



# Coding Challenge - Natural Language Processing

Expected turnaround time: 3 days

Write a text-processing tool that will read through a given text file, process its contents, and output your findings in text format.

## Input:

The file name to a text file. Assume the article is several hundred words in length with multiple paragraphs, such as an extract from Wikipedia.

## Processing:

- 1) Identify the articles used (i.e., 'a', 'an', 'the').
- 2) Identify dates (which might be expressed in a variety of ways such as '15/11/2012', '15/11/12', '15th March 1999', '15th March 99' or '20th of March, 1999').
- 3) Identify proper nouns used (names of persons, places, and things, etc).

This is a bit open ended, and somewhat intentionally so. The aim is for you to try to write something which figures out as many common patterns as possible, in which articles, dates, and names are present in text.

## Output:

Five lines of output for each given article.

First line -> number of occurrences of 'a'.

Second line -> number of occurrences of 'an'.

Third Line -> number of occurrences of 'the'.

Fourth Line -> number of occurrences of date information.

Fifth Line -> number of occurrences of proper nouns

Bonus (optional):

Separate the proper nouns into persons, places, and things.

**Presentation:**

The tool will be user initiated at the Linux command line. Your script should expect a file name of a plain text file as an input argument. You can assume that you will not know the content of that text file ahead of time.

The tool will read the text from the file and process it according to the guidelines above. The tool will then output the five lines of output as shown above. You may output more than 5 lines of text if you include the bonus output in your project.

You may use any technologies or software you feel you need to make the process work.

Once completed, check the code into your own github repository and share the link with Revcloud. Be sure that your repo includes the code needed for all pieces of the project.

You will then present the finished project with the Revcloud team via screen sharing over Skype or similar tools.

Be prepared to explain your project, the flow of your code, and any design or implementations decisions you have made.

**Good luck!**