**ENGL 4710**

**Prof. Mya Poe**

**Text Analysis using Lexos**

**Important Terminology**

* Corpus/corpora: a text or collection of texts that is used for analysis
* Stopwords: the most common words in a language
  + (English: the, a, she, her, it, they, etc.)

**Lexos website:** <http://lexos.wheatoncollege.edu/>

**What is Lexos?**

Lexos is a powerful web-based tool to help explore different corpora of digitized texts. Lexos provides a workflow of effective practices to help you be cognizant of choices you are making. For Lexos, separate .txt files are required.

Lexos provides a step-by-step guide for corpus uploading, preparation, and analysis.

**Step-by-Step Guide for using Lexos:**

1. Upload
   1. Upload your corpus (your separate .txt. files)
   2. Click “Browse” and select your entire corpus
2. Manage
   1. Make sure all documents in the corpus you want to use are selected
      1. Blue = selected
      2. Gray = not selected
3. Prepare
   1. Click “Prepare” and then “Scrub,” which transforms the text in your corpus and makes choices that will impact your results. Some possibilities:
      1. Make all your letters lowercase (the computer reads upper and lower case letters as two separate characters)
      2. Remove punctuation, which may influence results
      3. Remove a list of words, usually stopwords(Get a list of English stopwords here: <https://gist.github.com/sebleier/554280>**)**
   2. Once you have finished preparations, click “Apply” and wait a few minutes
   3. You can also download your corpus after preparation.
4. Visualize
   1. Click “Visualize” and a drop down of visualization choices will appear
   2. Click “Word Cloud” to visualize a word cloud across the entire corpus
   3. Click “Multi Cloud” to visualize word clouds for each individual text
   4. “Rolling Window” allows you to look at word trends across **one** document
5. Analyze
   1. The “Analyze” tab provides a drop down menu of more analytical options for analyzing your texts
   2. This includes comparing texts through functions like “Top Words” and “Similarity Query”
6. This is just a sample of the wide variety of tools Lexos offers for exploration