



RIGHT HIRES!

Talent Forecasting

The Right People, The Right Skills, The Right Jobs

Team 52 Final Report Executive Summary

Julia Barnhart, Brennen Chadburn, Jonathan McKim, David Pilkington, Mike Ryder

Problem Statement

Finding relevant applicants is increasingly difficult with the diversity of Data Science related jobs. Trouble hiring ill-qualified applicants are further exacerbated by varying skill sets, emerging technologies, and a general misunderstanding of market trends. Thus, employer job titles and required skills can quickly fall out of sync - leading to a lack of quality applicants and employee turnover. The result is this business landscape.

Figure 1: Business Landscape



Human Resource (HR) groups, hiring managers, headhunters, current/potential employees, students of data science, VC firms, and executives looking to determine the location of new Data Science groups are affected by job role ambiguity. These audiences look to better identify, locate, and budget for data science talent. These needs make up the RightHires product goals.

Competition and Potential Partnerships

RightHires reviewed the market for possible competitors and found two complementary offerings. The first is one from People Insights built for Walt Disney Company. The People Insights product is inward focused and provides a company with insights on retention problems for exiting employees. For example, what are the characteristics of those that might leave and the probability they will leave. This information can help a company change pay, environment, or



flexibility to lower the probability an employee will leave. Jobs covered are well defined and stable.

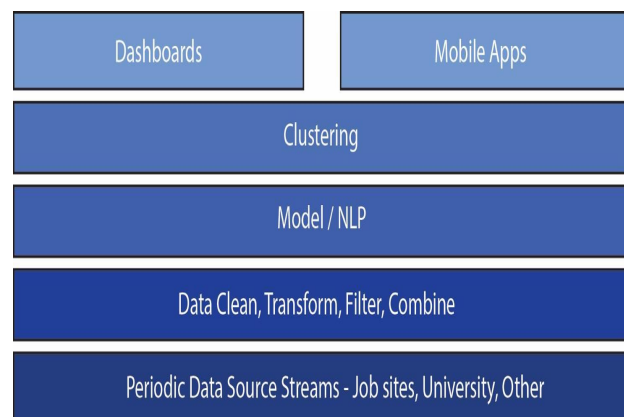
The RightHires offering has a market (external) focus in a quickly evolving industry (Data Science) where jobs and their requirements vary as new technologies become available. RightHires serves multiple internal and external facing customers including HR, VCs, executives, new hire candidates, and students. The product is used early in the hiring process to define a job for a successful outcome. Overall, the RightHires product is targeted to make the hire, find a job, find education process more efficient. The offering covers the US. Last, RightHires is subscription based, so constantly adding new value is important for keeping subscriber revenue growing.



The second competitor Aspiring Minds (www.aspiringminds.com/) is an AI driven tool to help screen candidates once a job is posted and before a candidate becomes an employee. Aspiring Minds provides on-line job simulations and coding evaluations as well as video assessments. While all three products don't overlap and are not direct competitors, each doesn't solve the full life cycle for our collective customer base. With that in mind, a partnership should be considered to offer customers a full suite of functionality. Workflow would move in the order shown above; customers begin with RightHires, move to aspiringminds, and once someone is hired to People Insights.

Proposed Solution

RightHires periodically scans U.S. job market postings and uses Natural Language Processing (NLP) and clustering to determine the most prevalent data science related job titles, demand by region, groups of related jobs, equivalent titles, the most important skills per job, and universities that offer master level data science programs. Results will be available for easy access through dynamic





dashboards and mobile applications to help organizations attract relevant job applicants with the required skill set and keep up with a rapidly evolving Data Science job market. Having the capability of comparing job titles, descriptions, and requirements to current market norms will drastically reduce the time and costs associated with building a Data Science team and reduce the problems highlighted in Figure 1. The remainder of this document summarizes RightHires conclusions, user experience, and recommendations.

