A Comprehensive and Subjective Job Searching Recommendation System

Presenter: Di Liu
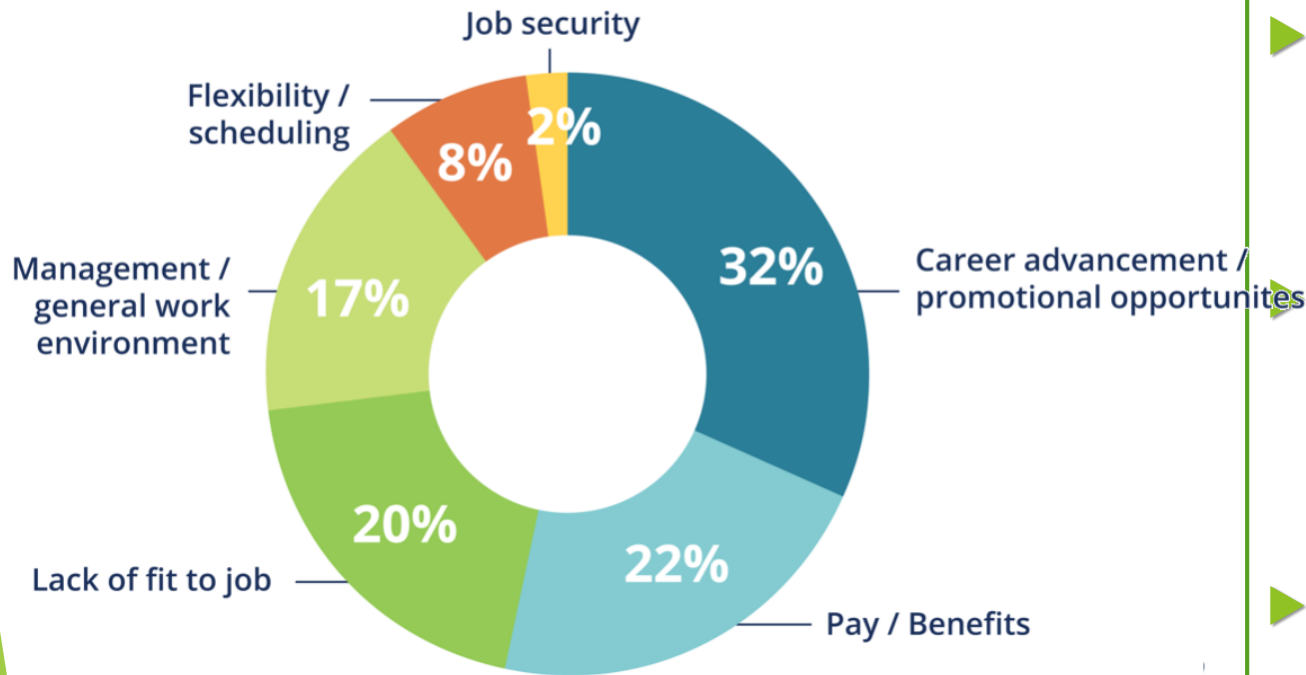
Ph.D. in Mechanical Engineering

