



Leading House « DUAL-T » Technologies for Vocational Education

# Detecting latent training needs from digital traces

Phase IIIb  
EPFL CHILI  
**Fu-Yin** Cherng

# Outline

- Motivation
- Literature Survey
- Research Projects
  - 1st and 2nd semester

ty law scotland wales northern ireland education media  
**Five ways work will change in the future**

15 jobs that are quickly disappearing  
STRATEGY

## Why we need to know training needs?

As the technology evolves, the demand for the job market and the required skills change rapidly.

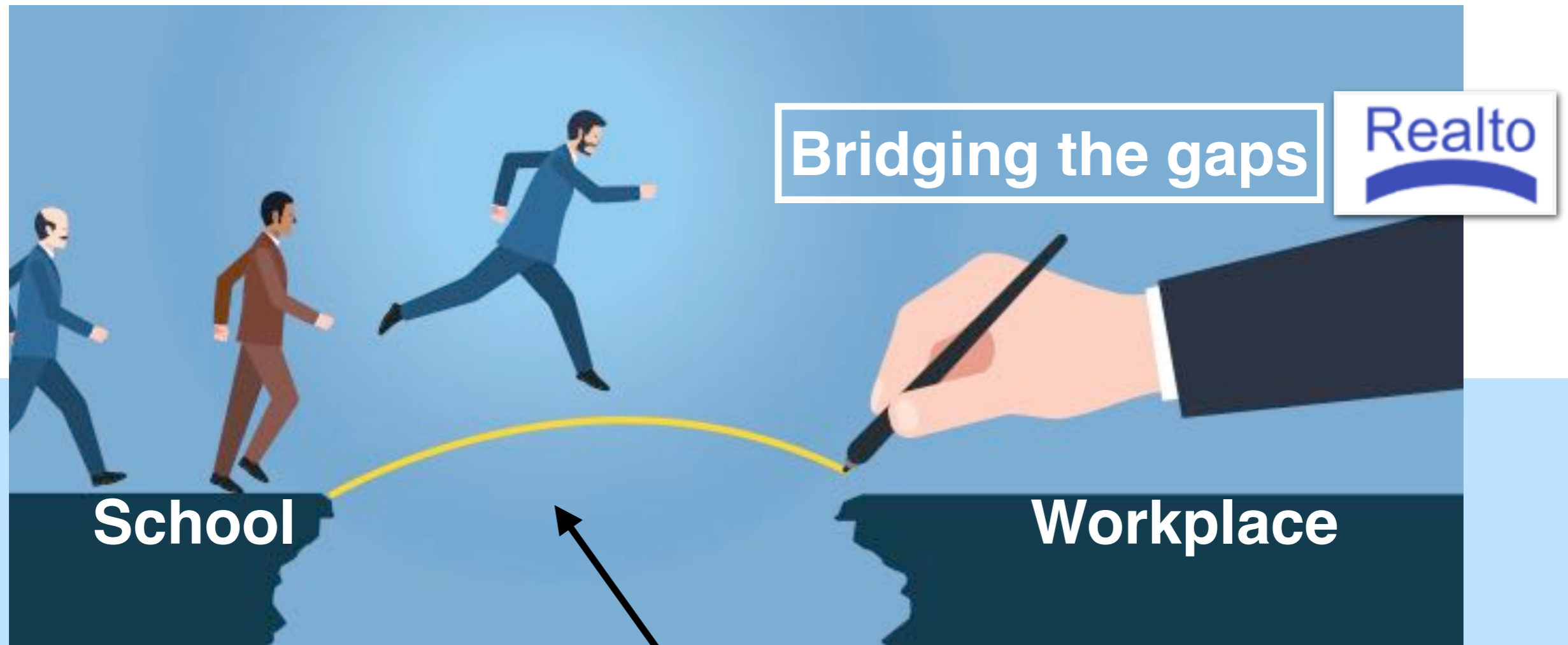
Identify training needs by survey or meeting

Set up new training offers

Take several years!

