CS 410 FA22 Project Proposal

Team: E.S.tension

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

Members:

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

Which topic and why:

The topic we have chosen is the intelligent browsing and we plan to build a chrome extension that can automatically extract email addresses from websites. Very often, it is time consuming to find the target email, especially in a large website. Our email finder will automatically extract all email addresses using web searching techniques and rank them based on the configuration set by the user.

Techniques and algorithms:

The main techniques/algorithms used in our chrome extension are web crawler, web indexing, and an appropriate ranker (raking function such as the Okapi BM25). The Google chrome.storage API will also be used to store, change, and track user data.

Software verification:

The email finder will extract all the email addresses from this site: https://cs.illinois.edu/about/contact-us. Then, rank the email addresses based on user's preferences (for example graduate advising vs undergraduate advising). The performance of the extension will be based on the user's explicit feedback.

Programming languages: HTML, CSS, and JavaScript

Workload justification:

Milestones	Hours for each	Hours * N (N =3)
Build the essential structures of chrome extension (without implementing any algorithms)	5	15
Implementation of web crawler and indexer	10	30
Implementation of the ranking function	5	15
Software verification and debugging	5	15
	Total:	75