For our application we decided to use Spring Boot, a SQL database and AngularJs for the UI.

In the database we have 2 tables:

Title: id, name (the title of the article)

Occurrences: id, word (top 10 words from the article), occurrences (the number of times the word appeared in the article), foreign key (to the name of the article)

This 2 tables are mapped using entities having the same name. We used 2 repositories for the CRUD operations with the database, both implementing JpaRepository interface.

For the part of the program that requires a single article title we used a controller (ArticleController) that has a method that maps the given URL and has a parameter (an article title). Its return type is a DTO that contains a list of Occurrences, the search time, the source (database or Wikipedia) and the name of the article. Also, in this method we verify if the title is in the database (in this case we get the information from there) or not (we first insert the data into the database and return the DTO).

The FileController is used when a file with multiple titles is uploaded, requires a MultipartFile and returns a map of words and occurrences.

Both controllers call methods from the services classes that are autowired using annotantions.

The services package contains 2 packages: impl and tools.

The impl package has the implementation of the interfaces for Occurrence and Title services and also a MainService that has 2 methods: showWordsForSingleTitle and showWordsForMultiTitles. These methods will further be used in the ArticleController and FileController in order to send the data to the UI.

The tools package contains the tools and logic of the application. FileReader, MapMerger and ReadURL are basic tools that are used throughout the application.

FirstWordsGenerator will process the content of an article and return the top 10 words using its methods.

SingleTitleProcess uses the FirstWordsGenerator methods to process the content and create a list of Occurence objects that will be used in MainService.

MultiTitlesProcessExecutor is a Runnable class that processes the content and creates a map of all the words. It doesn't contain only the top words because this will be used for multiple titles processing.

This will happen in MainService where using an Executor each title will be processed by a thread and then added to a common map which will be sorted and the top words will be retrieved.