Quicker way to identify potential training needs

# Skill Gaps



**Finding those gaps by data science and machine learning**

**NETFLIX**

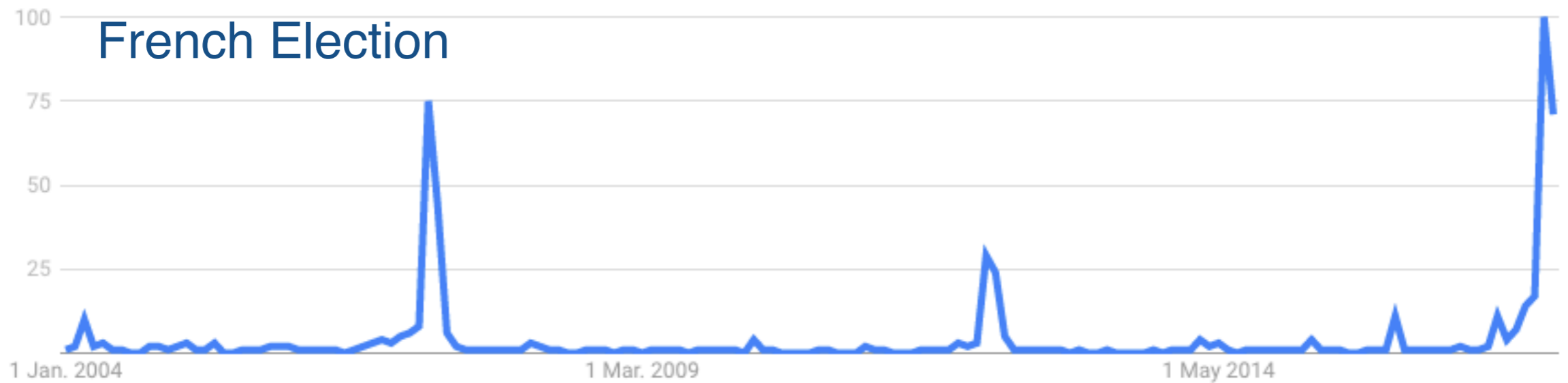


**amazon**



Understand customers opinion and interests

# Google Trends



# Mining Job Skills

- **Snapshot** data of job ads

**MONSTER**

**indeed**<sup>™</sup>  
one search. all jobs.

- Information technology (IT) job skills

- Compare importance of skills by their term frequency in job ads.

[Chuck, L. et al., IEEE 2010]

