

Team Captain : Rose Nowak - rnowak6

Calina Shaw - ceshaw2

Ayline Villegas - aylinev2

Sarah Menza - semenza2

Nicole Kolbasov - nsk3

1. What system have you chosen? Which subtopic(s) under the system?

Our group has chosen to expand the EducationalWeb System by adding functionalities to improve navigation between slides and uploading/downloading.

2. Briefly describe the datasets, algorithms or techniques you plan to use

Our datasets will include the entirety of the CS 410 slides from coursera. If time permits we may expand the dataset by adding slides from other UIUC courses such as HORT 106, but this is not the main priority of our project.

We will use similar techniques to the current Web of Slides. To connect new slides to the web framework we will use a combination of TF-IDF similarity and word embedding based similarity. We will also introduce some kind of data structure (probably a doubly linked list) to keep track of the student's path throughout the web so they can return to a previous slide.

3. If you are adding a function, how will you demonstrate that it works as expected? If you are improving a function, how will you show your implementation actually works better?

We would like to add a popup or sidebar that displays the most relevant similar documents as you are viewing one set of slides. The relevant documents would be calculated using one of the ranking algorithms.

We would like to add functionality to go "back" to your original slide after clicking related slides on the right hand side. This will be a good feature because that way you can go back to your original lecture after addressing anything you were confused about. We will demonstrate that this works by recording a short demo where we navigate away from the start slide and then hit the back button to return to it.

We would also like to add downloading multiple slides at a time. We will be able to add a "notebook" option where you can add any slides you like, and then download them at the end of your session. We will demonstrate that this functionality works by testing it locally by downloading a group of files.

Students will be able to upload their own pdf slides to perform analysis for them (if runtime permits, we will know more if this is possible after beginning). This could also be checked with local testing by uploading a pdf file that isn't included in the web yet.

4. How will your code communicate with or utilize the system? It is also fine to build your own systems, just please state your plan clearly

Our code will extract all of the Coursera slides and determine which slides are similar to the one we are currently viewing. It will then display the related slides on the side in a pop up or sidebar so you can easily access them. This popup/sidebar will then have an option to be closed or to help you navigate back to the original slides you were viewing.

5. Which programming language do you plan to use?

Python

6. Please justify that the workload of your topic is at least $20 \times N$ hours, N being the total number of students in your team. You may list the main tasks to be completed, and the estimated time cost for each task.

Main Tasks:

- Project Topic Brainstorming Session - 1 hour
- Project Proposal and Team Formation Submission - 3 hours
- Project Setup - 1 hour
- Add algorithm to help find most related documents - 1 hour
- Add pop up or sidebar that displays the results of the algorithm run - 5 hours
- Add functionality to go "back" to your original slide after clicking related slides on the right hand side - 20 hours
- Add functionality to download multiple slides at a time - 20 hours
- Team Check-In Meetings (over Zoom) - 5 hours
- Progress Report - 3 hours
- Add functionality to allow students to upload their own slides - 20 hours
- Documentation - 15 hours
- Improve front end functionality- 8 hours

Total Hours: 102