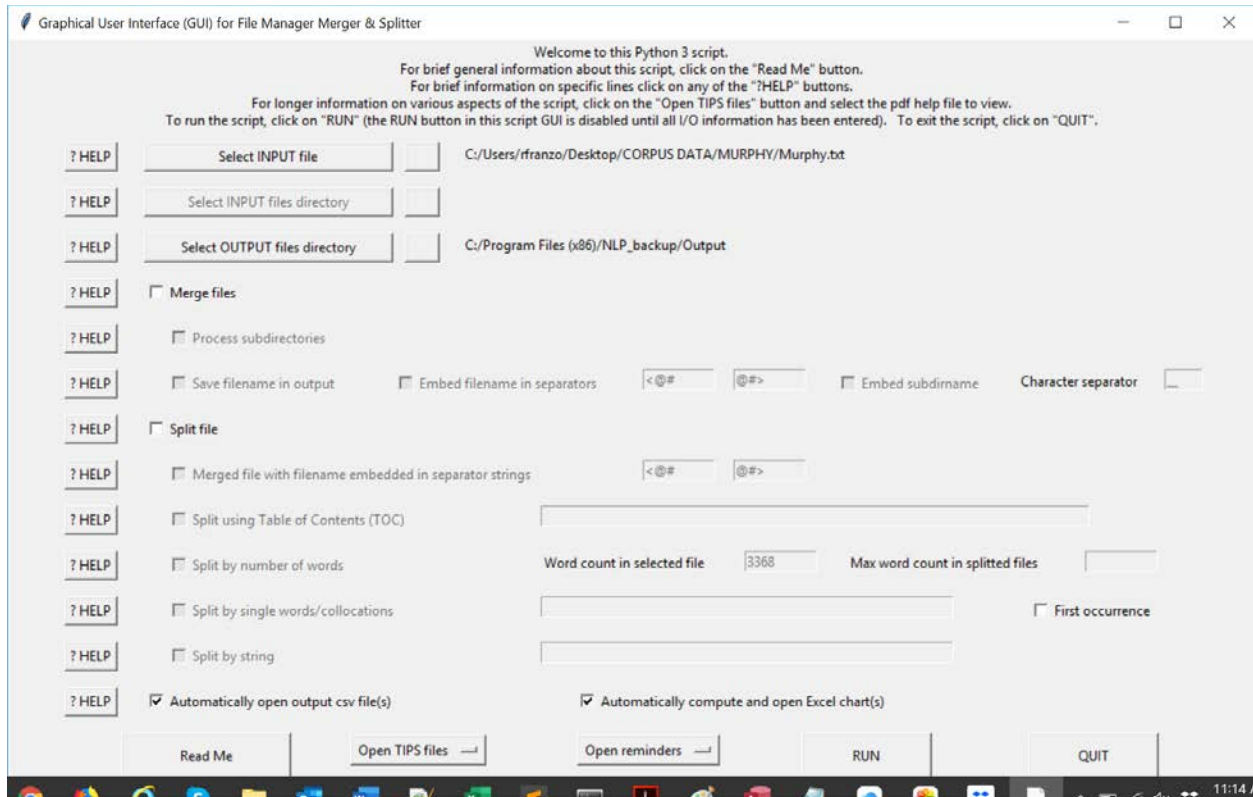
? HELP Check Files

? HELP Convert Files

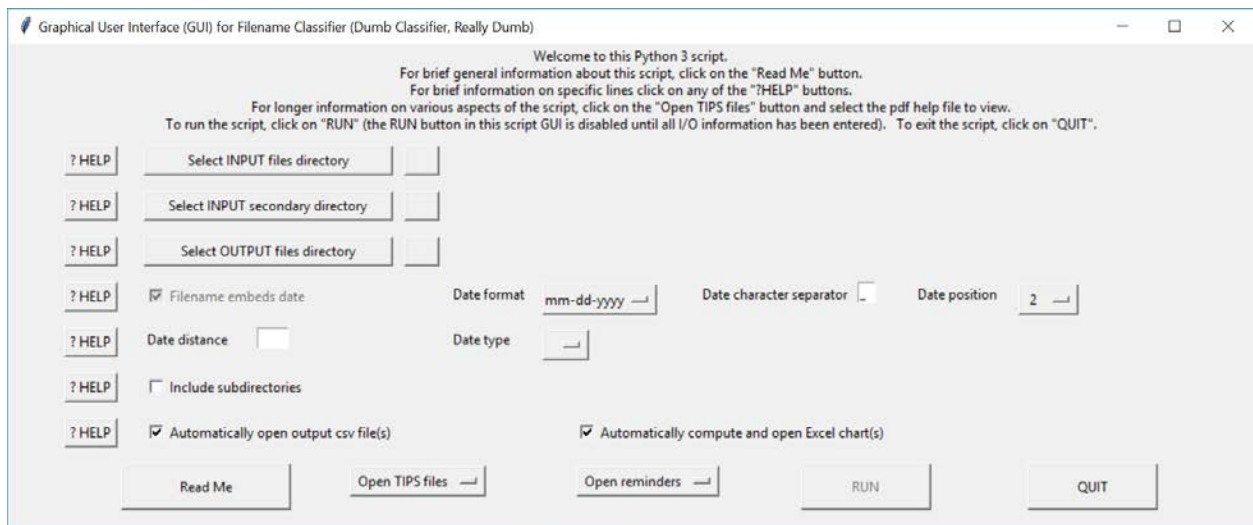
? HELP ☒ Automatically open output csv file(s) ☒ Automatically compute and open Excel chart(s)

Read Me Open TIPS files Open reminders RUN QUIT

File merger & splitter



File classifier (dumb classifier)



File manager

File Handling GUI

Welcome to this Python 3 script.
 For brief general information about this script, click on the "Read Me" button.
 For brief information on specific lines click on any of the "?HELP" buttons.
 For longer information on various aspects of the script, click on the "Open TIPS files" button and select the pdf help file to view.
 To run the script, click on "RUN" (the RUN button is disabled until all I/O information has been entered). To exit the script, click on "QUIT".

? HELP

? HELP

? HELP ☐ List files ☐ Rename files ☐ Copy files ☐ Move files ☐ Delete files

? HELP ☐ By file type

? HELP ☐ By creation & modification date ☐ By author

? HELP ☐ By suffix value ☐ By sub-string value Enter value New substring for renaming

? HELP ☐ Count character(s) embedded in filename

? HELP ☐ Include subdirectories

? HELP ☐ Filename embeds date Date format Date character separator Date position

Data manager

Data Manager GUI

Welcome to this Python 3 script.
 For brief general information about this script, click on the "Read Me" button.
 For brief information on specific lines click on any of the "?HELP" buttons.
 For longer information on various aspects of the script, click on the "Open TIPS files" button and select the pdf help file to view.
 To run the script, click on "RUN" (the RUN button in this script GUI is disabled until all I/O information has been entered). To exit the script, click on "QUIT".