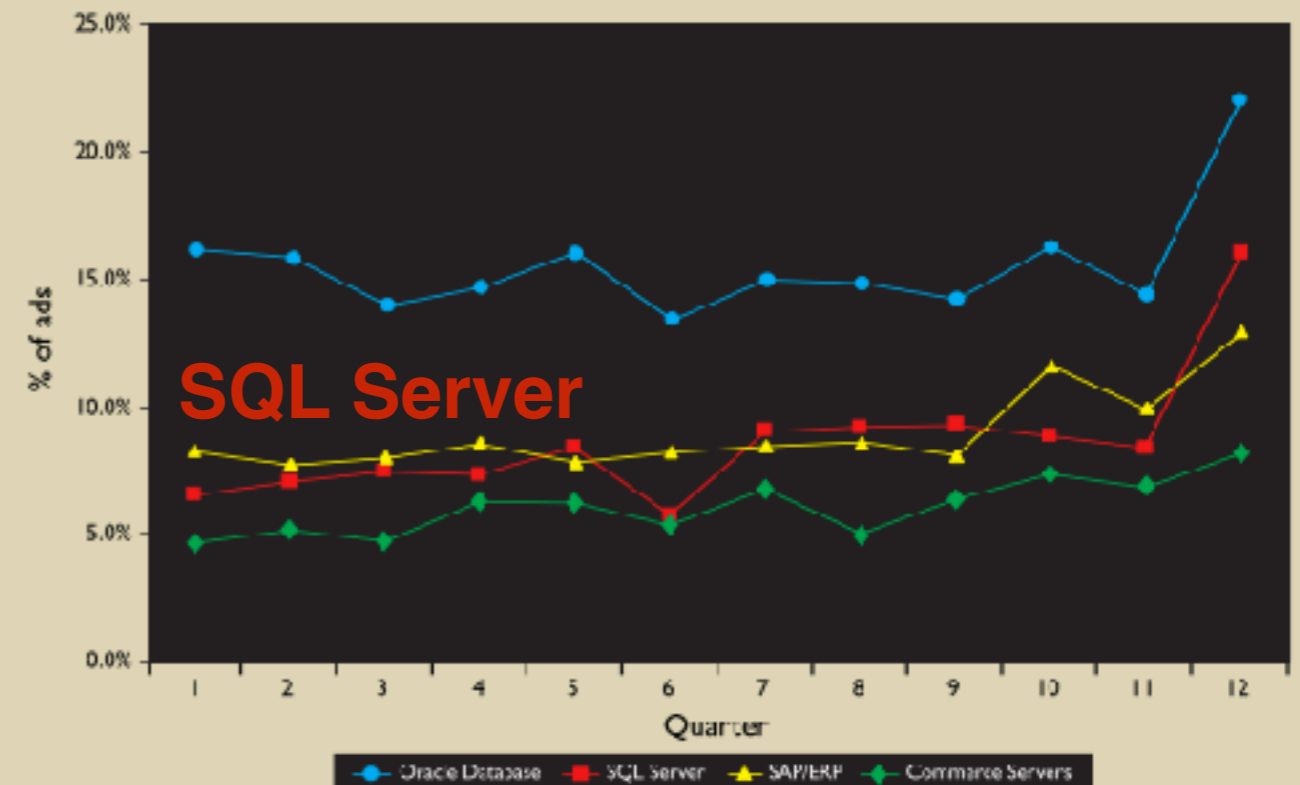
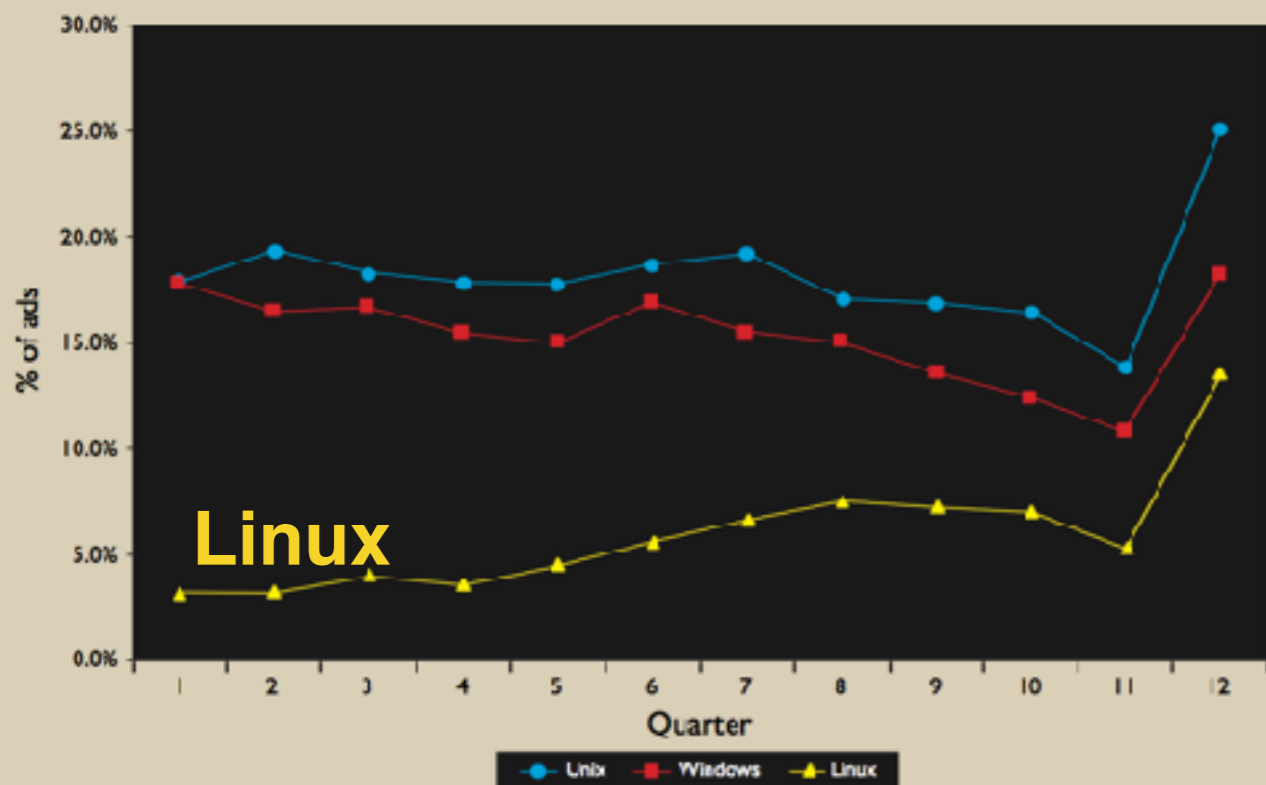
Skill	Frequency (%)
Security	33.29
C/C++	28.69
SQL	27.57
Programming	26.08
Microsoft operating systems	23.18
Java/Java 2 Enterprise Edition/Java to Python	21.09
Leadership	20.10



