



---

## Web-Based Tools for Text Analysis and Exploration

---

### 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.  
<http://databasic.io/en/wordcounter/>
- **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.  
<http://databasic.io/en/samediff/>
- **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. <http://jasondavies.com/wordtree/>
- **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. <http://lexos.wheatoncollege.edu/>
- **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.  
<http://storybench.shinyapps.io/textanalysis/> For more on how sentiment analysis works, see: <http://programminghistorian.org/en/lessons/sentiment-analysis>
- **Plot Mapper:** This is an experimental tool for exploring the shapes of plots.  
<http://nickbeauchamp.com/projects/plotmapper.php/>
- **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). <http://voyant-tools.org/>

### 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.  
<http://vep.cs.wisc.edu/serendip/#gettingstarted/>



- 
- **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. <http://lab.wwp.northeastern.edu/>
  - **NULab for Texts, Maps, and Networks, Resources:** This page contains a wide range of resources and datasets. <https://cssh.northeastern.edu/nulab/research/resources/>
  - **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. <http://programminghistorian.org/>