? HELP

? HELP

? HELP ☐ Current matched pair of csv filename and fields

? HELP File

? HELP

? HELP

? HELP ☐ Merge files (Join) Select field

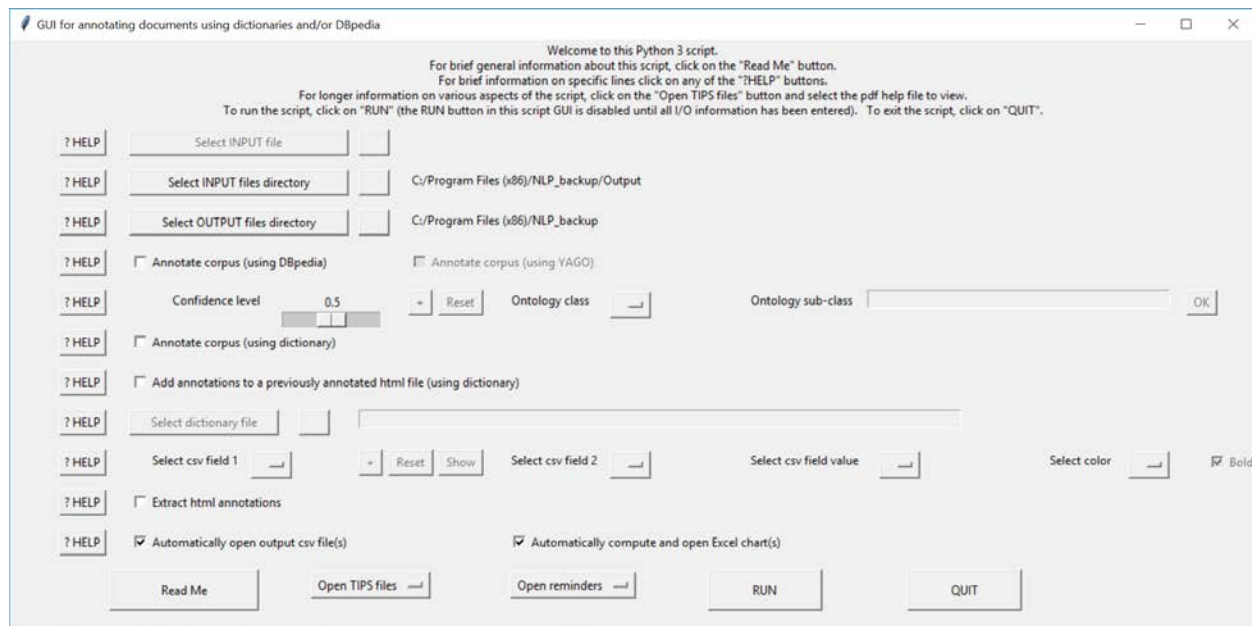
? HELP ☐ Concatenate field values Select field Character(s) separator

? HELP ☐ Append field values Select field

? HELP ☐ Extract field(s) from csv file Select field Enter value (WHERE) and/or

? HELP ☒ Automatically open output csv file(s) ☒ Automatically compute and open Excel chart(s)

Knowledge bases and annotator tools (DBpedia & YAGO, dictionaries)



Work ahead: What needs to be done

We continue to develop new tools. In particular, we want to develop the following set of tools:

1. Allow the use of different language packs in our implementation of Stanford CoreNLP
2. Fix and generalize the use of the Query CoNLL tool
3. Generalize the use of WordNet to allow for a more fine_grained WordNet searches
4. Quote extractor via Stanford CoreNLP
5. Date extractor via Stanford CoreNLP date annotator
6. Generalize the use of the PyTesseract pdf converter
7. Expand the algorithms for sentence complexity
8. All our Java scripts that are based on the Stanford CoreNLP contain inside them the current release of Stanford CoreNLP with 2 negative consequences:
 - a. The size of these Java scripts becomes enormous
 - b. These Java scripts become quickly obsolete as every new release of Stanford CoreNLP would require constant updates of our Java scripts.

We must try to make a call to CoreNLP downloaded on the hard drive rather than embedding it.
9. Implement Semantic Role Labeling (SRL)
10. Update our Java spelling checker
11. Connect the Python Word2Vec library to the NLP suite
12. Text annotator via dictionary, Dbpedia, YAGO
13. Topic segmentation
14. Automatic classification of various aspects of a story: description, action, emotions/feelings, dialogue, evaluation.

More generally, we want to improve the overall organizational structure of NLP suite.










- a. Currently, for reasons of haphazard growth of the scripts, each tool (e.g., geocoder_Google_Earth.py) has its own specialized GUI (e.g., geocoder_Google_Earth_GUI.py) but these GUIs depend upon a more general GUI script (GUI_util.py). GUI_util.py contains references to all available GUIs for all scripts. This is a cumbersome design. There should be no script containing references to all scripts.
- b. For similar reasons, the cumbersome design of config_util.py, containing references to all scripts should be streamlined by passing a parameter for the number of lines to be rewritten in the output config filename.

How to run the NLP suite



















The NLP Suite is launched in command line via Python NLP_main.py; this will fire up the NLP_GUI.py with a convenient Graphical User Interface (GUI) that provides access to the dozens of NLP and visualization tools available to the users.

Download and install NLP Suite.zip













Download the NLP Suite.zip depository and extract the content in an NLP directory, e.g., C:\Program Files (x86)\NLP. Your NLP folder will look like this.

