

*Duration:* 30 minutes + 5 minutes for uploading to an open git repository (github, gitlab, ...)

### **Programming Task**

Write an application that analyzes two text files, with the following functional requirements:

- read two text files loaded from file system (or hard-code the strings for a quick start)
- generate a list of words that appear in BOTH documents
- all words must first be converted to lowercase
- write the output in CSV with the following format: <word>, <frequency in file1>, <frequency in file2>, <total frequency>

### **Optional Requirements**

- **Filtering.** All output words must only contain letters (no numbers or special characters)
- **Sorting.** Allow the user to sort the output either: a) alphabetically, b) by word length or c) by total word frequency.

You can use any libraries or third-party components available under an appropriate license (BSD, Apache, or similar), as long as you reference them properly.

### **Evaluation Criteria**

Before you start writing the actual code, please take the time and use pseudo-code to outline your plan and include this as a comment in your implementation.

Please note that we do not expect you to complete all the listed tasks. Concentrate on the main task first, and then tackle those optional task(s) that you consider feasible to address with the remaining time. Criteria for this assessment include functionality and stability of the tool, as well as the structure and documentation of the source code.

Once finished, please send a link to the publicly accessible git repository to [philipp.kuntschik@fhgr.ch](mailto:philipp.kuntschik@fhgr.ch).