Studies Case: Words search

1.Decomposition

- Searching the synonyms of the word.
- Iterating through each of the synonym and the keyword in a corpus and counting the occurrence.

2.Pattern Recognition

- Counting the number of occurrences of each synonym and the original word of a document. This process will need to be repeated for the keyword and each synonym.
- Using the pattern above to searching each document.

3.Data Abstraction

- The keyword should probably be a user-input, so maybe data that is given by the user if it's a software.
    - This word needs to be valid, so maybe check in a dictionary database.
- Thesaurus is a dictionary in the data structure sense of the term (key:value) where the key is the word, and the value is a list of synonyms.
    - For example, each keyword written by the user result in an entry like this {keyword : "synonym1, synonym2, synonym3}
- The corpus is a set of documents. In the program, we will probably create a list of paths so we can iterate through the list to iterate through each document. The corpus could also be decide by a user but we will probably not see that in this course.

4.Algorithm

1. Select the word (and store it in a variable)
    a. Confirm the validity of the word
    b. If word not valid, loop 1 and 1a.
2. Search for the synonym and create the thesaurus with all the synonyms.
3. Iterate through the list of documents
    a. Iterate through each document
        i. Count every occurrence of the keywords and synonyms.
4. Create a report of the occurrence

5. Real life problem

I used to study ancient text in ancient Greek. Because the language is dead, we need to be sure the translation seems accurate through the comparison of occurrence of a word. So, it's would have been useful to iterate through a document and not only count but give an index or a link to compare later on each of the occurrence.