Name	Date modified	Type
 config	6/13/2020 10:37 AM	File folder
 lib	6/13/2020 4:39 AM	File folder
 reminders	6/5/2020 5:03 AM	File folder
 sampleData	6/14/2020 5:27 AM	File folder
 src	6/13/2020 4:37 AM	File folder
 TIPS	6/8/2020 10:37 AM	File folder
 NLP Suite.docx	6/14/2020 8:57 AM	Microsoft Word Doc...
 readme installation.txt	6/13/2020 5:13 AM	TXT File
 requirements.txt	6/13/2020 5:10 AM	TXT File

The **config subdirectory** will contain a bunch of config.txt files with the Input/Output options for each script (e.g., Stanford-CoreNLP-config.txt or geocoder-GoogleEarth-config.txt). Each config file will have as suffix the name of the script followed by -config.txt.

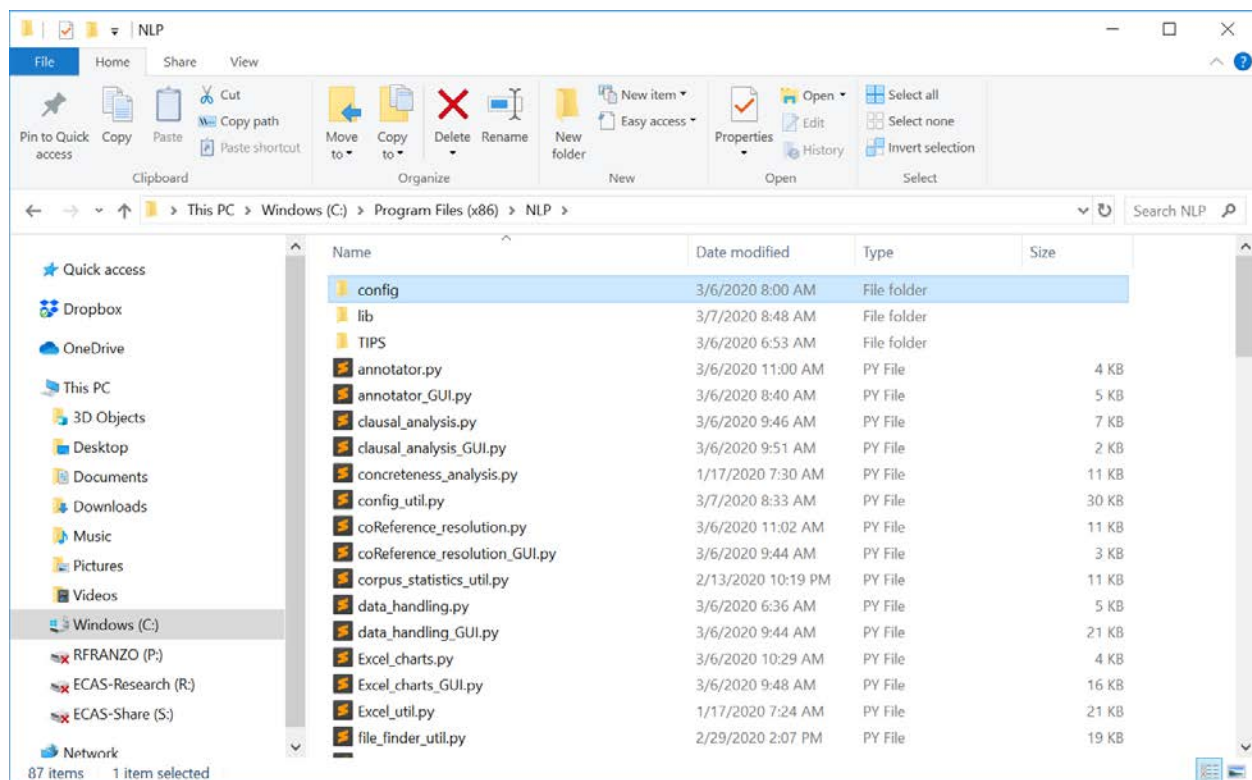
Name	Date modified	Type	Size
 annotator-config.txt	11/26/2019 5:49 PM	TXT File	1 KB
 clausal-analysis-config.txt	9/25/2019 10:25 AM	TXT File	1 KB
 Excel-config.txt	11/15/2019 3:45 PM	TXT File	1 KB
 file-handling-config.txt	11/16/2019 7:38 PM	TXT File	1 KB
 function-words-config.txt	10/26/2019 3:47 PM	TXT File	1 KB
 geocoder-GoogleEarth-config.txt	12/5/2019 12:31 PM	TXT File	1 KB
 KWIC-config.txt	10/16/2019 11:45 AM	TXT File	1 KB
 n-grams-word-co-occurrences-config.txt	11/18/2019 2:38 PM	TXT File	1 KB
 NLP-config.txt	12/2/2019 5:45 PM	TXT File	1 KB
 nominalization-config.txt	11/8/2019 12:44 PM	TXT File	1 KB
 noun-verb-config.txt	12/3/2019 6:31 PM	TXT File	1 KB
 query-conll-config.txt	10/1/2019 12:25 PM	TXT File	1 KB
 sentiment-concreteness-analysis-config.txt	10/31/2019 1:32 PM	TXT File	1 KB
 shape-of-stories-config.txt	10/31/2019 8:48 AM	TXT File	1 KB
 social-science-research-config.txt	11/1/2019 8:32 PM	TXT File	1 KB
 Stanford-CoreNLP-config.txt	12/5/2019 7:00 AM	TXT File	1 KB
 SVO-config.txt	11/25/2019 10:23 AM	TXT File	1 KB
 topic-modeling-gensim-config.txt	10/31/2019 8:48 AM	TXT File	1 KB

