University of Windsor

COMP-8547 Advanced Computing Concepts
Fall 2024 - Final Project

Deadlines:

- Presentation: **Week 11/12** during the class
- Upload final project report, presentation, source code before the deadline.

Guidelines: This project will be submitted as a group of 4 - 5 students (recommended). However, group participants will be evaluated individually and may receive a different mark from the other group members.

Goals: The aim of the project is to apply the concepts learned in class to create a real system according to the variant chosen. Students will obtain hands on experience in developing a real system using the data structures and algorithms studied and evaluating the system based on the methods for analysis, and the use of Java and Eclipse.

Description: The project involves developing a back-end software application that uses concepts from three to five different concepts learned in the class.

Instructions:

1. Each group should choose their variant of the final project.
2. One variant can be assigned to one group only. In order to select your variant, please put the name of your group in the corresponding column (see online document).
3. Students are allowed to suggest their own variant. In this case, they must submit a project proposal. Students can proceed with their own variant **ONLY** upon receiving a confirmation from their Instructor or GA.
3. All projects must be completed using **Java**.
4. GUI is not required for this project and **cannot be counted as a feature**.
5. Each project should contain the following minimum features (each student implements 1-2 features):

- ✓ Web crawler;
- ✓ HTML parser;
- ✓ Spell checking;
  Spell checking can be achieved by constructing a vocabulary based on all existing words in text files.
  Alternative word suggestions should be provided if no results are found.
  Edit distance algorithm can be used to compare the user's input with existing word from source files.
- ✓ Word completion;
- ✓ Frequency count;
  Frequency count shows the user the number of occurrences of a word in a specific url.
- ✓ Search frequency;
  Ability to show the word that has been searched before as well as the number of times the word has been searched.
- ✓ Page ranking;
  Page ranking is used to measure importance of a search result based on the number of occurrences.
  Search keywords that are repeated more within a web page will be ranked higher than the others.
  Ranking web pages can be performed using sorting, heaps or other data structures.
- ✓ Inverted indexing
  Inverted indexing allows us to perform quick searches without going through all the files. This can be represented as an index data structure storing a mapping from content, such as words or numbers, to its locations in a set of documents.
- ✓ Data validation using regular expressions;
- ✓ Finding patterns using regular expressions.

Final Report & Presentation

The **report** should contain the following elements:

1. Name of the group, group members and your variant (please copy the task to the report).
2. Names of group members and their contribution with the names of java files in your source code.
If there is a feature designed by two or more students, each student **should specify their contribution**.
3. Additional features which students developed (if applicable).
4. Description of data structures/algorithms used for each feature developed.
5. Screenshots showing the demonstration of each feature.
6. References.

As for the **presentation**, students are free to choose any style, but are required to include the following elements:
1. Name of the group, group members and their variant (please copy the task to the presentation).
2. Names of group members and their contribution with the names of java files in the project source code.
3. Additional features which students developed (if applicable).
4. Explanation how students completed the task (including what features they have, how they developed them, algorithms/data structures used).
5. Live demo of their project.
6. References.

Project presentation should be no longer than **7-10 minutes (live demo included)**. Points will be deducted if the presentation is longer than 10 minutes and shorter than 7 minutes. After the presentation each student will be asked questions regarding the project.

Grading Distribution and Deadlines:

Phase	Task	Deadline	Notes
P1	Forming a group	Sec 1: Sept 16; Sec 2: Sept 18; Sec 3: Sept 19.	
P2	Choosing the variant of your final project	Sec 1: Sept 23; Sec 2: Sept 25; Sec 3: Sept 26.	
P3	Milestone Report I*	TBA on Brightspace (approximately Week 5)	Penalty 10% may apply*
P4.1	Final project report	During your class in the 11th and 12th week (300 Ouellette Ave)	
P4.2	Presentation and Q&A		

*** For not submitting their milestone reports, students will be faced a penalty of 10% deducted from their final project total mark.**