Exercise #21

- Stopword elimination problem
  - Input:
    - A large textual file containing one sentence per line
    - A small file containing a set of stopwords
    - One stopword per line
  - Output:
    - A textual file containing the same sentences of the large input file without the words appearing in the small file
    - The order of the sentences in the output file can be different from the order of the sentences in the input file

Exercise #21 - Example

- Input files
  - Large file
    - This is the first sentence and it contains some stopwords
    - Second sentence with a stopword here and another here
    - Third sentence of this stopword example

- Stopword file
  - a
  - an
  - the

Exercise #21 - Example

- Output file
  - This is first sentence it contains some stopwords
  - Second sentence with stopword here another here
  - Third sentence of stopword example