



Case Study On Indeed

BY

JAEMON THYPARAMBIL THOMAS

STUDENT ID : 29566428

What is indeed

- ▶ One among most popular site for job portal
- ▶ Spread across 60 countries
- ▶ Available in 28 different languages

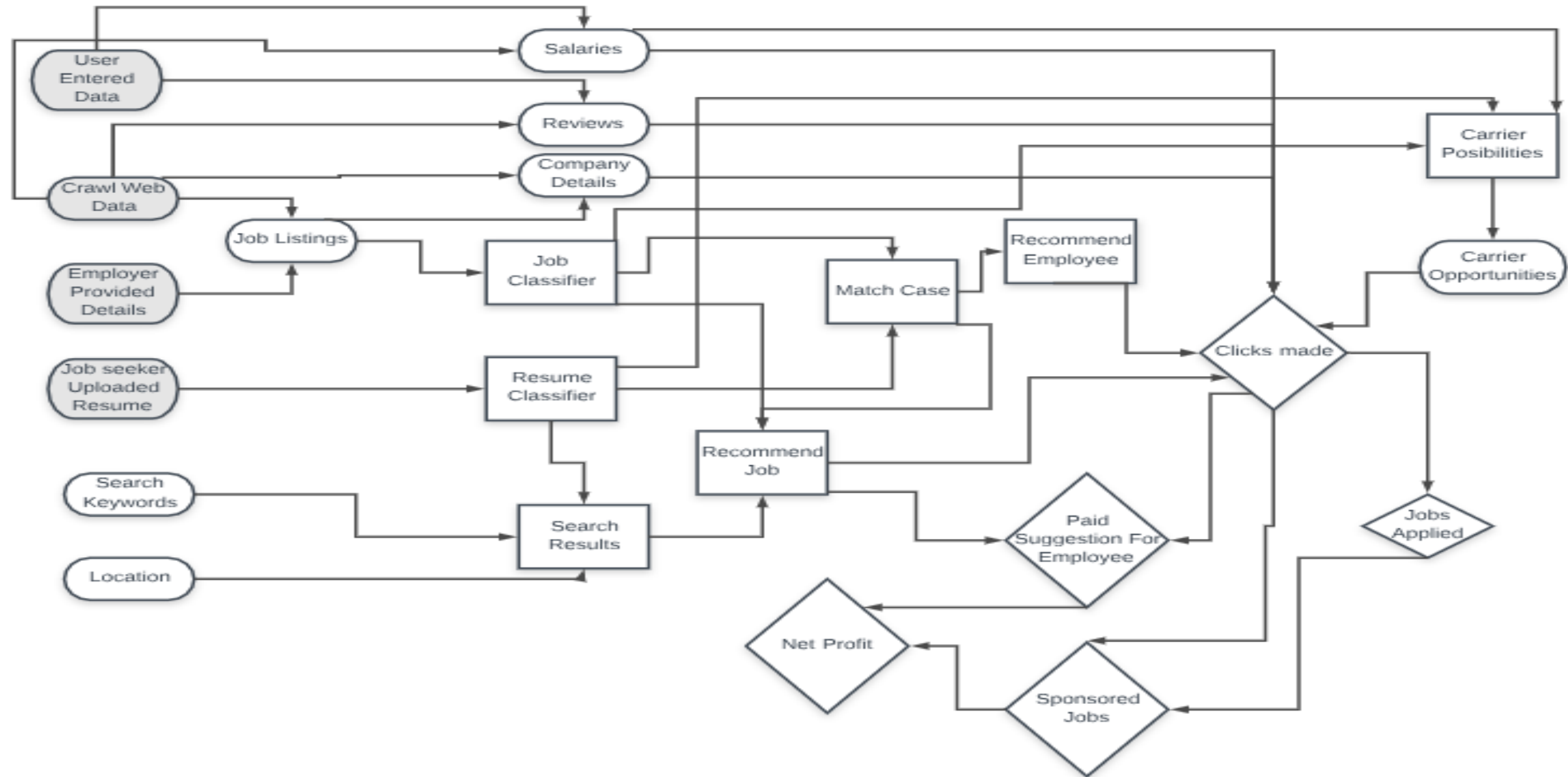
Why I choose indeed

- ▶ Data Driven
 - ▶ 200 million unique monthly users
 - ▶ 9.8 jobs added per sec
- ▶ Data Roles
 - ▶ Data Scientist
 - ▶ Data Engineer
 - ▶ Product Scientist
 - ▶ Business Intelligence Analyst

Data Characteristics

- ▶ Volume
 - ▶ 100 million resumes
 - ▶ 72 million company ratings
 - ▶ 500 million salaries
- ▶ **Velocity**
 - ▶ 200 Million unique monthly visitors
 - ▶ 9.8 jobs added per second globally
- ▶ **Variety**
 - ▶ Data in pdf, Photos, Videos etc
 - ▶ Portability in 28 different languages
- ▶ **Veracity**
 - ▶ Precision - eg architect (building or software)
 - ▶ Recall – eg hr and human resource
 - ▶ Tokenizer issues due to different language

Data Model



Challenges faced

- ▶ Scaling into different countries
- ▶ Job seeker might not know correct job description
- ▶ Detecting Language and trying to tokenize data
- ▶ Localisation may not provide correct data

Questions



THANK YOU