Platform Conclusions (Data, Model, Cluster)

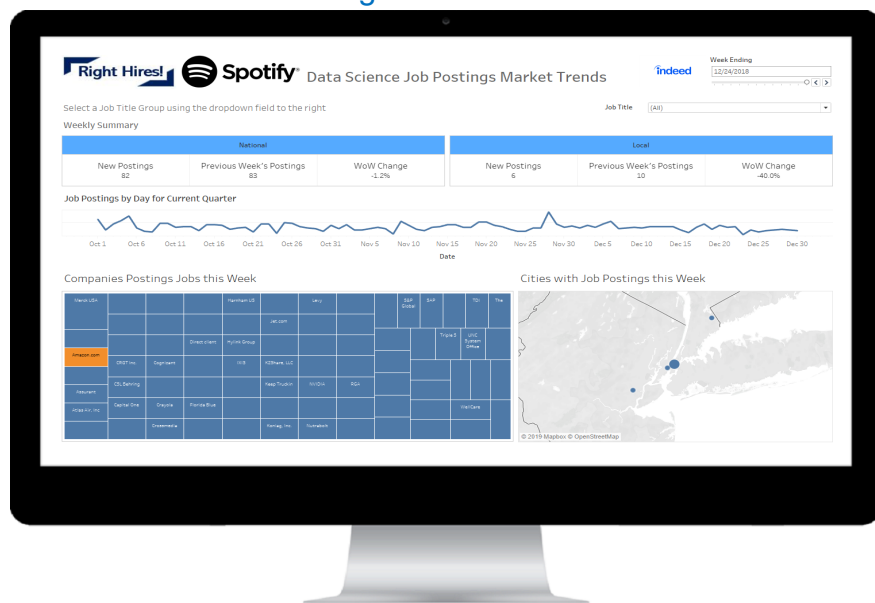
The result of the data sourcing, data preparation, multiple modeling, clustering and refinement processes (fully described in the Final Report) are 3 high level clusters that contain jobs targeted for the initial release of RightHires. Characteristics of those clusters are shown below.

Cluster	1	2	3
Job Classification	BI & Analyst	Data Science & Data Engineer	Healthcare
% Covered	41.5%	31.1%	27.4%
Alternative job titles in cluster	bi analyst data analyst senior bi analyst tableau developer data engineer senior data analyst senior tableau dev business data analyst bi analyst ii data scientist	bi analyst data scientist data analyst senior bi analyst data engineer senior data scientist senior data analyst senior data engineer marketing data analyst bi analyst ii	data analyst bi analyst bioinformatics scientist data scientist director biostatistics research data analyst bioinformatics specialist senior bi analyst healthcare data analyst scientist bioinformatics
Top terms and skills in order of importance	Report, requirements, manage, development, work process, BI, knowledge management, Tableau, SQL, analyst, design, performance	Machine, learning, design, engine, python, analyze, model, big data, technology, innovate, data scientist, drive, make, build, insight, market, passion	Research, clinic, patient, public, population, care, applications, study, strategist, state, medical, investigate, requirements, staff, federal, scientific, education

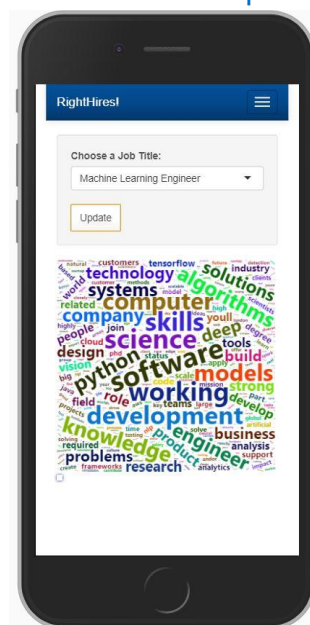
The User Experience

Insights driven through our model/clustering are made available through an interactive web and mobile application experience for two user groups. These are people looking for Data Science jobs, and those looking to hire them. RightHires plans to focus the web experience towards a corporate audience and the mobile for people looking for a job. Cross offerings will be tested in a Phase 1 pilot. A sampling of the functionality available through each interface is shown below (more in Final Report and CEO demo).

