

## CS410 Project Final Report: Free Topic – Topic Mining

David Tan Sang Tran ([davidtt2@illinois.edu](mailto:davidtt2@illinois.edu))

<https://github.com/davidtt2/CourseProject>

### ## Description

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.

### ## 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

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.

Please contact me for any assistance or comments.

Current UI:

CS 410 Project - David Tran










localhost:4200

Not syncing

UIUC CS410 Project - Fall 2020 - David Tran

Software Company Database

Company Search

Logo	Company Name	Headquarters	Founded Date	Founders
	<a href="#">Alphabet</a>	Googleplex, Mountain View, California, U.S.	October 2, 2015	Larry Page, Sergey Brin
	<a href="#">Apple</a>	1 Apple Park Way, Cupertino, California, U.S.	April 1, 1976	Steve Jobs, Steve Wozniak, Ronald Wayne
	<a href="#">Dell Technologies</a>	One Dell Way, Round Rock, Texas, U.S.	September 7, 2016	Michael Dell
	<a href="#">Facebook</a>	Menlo Park, California, U.S.	February 4, 2004	Mark Zuckerberg, Eduardo Saverin, Andrew McCollum, Dustin Moskovitz
	<a href="#">Foxconn</a>	Tucheng District, New Taipei, Taiwan	20 February 1974	Terry Gou
	<a href="#">HP</a>	Palo Alto, California, United States	January 1, 1939	Bill Hewlett, David Packard[1]
	<a href="#">Hitachi</a>	Marunouchi, Chiyoda-ku, Tokyo, Japan	1910	Namihei Odaira
	<a href="#">Huawei</a>	Shenzhen, China	September 15, 1987	Ren Zhengfei
	<a href="#">IBM</a>	Armonk, New York, U.S.	June 16, 1911	Charles Ranlett Flint, Thomas J. Watson