From *Issues in Science and Technology Librarianship*
http://www.istl.org/17-spring/refereed3-appendix.docx

Appendix

All data files, analyzed results, MATLAB and Mathematica scripts (matlab_TI_analysis and mathematica_TI_analyis) are available on figshare. The data and results are assigned a CC-BY license, while the scripts are MIT license. The permanent DOI link for the entire collection of supporting information is https://doi.org/10.6084/m9.figshare.c.3279968.v2. The collection includes:

Data Files - Text Analysis of Chemistry Thesis and Dissertation Titles
https://doi.org/10.6084/m9.figshare.3469121.v1

Results - Text Analysis of Chemistry Thesis and Dissertation Titles

https://doi.org/10.6084/m9.figshare.3469166.v2

Scripts - Text Analysis of Chemistry Thesis and Dissertation Titles

https://doi.org/10.6084/m9.figshare.3469169.v2

Readers are encouraged to download the datasets and programming scripts to use in their own work, with appropriate credit to this publication and author. These scripts can be run without expert knowledge of MATLAB and Mathematica.

Instructions for executing matlab_TI_analysis

1. Copy the matlab_TI_analysis.m file to an appropriate file directory on your computer.
2. Copy the dataset text file you wish to analyze into the same file directory.
3. Open MATLAB, and navigate to the file directory where you saved the matlab_TI_analysis.m and dataset text file. Right click on these files and select "`Add to Path`".

4. Open the matlab_TI_analysis.m file by double clicking the file, or navigating to open file menu.
5. Replace "`filename.txt`" with your dataset filename.
6. In the command window, type `matlab_TI_analysis,` then click `enter`.

Instructions for executing mathematica_TI_analysis

1. Copy the mathematica_TI_analysis.nb file to an appropriate file directory on your computer.
2. Copy the dataset text file you wish to analyze into the same file directory.
3. Open the mathematica_TI_analysis file in Mathematica
4. Set the directory path to the saved mathematica_TI_analysis and dataset. For example, "C:\\Documents\Mathematica". Next, replace "`filename.txt`" with your dataset filename.
5. On the right side of the Mathematica notebook, right click on the open bracket and select "Evaluate Cell". Repeat this for each cell (there are three), starting from the top.

Instructions for Adapting Code with Custom Stop Words and Variant Word Combinations

1. Follow the regular expression syntax structure present in the matlab_TI_analysis and mathematica_TI_analysis scripts for either deleting a word or combining words.
2. For example, in MATLAB, to delete a word, the code is as follows:

```
data = regexprep(data,'\<the\>','');
```

   This above line of code would delete all "the" words in the dataset.

3. To combine words in MATLAB, the syntax is as follows:

```
data = regexprep(data,'\<application\>','application*');
data = regexprep(data,'\<applications\>','application*');
```

These two lines of code combine all instances of “application” and “applications” in the dataset into application*. Note: During the production of this manuscript, MathWorks released a new string datatype and several new functions for working with text in MATLAB. This new datatype along with the new text functions may provide an alternative approach to delete and combine words in MATLAB. Ref: http://blogs.mathworks.com/loren/2016/12/22/singing-the-praises-of-strings/

4. An analogous approach can be used in Mathematica, with slight differences in the syntax, see mathematica_TI_analysis file for examples.