The **lib subdirectory** will contain a set of files used by various scripts.

Name	Date modified	Type	Size
 gexf	3/7/2020 8:48 AM	File folder	
 100 sight English words.txt	2/27/2018 6:08 AM	TXT File	1 KB
 300 sight English words.txt	2/27/2018 6:15 AM	TXT File	2 KB
 Concreteness_ratings_Brysbaert_et_al_BRM.csv	2/12/2018 4:34 PM	Microsoft Excel Com...	1,569 KB
 EnglishANEW.csv	2/22/2018 4:51 PM	Microsoft Excel Com...	3,593 KB
 EnglishShortenedANEW.csv	2/22/2018 4:34 PM	Microsoft Excel Com...	319 KB
 hedonometer.json	2/22/2018 4:34 PM	JSON File	1,724 KB
 social-actor-list.csv	12/1/2019 8:18 PM	Microsoft Excel Com...	138 KB
 social-actor-list-SORTED.txt	12/1/2019 8:18 PM	TXT File	138 KB
 stopwords.txt	5/21/2019 3:25 AM	TXT File	4 KB
 sw.txt	10/7/2016 6:51 PM	TXT File	2 KB
 vader_lexicon.txt	3/2/2016 4:51 PM	TXT File	424 KB

The **TIPS subdirectory** will contain some 50 pdf files of help an available to the various NLP tools.

Name	Date modified	Type	Size
TIPS_NLP_Clausal tags.pdf	8/25/2019 2:18 AM	Adobe Acrobat Docu...	44 KB
TIPS_NLP_Coreference Resolution.pdf	5/15/2019 3:46 AM	Adobe Acrobat Docu...	782 KB
TIPS_NLP_DEPREL (Dependency Relations) Stanf...	5/15/2019 3:47 AM	Adobe Acrobat Docu...	307 KB
TIPS_NLP_Excel Charts.pdf	9/30/2019 11:12 AM	Adobe Acrobat Docu...	404 KB
TIPS_NLP_Function Words Analysis.pdf	5/15/2019 3:47 AM	Adobe Acrobat Docu...	287 KB
TIPS_NLP_Gephi Displaying network graphs.pdf	5/15/2019 3:48 AM	Adobe Acrobat Docu...	2,018 KB
TIPS_NLP_Gephi Dynamic network graphs.pdf	5/15/2019 3:48 AM	Adobe Acrobat Docu...	856 KB
TIPS_NLP_Google Earth Pro Description.pdf	5/15/2019 3:48 AM	Adobe Acrobat Docu...	2,868 KB
TIPS_NLP_Google Earth Pro From KML to Excel...	5/15/2019 3:48 AM	Adobe Acrobat Docu...	229 KB
TIPS_NLP_Google Earth Pro HTML.pdf	5/15/2019 3:48 AM	Adobe Acrobat Docu...	324 KB
TIPS_NLP_Google Earth Pro Icon.pdf	5/15/2019 3:48 AM	Adobe Acrobat Docu...	660 KB
TIPS_NLP_Google Earth Pro KML options.pdf	5/15/2019 3:48 AM	Adobe Acrobat Docu...	1,178 KB
TIPS_NLP_Google Earth Pro.pdf	5/15/2019 3:48 AM	Adobe Acrobat Docu...	801 KB
TIPS_NLP_Java download install run.pdf	8/29/2019 8:59 AM	Adobe Acrobat Docu...	92 KB
TIPS_NLP_KWIC (Key Words In Context).pdf	5/15/2019 3:47 AM	Adobe Acrobat Docu...	110 KB
TIPS_NLP_Ngrams and Word co-occurrences.pdf	5/15/2019 3:47 AM	Adobe Acrobat Docu...	286 KB
TIPS_NLP_NLP Basic Language.pdf	5/15/2019 3:47 AM	Adobe Acrobat Docu...	337 KB