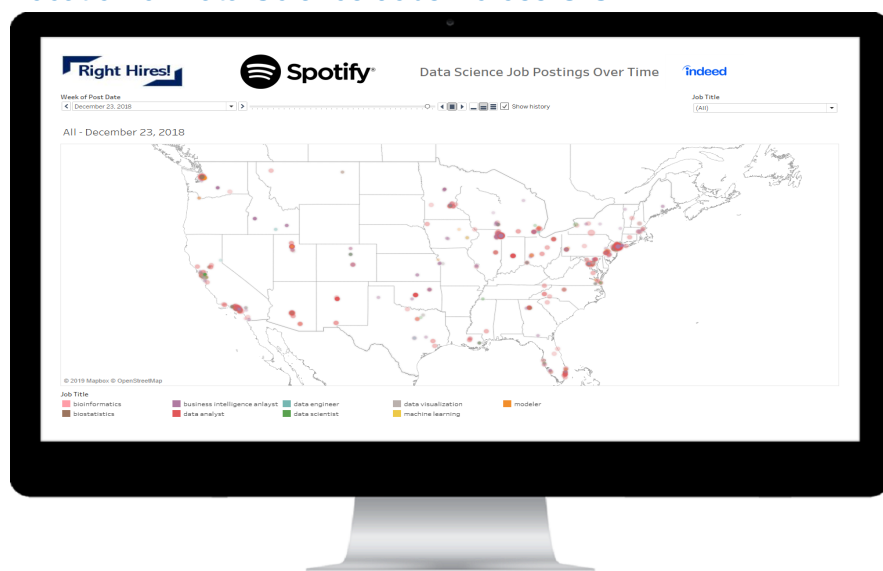
Data Science Job Posting Trends



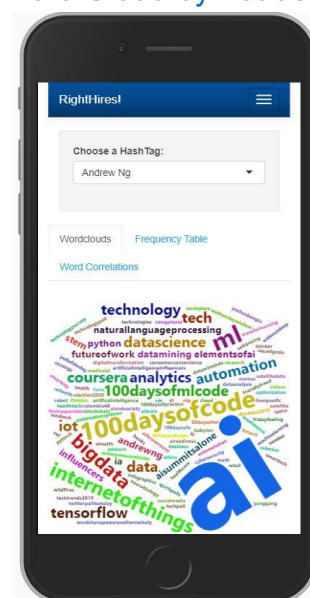
Job Word Map



Location of Data Science Jobs Across U.S



Word Cloud by Leader



Recommendations – Phase 1

During this project the team has proven out strategies around business need, technology selection for modeling/clustering, jobs to target, data sourcing options, multi-user interfaces, and infrastructure required to support demand. As a result, the team makes the following recommendations to the CEO for taking RightHires into the market (Phase 1) and extending the offering (Phase 2). The figure below shows a sequence of steps for Phase 1. A set of Phase 2 recommendations is also summarized that would add more value, increase revenue potential, and expand the RightHires customer base.



1. **Create Marketing/Pricing Materials** – Leveraging findings from the problem statement, environment stats, goals, findings work done to date.
2. **Identity Pilot Customers** – for the targeted spaces, namely HR, Headhunters, VC firms, job candidates, and students.
3. **Evolve Data Sourcing** – Move data acquisition in house by partnering directly with job posting websites to obtain API access or invest more in our data scraping capabilities for use across additional job posting websites.

4. **Put Model/Cluster into Service As Is** – The tested design that leveraged data cleaning steps, TF-IDF, Word2Vec and T-SNE, and K-Means has worked well for the jobs targeted. Put those in production for the pilot.
5. **Infrastructure Build Out** – Implement the cloud Infrastructure as specified to handle the workload of running the model and clustering on the latest data every week.
6. **Launch Dashboards** – Target dashboards built to date for our business customers while collecting feedback over the pilot to provide insight into required enhancements before General Availability (GA).
7. **Launch Mobile** – Target mobile application for all pilot customers. This is being done to test the original theory that mobile is preferred by non-business users. Feedback collected will be used to tweak the offering (and user population targeted) for GA.

8. **Run Pilot** – Run pilot, collect feedback through UIs and user groups, use the feedback to determine next steps before GA.
9. **Prioritize Phase 2** – Over the above steps, gather data to help prioritize and enhance the list of Phase 2 recommendations. Use that insight to determine the sequence and content of Phase 2 potential solutions.

Phase 2 Recommendation Summary

Automatically
generate job
postings



Voice Mining and
Analysis



Add Salary



Expand Education
Info and
Partnerships



Expand base
functionality through
Partnerships



Expand Job Listing
sources in US and
other countries



Cross Web and
Mobile Functionality



Full description of
Phase 2
recommendations is
available in the Final
Report.