

FAANGMULA Job Search Engine

1 TEAM MEMBERS

The team members are Yang Pan (yangpan3), Wenjun Peng (wenjunp2), and Tony Mu (tongm3). Tony is the captain of the team.

2 PROJECT OVERVIEW

We want to implement a “FAANGMULA Job Search Engine”. It is a search engine that indexes and retrieves relevant FAANGMULA job links based on a given keyword query. There are existing job search engines that provide similar functionality. Some examples are Indeed, Linkedin, Angelist, etc. However, these job search engines are not specified in any fields, such as software engineer jobs. In other words, it can be hard to filter and find relevant information from such generic search engines. The indexed links also suffer from staleness. Our project aims to solve these issues by providing targeted job links and updating indexes frequently (daily).

3 HIGH-LEVEL DESIGN

We are planning on implementing the system in three separate parts. First, we will implement a scraper that scrapes FAANGMULA job sites and downloads job description pages. It will also parse them into tokens, and output a CSV file that serves as a 1:1 mapping between the tokens and a web link as an ID. We will develop this scraper using python and metapy. Second, we will implement an information retrieval system that takes in a keyword query from users, and outputs a ranked list of relevant job links based on the job descriptions. We will implement this system using MetaPy. Lastly, we will implement a simple front end view that takes user queries and displays the list of relevant links. We will implement this part of the system using Python and React. We will evaluate the outcome of the project by user input (explicit feedback).

4 TASK BREAKDOWN

We estimate the total effort required to be 70 hours, where

- Part 1 - scraper and tokenizer will take about 30 hours to complete
- Part 2 - indexer and ranker will take about 30 hours to complete
- Part 3 - front end integration will take about 10 hours to complete