

## 1 Completed Tasks

- 1) JavaScript is used for extension frontend: A dummy client is created. There is basic UI to communicate with the Python server. The url of the current page would be sent to the server.
- 2) Python is used for backend server: A web scraper is created to scrap all links on given page, then store the text files on disk.
- 3) Java and library Lucene is used for search engine. It has two parts: Indexer and Searcher. Currently the Indexer is completed, it loads text files from disk and create inverted index for further search.

## 2 Pending Tasks

- 1) The communication chain is not completed, currently the frontend can send the url to Python server, but Python server only respond with placeholder message. Also the Python server does not make use of the Java search engine.
- 2) Java searcher is not completed. Once the searcher can send some meaningful messages back. The frontend UI need to be modified to display these search results.
- 3) A PLSA topic analyzer should be built with Java.

## 3 Challenges

- 1) I did not have JavaScript experience, so it took me some time to create the frontend elements.
- 2) The communication between different parts is tricky. I used WebSocket for JS-Python communication and trying to use Python subprocess library to call Java binary program (still learning).