

CS410: Text Information Systems

Project Progress Report

List of Total Tasks Done/Remaining

- Fork over Gitlab repo: Done
- Start Building Campuswire Scraper: Done
- Building an inverted index/Retrieving relevant information based on a user's query (Campuswire): Done
- Start Building Coursera Scraper: Done
- Building an inverted index/Retrieving relevant information based on a user's query (Coursera): Done
- Start Building Extension Skeleton: Done
- Finish Building HTML/CSS/JS to display test information: TBD
- Convert Scraper Functions into AWS Lambda Functions, setup API Gateway, to be used as an API: TBD
- Setup the HTTP POST calls from Extension to Lambda's APIs: TBD
- Final Adjustments and Testing: TBD
- Write Final Documentation: TBD
- Write Final Presentation: TBD

We're mostly finished with the project so far, with only the Google Extension and Lambda functions being the main parts of the project that aren't implemented yet. We still have a lot of tweaking and testing to do ahead of us, but we are well ahead of where we thought we were going to be. We aren't going through any major issues at this point, but we suspect that converting the scraper functions into AWS Lambda Functions might be a little tricky to successfully pull off, an essential step that is needed as this

would be the only way the Google Chrome Extension could retrieve data from the scrapers. Other than that, we have some final tweaking and testing to do to make sure everything is working as properly as possible before submission.