CS410 Project Final Report: Free Topic – Topic Mining

David Tan Sang Tran (davidtt2@illinois.edu)

https://github.com/davidtt2/CourseProject

Overview of the Function of the Code

This project uses a text retrieval method through Python's pandas & Selenium in order to gain information about the top companies in the technology industry. After retrieving that information, the Python file will generate a json file within the Angular project that will be read and displayed in the user interface. From the user interface, the user can search for technology companies by name.

The expected results from this code is that whenever the Python file is ran, it will obtain the top technology companies for that year. For example, if this code was run in 2021, it will produce similar results with those top companies without any failure or bugs.

Software Implementation (How to Run)

1) Clone project (git clone https://github.com/davidtt2/CourseProject.git)
2) Run "CS410 Project Data.py" (keep file structure unchanged)
    - Requires file to be in same directory as cs410-project
    - Requires chromedriver.exe in same directory as py file
        (also requires Chrome)
        - Different versions can be downloaded at https://chromedriver.chromium.org/downloads
    - Python file will run Selenium webdriver scripts
    - After, it will create a companies.json info file
      at root and in Angular project
3) cd to /cs410-project/src
4) Run the Angular script (ng serve -o)
5) UI will open in browser

Documented Usages for this Project

This project relies on Python and Angular.

Modules that may need to be imported/installed to run:

- npm install @angular/cli
- ng add @angular/material
- pip install pandas
- pip install selenium
- others

Future Goals

The current future goals are to add a dropdown for each company that shows their career websites and have useful information to prospective students looking for a place to join. Besides having this dropdown with extra information, I plan to implement a recommendation system based on the companies that the user has searched for.

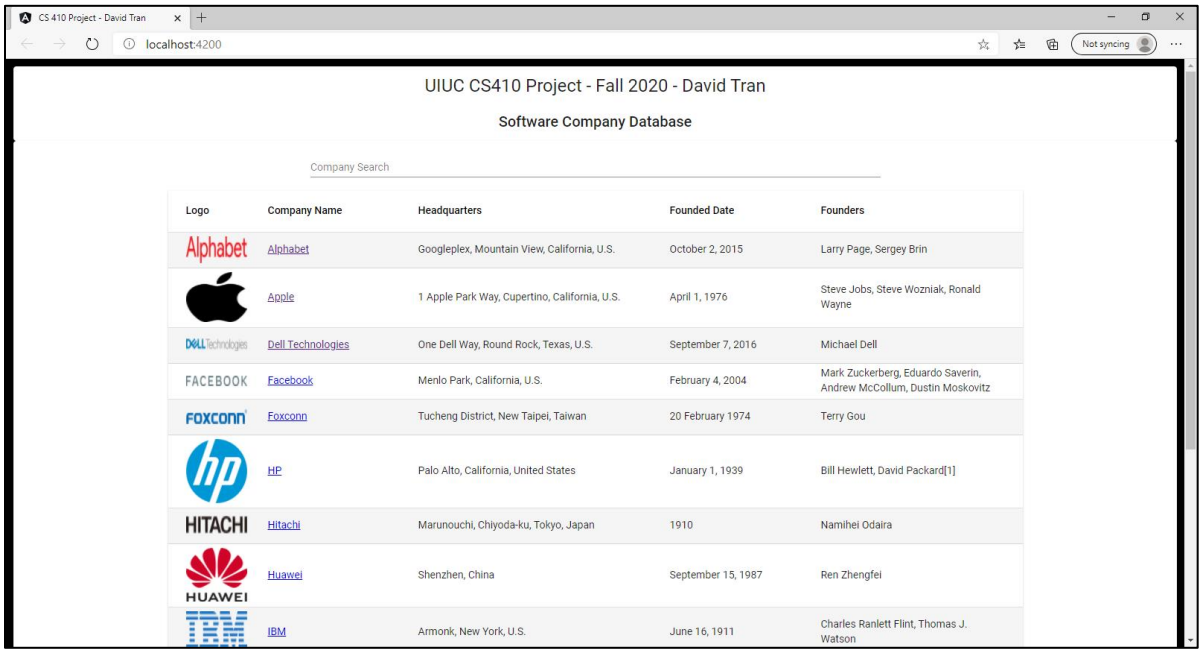
Team Member Contributions

Because this was a single member team, I completed all the work on my own.

Please contact me for any assistance or comments.

Task	**Project Hours**
Research and UI Mockup	5 Hours
Topic Mining to Retrieve Data in Python	15 Hours
Parsing Useful Information from Data Retrieved	10 Hours
Developing the UI in Angular	10 Hours
Connecting the Angular UI with the Data	5 Hours
Testing the UI and Python	5 Hours
Total:	**50 Hours**

Current User Interface



Please contact me for any comments or questions.