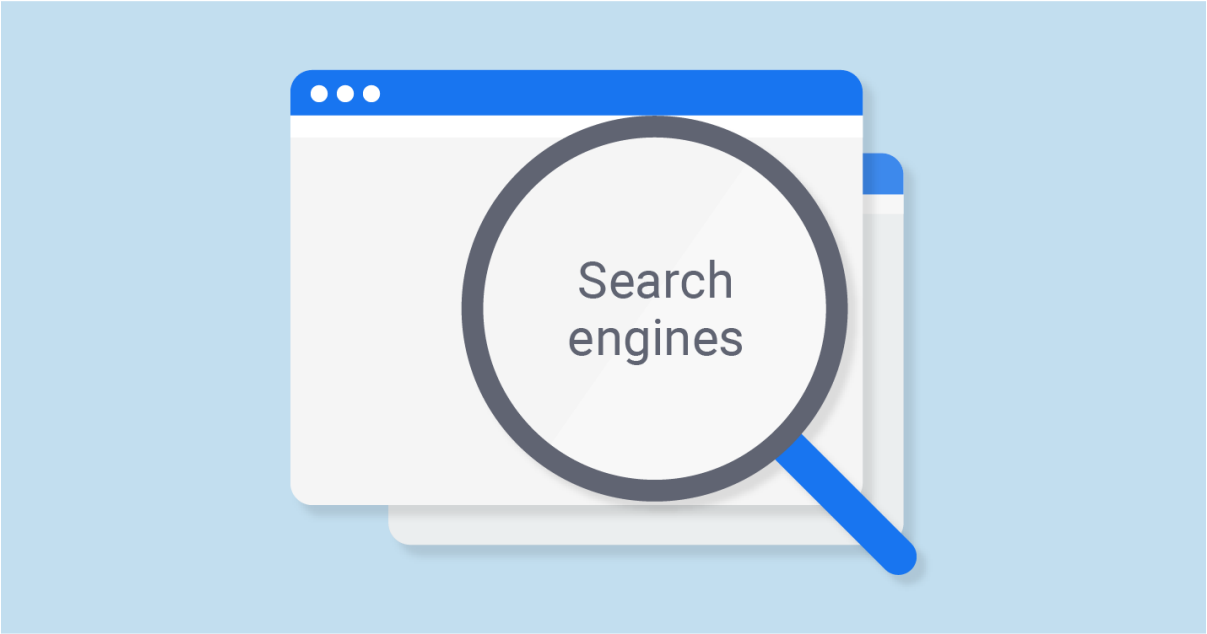
Final Project



Project Description

In this final project, you will be tasked with building a search engine using Python. The search engine should take a user's query and search through multiple files to find all instances of the query. The search results should be sorted and presented to the user in a user-friendly manner. Students are encouraged to choose their own sorting algorithm to implement.

Minimum Requirements

To successfully complete this project, you should break it down into several well-defined components. These components can be implemented step-by-step, and completing each component will contribute to your final grade. Here are the key components of the project:

1. User Input: Implement a function that takes user input for the search query. You can use the input() function in Python to read user input from the command line.

2. File Reading: Develop a function that reads multiple files and extracts their content. You can use the open() function in Python to open and read files. Consider storing the content in a data structure that allows for efficient searching.

3. Query Search: Create a function that searches through the extracted content for the query. You will need to implement a search algorithm to find **all occurrences** of the query in the files.

4. Sorting Algorithm: Design and implement a sorting algorithm to sort the search results based on the criteria of your choice.

5. Presentation of Results: Develop a function that presents the search results to the user in a visually appealing and user-friendly manner. Consider displaying the file names, the line numbers where the query was found, and some context around each occurrence.

Example

Assume we have some files with the according contents:
- file1.txt: "I love programming in Python."
- file2.txt: "There are many different programming languages. For example, Python is a popular programming language."
- file3.txt: "Java and C++ are both widely used."

The user can search a query:
Query: "python"

Search Results:
- Result 1: File: file1.txt, Line: 1, Text: "...languages. For example, Python is a popular…."
- Result 2: File: file2.txt, Line: 1, Text: "Python is a popular programming language."

Going Further

After you made sure that you reached all of the program requirements, you can think about some of the following questions to further improve your project. For example:

- What is the best way to sort the results? For instance, is there a way to sort the results based on relevance (to what the user is looking for)?
- How should the result be presented to the user so that it is more informative and user-friendly?
- What happens if the user's query has a typo? What if the user searches for a phrase (sequence of words) that does not exactly exists in any of the documents, but there are still some roughly similar phrases in the documents?
- Usually, when you search for something on Google, you see something like "*x* results found in *y* seconds". Wouldn't it be cool if your project also has the same thing? How fast is your program? Can it become faster?

Do not forget that a complete project is always much better than a project with tens of fancy features that don't work. ✓

Evaluation

Throughout the project, it's important to write clean and well-documented code. Consider using functions to modularize your code and make it more maintainable. Document your code using comments to explain the purpose and functionality of each component.

During your 5-minute presentation, you will demonstrate your project and discuss how you were able to build it. You will also have the chance to tell us what you have learned from this course and why you think you are ready to start your advanced studies. We are also interested in hearing about your plans, goals, and ambitions with regard to programming. Only, don't forget to **manage your time** properly.

Submission

Please upload (only) the code of your project as a ".py" on Karyar's LMS. The deadline for this project is on the **14th of Mordad** (5th August). The

presentation session will be held on the following day (more information will be given).
Your code will be evaluated based on the file that you will submit, so make sure that you submit the correct final version of your code (we all know that you have another version on your computer that works better…).

Good luck with your project, and have fun exploring the world of search engines in Python!

دوره برنامه نویسی مقدماتی کاریار
بهار 1402
مهبد نوری، عرفان میرزایی