Virginia Commonwealth University, Richmond, Virginia

2019/2/14

Motivation

TOP VOLUNTARY JOB-RELATED REASONS FOR CHANGING JOBS

Job security 2%

Flexibility / scheduling 8%

Management / general work environment 17%

Lack of fit to job 20%

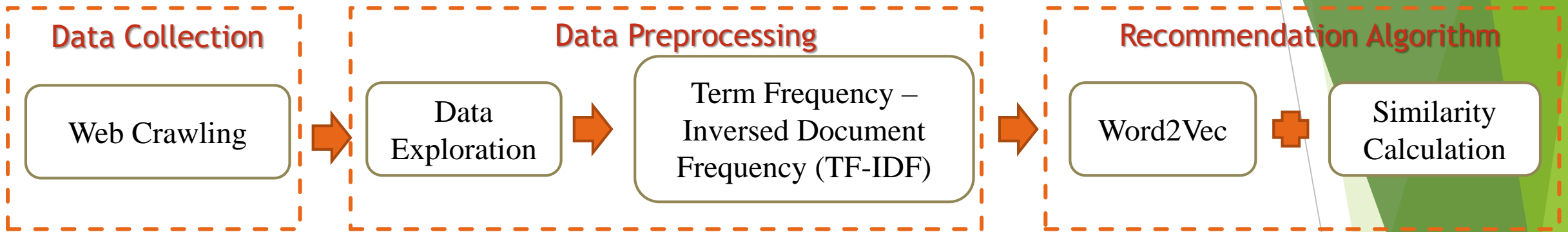Pay / Benefits 22%

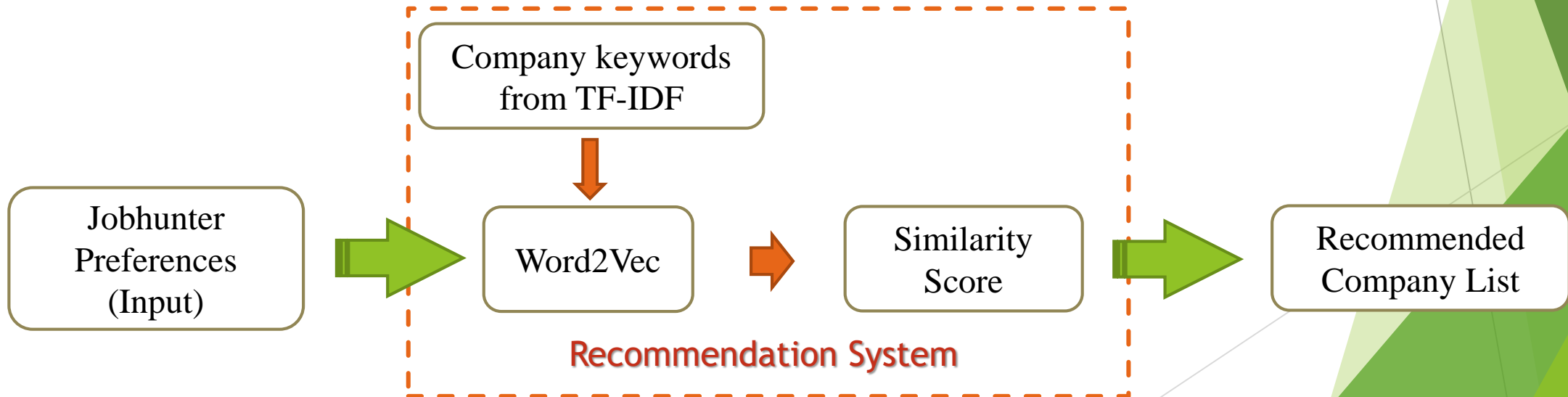Career advancement / promotional opportunites 32%

Source: Gallup
(http://myaliat.com/news/employees-mentally-checked-out/top-voluntary-job-related-reasons-changing-jobs-gallup-aliat/)

▶ **Frequently changing jobs has adverse effects on both employers and employees themselves.**

▶ **About 80% reasons are related to the features of the companies, such as culture, value, working environment, etc.**

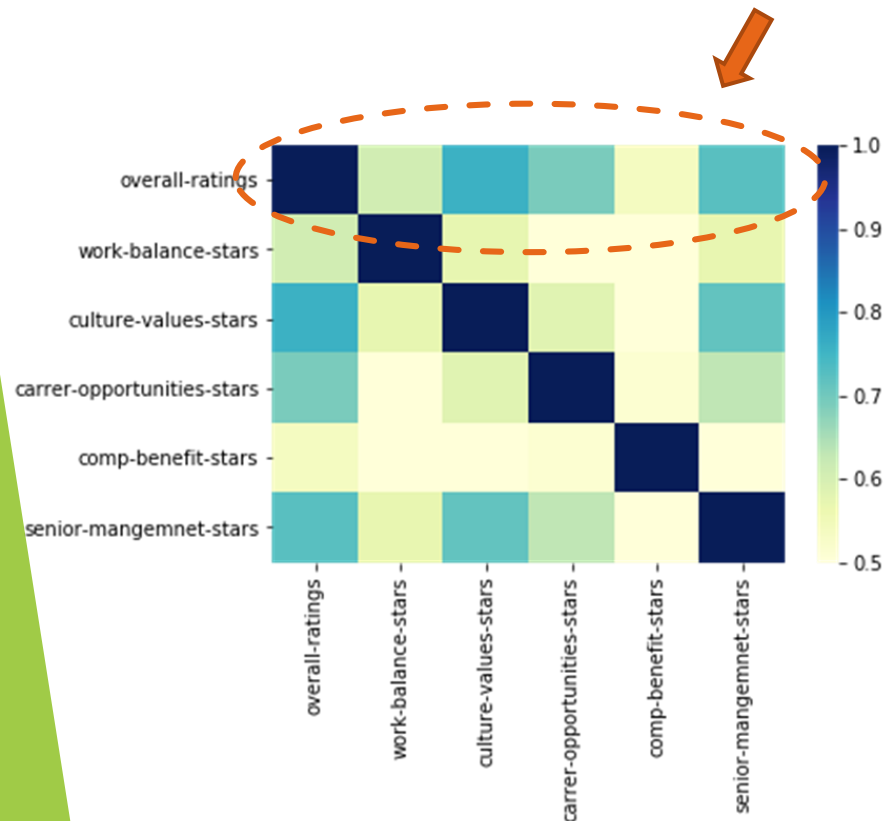▶ **Comprehensive and subjective understanding of the companies during the job searching process is important.**

