

Bryson Li (brysonl2)

CS 410 Project **Progress Report**: Intelligent Browsing

Fall 2022

The progress report should give us an idea of how you're implementing your proposal. It should answer 3 main questions: 1) Which tasks have been completed? 2) Which tasks are pending? 3) Are you facing any challenges?

As a reminder: This project that I am proposing is to create an extension to index the current page and allow users to search over the page using a common retrieval function (ie. BM25), possibly also extracting some of the topics from the page [or some additional functionality based on what we have learned in this class].

Here are some tasks that I will need to complete in order to finish the project (Exact details for each of these tasks may vary as I look into them)

1. [3-5 hours] simple proof of concept - python program with python server - scrape from selected webpage(s) and send to local python server
2. [1-2 hour] look into building extensions
3. [5 hours] simple extension (javascript) - take current webpage send to server (local server for simplicity of debugging)
4. [3 hours] add search functionality (if the extension part was done well then this will mostly be for using BM25/search algorithm and formatting the results page, etc.)
5. [5 hours] add additional functionality [ie. topic extraction or sentiment analysis after the user has chosen/ saved multiple articles]

Progress Made

- Worked on/completed the proof of concept - involving a simple python server/ very simple version of the application
- Went through some browser extensions tutorials, added some basic functionality
- Looked into some options for online servers/VMs to eventually host my application (at least for the duration of the project)
- Worked on the extension itself and thought more about the end result, and modified as needed

Tasks that are pending

- Working on finishing up the extension itself
- Doing all the necessary configurations for server to work properly (exact details to be finalized soon)
- Adding remaining functionality which should be relatively straightforward after finishing the core of the extension

Challenges faced

- in general, just learning about extensions and how to write them and also researching some VMs to possibly use