# Longitudinal Analysis

- Collected IT job ads over **3-years** period (**2002-2005**) from online job board
- Compare how the % of each skill change over time

Prabhakar, B. et al., "IT skills in a tough job market." Communications of the ACM 48.10 (2005)



*Operational system*

*Database*

# Skill Gaps

- Questionnaire: **perceived** skill gaps  
[Aken, A. et al., SIGMIS 2007]
- **Mismatch** of perceived importance between **student** and **employee**
- Compare skill sets from **job ads** and **curricula** of related disciplines. (e.g., ACM Curriculum Committee)

[Richard, W. et al. Higher Education Theory and Practice 2016]

What kind of data can reflect the training needs?

# Digital trace about training needs

## Q&A sites

**course**ra

Stack**Exchange** 

# Digital trace about training needs

## Q&A sites

**course**era

Stack**Exchange** 

## Job Ads

**indeed**<sup>TM</sup>  
one search. all jobs.

**MONSTER**

# Digital trace about training needs

## Q&A sites

**course**era

**StackExchange** 

## Job Ads

**indeed**<sup>TM</sup>  
one search. all jobs.

**MONSTER**

## General Interests

**Google** Trends



**WIKIPEDIA**  
Die freie Enzyklopädie

# Digital trace about training needs

## Q&A sites

**course**ra

**StackExchange** 

## Job Ads

**indeed**<sup>TM</sup>  
one search. all jobs.

**MONSTER**

## General Interests

**Google** Trends



**WIKIPEDIA**  
Die freie Enzyklopädie

# Finding Question Topics in MOOCs Forum by Using Stack Exchange as Training Data

2016 fall



# How do I improve LMS / NLMS filter performance?



14



Are there ways to increase computational performance of a normalized least squares (NLMS) filter? **Multidelay block frequency-domain (MDF)** filters have been proposed to do this, but they also take away from convergence speed and accuracy, because they only update the estimated impulse response once every block, not once every sample. Are there any other methods?

adaptive-filters

performance

least-squares



share improve this question

1

edited Jun 6 '12 at 13:43



Jav\_Rock

693 2 7 22

asked Aug 16 '11 at 19:13



Phonon

3,010 2 20 56

predict tags of questions in MOOC Forum

tags

which **topics** that students have the most questions of?