How do I build the system?

Data Collection

Web Crawling

Data Preprocessing

Data Exploration

Term Frequency – Inversed Document Frequency (TF-IDF)

Recommendation Algorithm

Word2Vec + Similarity Calculation

How to use the system?

Company keywords from TF-IDF

Jobhunter Preferences (Input)

Word2Vec

Similarity Score

Recommended Company List

Recommendation System

Takeaways so far...

web crawling → Data Exploration → Term Frequency – Inversed Document Frequency (TF-IDF) → Word2Vec + Similarity Calculation

* Darker color indicates higher correlation between the two features.
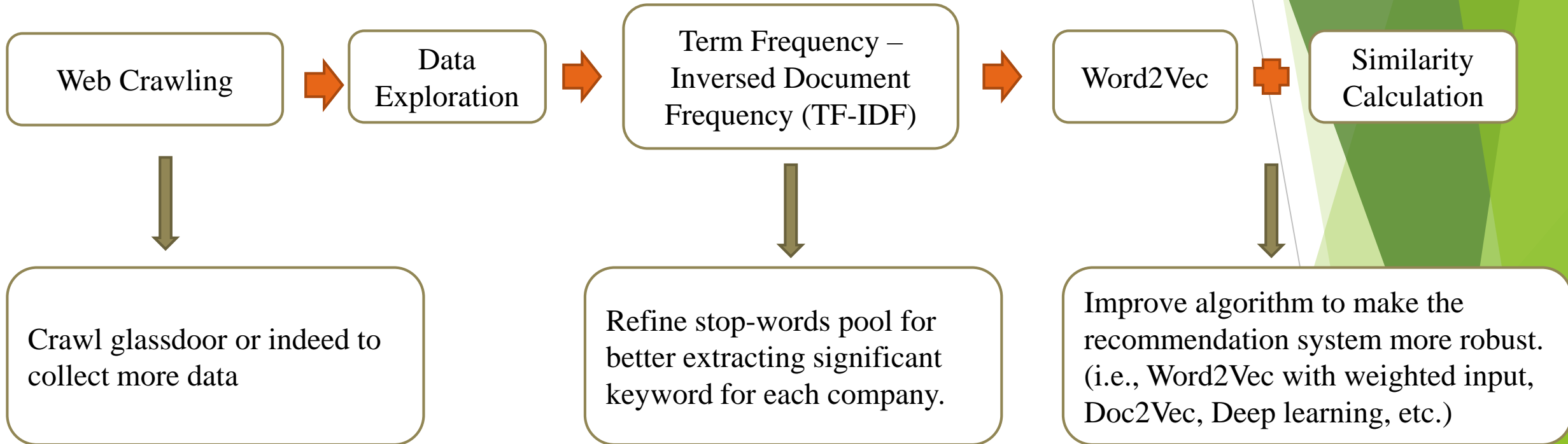

> ➤ **The overall ratings are more related to the company culture and management rather than 'money'.**

> ➤ **Wordclouds including the keywords have been plotted for companies.**

On-going Work…

Web Crawling	→ Data Exploration	→ Term Frequency – Inversed Document Frequency (TF-IDF)	→ Word2Vec	+ Similarity Calculation

Crawl glassdoor or indeed to collect more data

Refine stop-words pool for better extracting significant keyword for each company.

Improve algorithm to make the recommendation system more robust. (i.e., Word2Vec with weighted input, Doc2Vec, Deep learning, etc.)

*There is no **Mr.Wrong**,*

*but **Mr.Right** in the **WRONG** place!*

Thank you!
& Question?

* For more information, please refer to my github link: https://github.com/liudi1025/TDI