Download and install Python and Java

To run the Python and Java NLP tools you will also need to install both Python and Java. Read the TIPS files

TIPS_NLP_Python download install run.pdf

TIPS_NLP_Java download install run.pdf

Install Python libraries

There is an important file inside your NLP directory: requirements.txt (on the function of a requirements.txt file, see TIPS_NLP_Python download install run.pdf). The file contains all the Python libraries used by the various scripts of the NLP suite. Run it, to install automatically all the libraries once and for all.

```
pip install -r requirements.txt
```

Download and install Stanford CoreNLP: A language parser and more

Depending upon which tools you use you will also need to install various other **freeware** tools. One such fundamental tool is the suite of NLP tools by Stanford CoreNLP (<https://stanfordnlp.github.io/CoreNLP/download.html>). Read the TIPS file

TIPS_NLP_Stanford CoreNLP download install run.pdf

Download and install Gephi: A network software

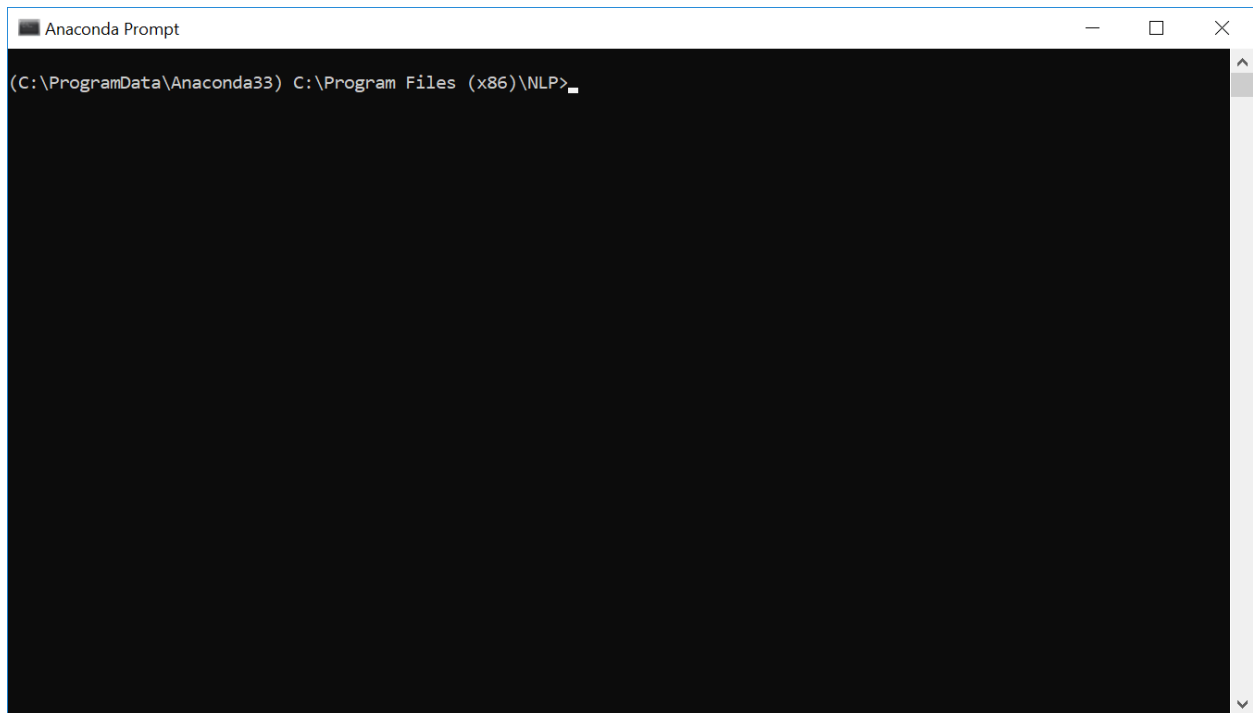
You will also want to download and install Gephi, a freeware network software (<https://gephi.org/users/download/>).