**Training needs**

# SIGNAL PROCESSING

Posts: 10,105  
Date: 08/2011 - 09/2016

Signal Processing Stack Exchange is a question and answer site for practitioners of the art and science of signal, image and video processing. Join them; it only takes a minute:

Sign up

## Explore Our Questions

- image-processing
- matlab
- fft
- filters
- discrete-signals
- fourier-transform
- frequency-spectrum
- computer-vision

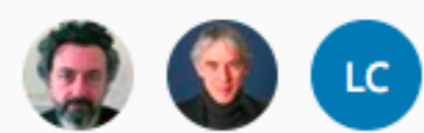
0 votes    2 answers    166 views    Improvement of Minimum description length

linear-algebra

0 votes    0 answers    5 views    Why do linear phase filters have sy

# Digital Signal Processing

Posts: 5,910  
Date: 02/2013 - 03/2015



Welcome to Digital Signal Processing! You're joining thousands of learners cur

More



# Challenge

Whether the posts of Stack Exchange and MOOC forum are **similar**?

**course**ra  
signal processing

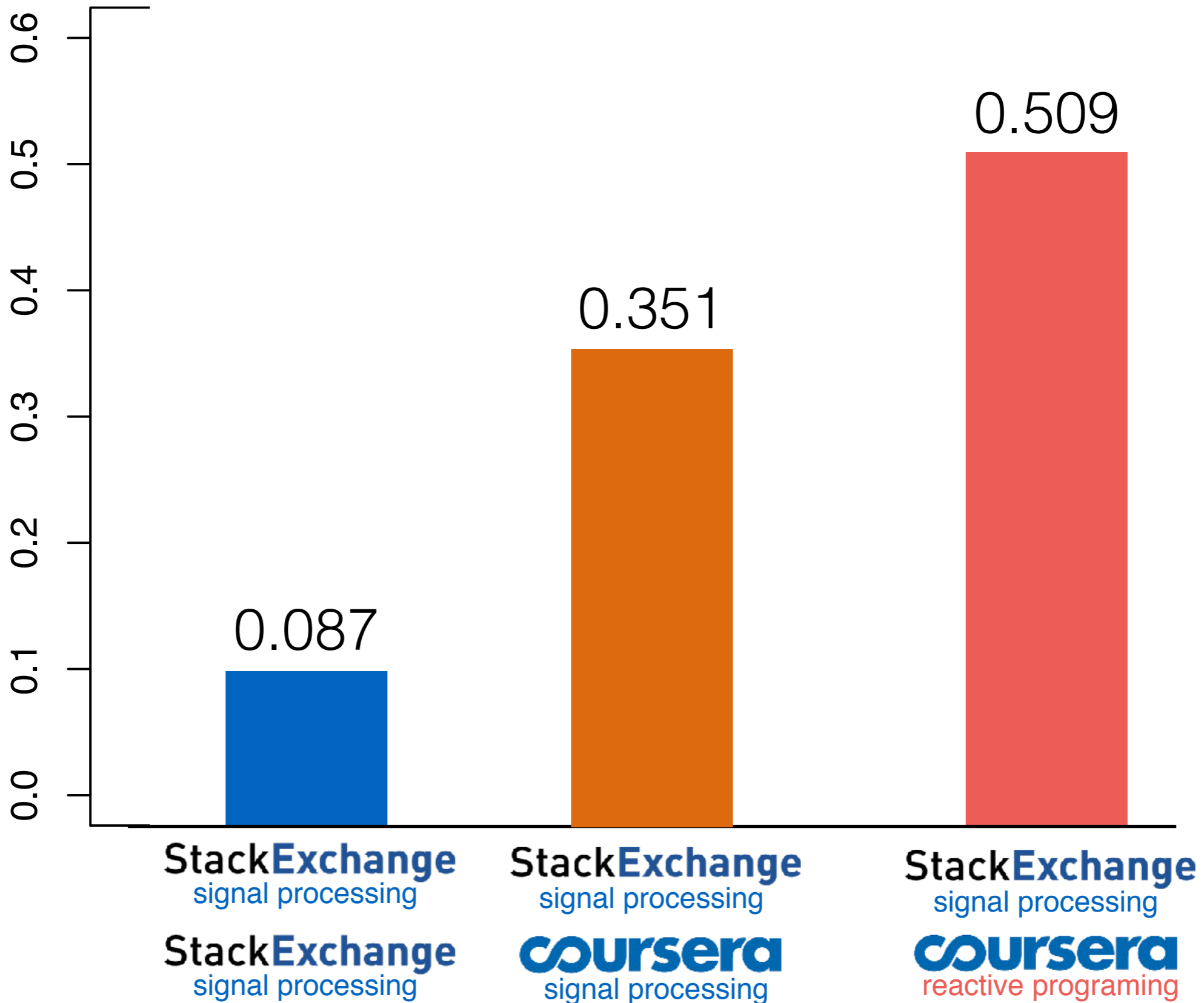
**StackExchange**  
signal processing



# JS divergence

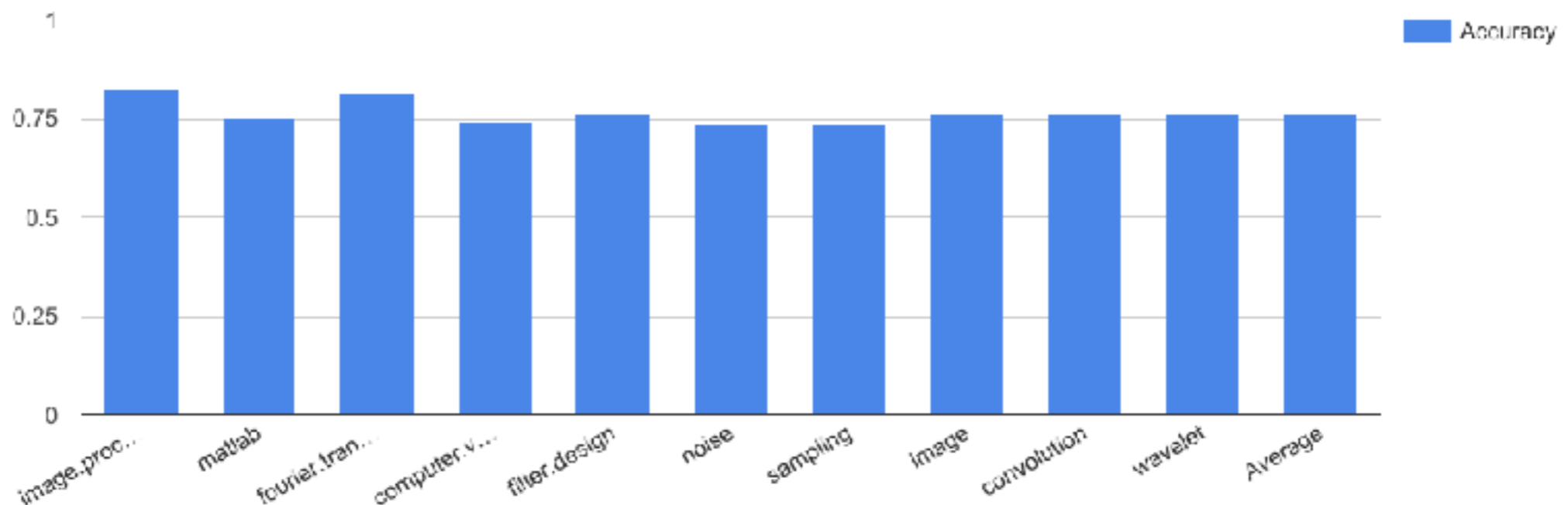
Divergent

Similar



# Summary

- Two data sets are **not similar enough to** predict tags across platforms.
- In the context of the same platform, we can reach **75% accuracy** of predicting tags of new posts in Stack Exchange.



