COMP3506 Assignment 2 Development Notebook

**Word Count**

Need to create a data structure that holds the list of used words (and their count and location).

* Need to process the text to exclude certain grammar characters that change the nature of the word
* Need to ignore stop words

Trie could be useful here. Going to start with this, and see how it works from there.

Trie is not a binary tree. Number of children will be equal to the number of characters in the alphabet.

Linked list implementation seems to make sense. Worse case scenario is still O(n).

Node must have:

A character representing the node’s value

An iterable list of children nodes

An int representing the number of times the node signifies its end-of-word position (and therefore count of word)

A list of word coordinates (the word’s positions in the text).