

CS 410 FA22 Project Progress Report

Team: E.S.tension

GitHub link: <https://github.com/yuemingpang/CourseProject.git>

Members	NetID	Role
Yueming Pang	ypang10	Captain
Yuling Gao	yg28	Developer
Schillaci Mcinnis	mcinnis3	Developer

1. Progress made thus far:

We have built the essential structures of chrome extension. Google Chrome extension mainly contains 4 parts, which are manifest, background, content, popup. The manifest records important metadata, defines resources, declares permissions, and identifies which files to run in the background and on the page. Background uses all Chrome APIs, but it cannot interact directly with the content of web pages. Content can execute JavaScript in the context of a web page. Popup contains HTML, CSS, and JavaScript file that controls the logics and appearance of the extension popup.

We have also completed the implementation of web crawler and indexer. This was done mainly in the content.js file, where we used regular expression to find and match the HTML element of email addresses. Once found, all the emails are displayed on the popup.

2. Remaining tasks:

- Implementation of the ranking function to sort the extracted emails based on a query from the user. We may extract emails from multiple websites into our database (chrome.storage API) and rank them using appropriate ranker.
- Software verification and debugging

3. Any challenges/issues being faced:

- We are still working on ensuring 100% extraction of email addresses from the website.
- In order to rank the emails, we need to label each email address. How to label them is a challenge during implementation.
- Picking an effective ranking model to rank emails based on user's query.