Web-Based Tools for Text Analysis and Exploration

Click to schedule a meeting with the DITI Team

Tools for Uploading/Pasting & Analyzing Texts

- **Word Counter**: This is a user-friendly basic word counting tool; it allows you to count single words, bigrams, and trigrams in plain text files and to download spreadsheets with your results. The max file upload is 10MB. Click here to navigate to Word Counter
- **Same Diff**: With this tool, you can upload two files to see which words appear in both, as well as which words are unique to each file; you can download spreadsheets with the counts for each text. Max file upload is 10MB. Click here to navigate to Same Diff
- **Word Trees**: This is a good way to see patterns in word usage, based on words that appear before and after a term or terms of interest. There are some restrictions in size; fewer than 1 million words should work, but loading that much text in might be slow. Click here to navigate to Word Trees
- **Lexos:** This is a tool for preparing and analyzing digital texts; it offers several options for text preparation, and a wide range of different analytical possibilities. Importantly, Lexos also preserves all the changes that are made to a text, so that any results can be reproduced. Click here to navigate to Lexos
- **Drag-and-Drop Sentiment Analysis**: This is an exploratory tool that lets you see the top negative and positive words, as well as common bigrams and trigrams. Click here to navigate to Drag-and-Drop Sentiment Analysis. For more on how sentiment analysis works, click here to navigate to a lesson on sentiment analysis
- **Plot Mapper:** This is an experimental tool for exploring the shapes of plots. Click here to navigate to Plot Mapper
- **Voyant:** This suite of tools gives you counts of words and lets you compare patterns in word locations and frequencies, or examine keywords in context, along with a few other options. Voyant will let you upload larger files than most other interfaces (up to as many as 4 million words, though it may take more than one try to successfully upload very large files). Click here to navigate to Voyant

Developed by: Sarah Connell
Questions? Contact nulab.info@gmail.com

Additional Tools and Resources

- **Serendip**: This is a tool that supports topic modeling, which is a method that uses machine learning to discover "topics," or sets of related terms, in collections of texts. Click here to navigate to Serendip
- **Women Writers Vector Toolkit:** This is a resource developed by the Women Writers Project that lets you explore word embedding models, which are a machine-learning based method for discovering relationships between words in large collections of texts. Click here to navigate to the Women Writers Vector Toolkit
- **NULab for Digital Humanities and Computational Social Science Resources:** This page contains a wide range of resources and datasets. Click here to navigate to the NULab's resources page
- **Programming Historian:** This site includes many different tutorials for a broad range of methods of digital analysis; it is not exclusively focused on history and includes materials that are useful for literary studies as well. Click here to navigate to Programming Historian

Northeastern-specific resources

The resources below are available through the Northeastern University Library.

- **Constellate:** This is a platform for learning and performing text analysis, with both code notebooks and a web-based interface. A free JSTOR account is needed for Northeastern users to log in to Constellate. To access Constellate, Click here for the Library's list of databases and log in with your Northeastern credentials. Click here to navigate to Constellate
- **ProQuest TDM Studio**: This resource provides both web-based and code-based options for exploring textual analyses with ProQuest datasets. To get started, create an account with your Northeastern email. Click here to navigate to ProQuest TDM Studio
- **Library guides**: The Library also offers guides and links to resources for text analysis; click here for a guide to getting started, and click here for a guide on vendor policies.

Developed by: Sarah Connell
Questions? Contact nulab.info@gmail.com