Download and install Google Earth Pro: A Geographic Information System (GIS) software

If you want to automatically visualize maps, you will need to download and install Google Earth Pro, a freeware Geographic Information System (GIS) software (<https://www.google.com/earth/versions/>).

Run NLP.py

Go to command line



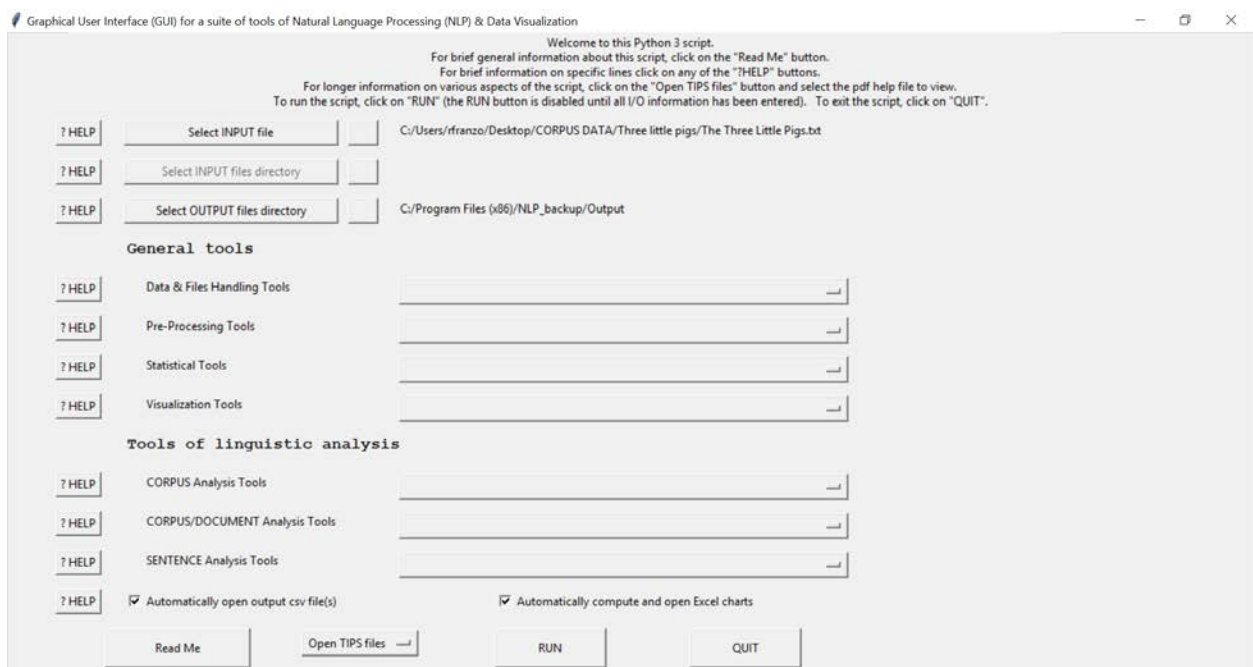
change directory to your NLP directory

cd C:\Program Files (x86)\NLP

type

python NLP.py

hit return and, with a little bit of luck, you will see this:



This GUI (Graphical User Interface) will give you access to all NLP tools available in the NLP suite (see all sections below). You can run any of the scripts that have an embedded GUI directly, without accessing first the NLP GUI. Just type, for instance,

`Python Stanford_CoreNLP.py`

Or

`Python geocoder_Google_Earth.py`

And their respective GUIs will open up.