# Digital trace about training needs

Q&A sites

**coursera**

StackExchange 

**Job Ads**

**indeed**<sup>TM</sup>  
one search. all jobs.

**MONSTER**

**General Interests**

**Google** Trends



**WIKIPEDIA**  
Die freie Enzyklopädie

# Online Job Ads Analysis

2017 spring  
work in progress



Find Jobs Find Salaries Company Reviews Find Resumes Employers / Post Job

**indeed**

**what**  **where**

job title, keywords or company city, state, or zip

**engineering jobs in San Francisco, CA**

Upload your resume - Let employers find you

Show: **all jobs** - 1,019 new jobs

**Entry Level Patent Agent or Scientific/Technical Analyst (SF... - new**  
Morrison & Foerster - ★★★★★ 12 reviews -  
San Francisco, CA 94105 (Financial District area)  
Life sciences (grad level preferred), chemistry (grad level preferred), electrical engineering, computer science, physics, and mechanical engineering ...  
1 day ago - save job - more...

**Operations Engineer**  
Slack - ★★★★★ 2 reviews - San Francisco, CA  
It is based in San Francisco and reports to one of our engineering managers. At least years operations engineering experience, a large portion of which in a...  
6 days ago - save job - more...

**Design Engineer - new**  
Autodesk - ★★★★★ 171 reviews - San Francisco, CA  
Design Engineer 17WD23536 Americas United States of America California San Fran  
Description Principal Duties and Responsibilities Note: \* Lower  
1 day ago - save job - more...

**Engineering Manager - new**  
Twitter - ★★★★★ 36 reviews - San Francisco, CA 94103 (South Of Market area)  
A successful engineering manager will: A successful engineering manager will  
communicate fearlessly to build trust...

Recommended Jobs - 49 new

My recent searches

- engineering - Lausanne
- engineering - 32,760 new
- engineering - NY - 1,766 new
- engineering - NYC - 1,625 new
- data - NYC - 3,338 new
- computer - NYC - 3,441 new
- "big Data" - 1,863 new
- data scientist - 2,011 new

Sort by: **relevance** - date

Distance:

Salary Estimate: \$70,000+ (11162)

## Operations Engineer

San Francisco

This is a senior technical operations position. It is based in San Francisco and reports to one of our engineering managers.

Use your extensive knowledge of Linux and your desire to make everything 'just work' to keep a cc high-volume web-app running optimally.

From tracking statistics and managing alerts to deploying new hosts and debugging intricate prod issues, this role requires managing a large and varied infrastructure as it grows. The position also i building out the appropriate tools to make your own job easier, planning for future platform growth anything else that comes up.

### Requirements

- A deep and abiding love of Linux. Or a respectful hatred for it. Either way, you'll need to be ab sed/awk/grep like a pro.
- Strong background in deployment automation/configuration management, especially Chef
- Experience with virtualized environments (AWS experience a plus)
- Ability to analyze and optimize performance in high-traffic internet applications

“Job ads provide information about employers’ expectations of potential employees and other information about evolving skill sets required in the technology workforce.”

[Hauyan H. et al., SIGMIS'09]

# Research Goal

Analysis job skill sets using job ads and find what kind of skills emerged lately.

# Challenges

- Online job ads collections
- Extract skill terms from job description

Have a MSc in **Computer Science**, Electrical Engineering or related disciplines.

Ph.D. degree is a plus Have advanced **Java, Scala or Python** programming skills.

Experience with real-world data is a plus Have advanced **data management skills**, including knowledge of data curation, data governance and data stewardship Have hands-on experience with data manipulation using **SparkSQL, Hive, HBase, Phoenix, Solr** or similar using a leading **Apache Hadoop** distribution.

# Challenges

- Online job ads collections
- Extract skill terms from job description

Have a MSc in Computer Science, Electrical Engineering or related disciplines.  
Ph.D. degree is a plus Have advanced Java, Scala or Python programming skills.  
Experience with real-world data is a plus Have advanced data management skills,  
including knowledge of data curation, data governance and data stewardship Have  
hands-on experience with data manipulation using SparkSQL, Hive, HBase,  
Phoenix, Solr or similar using a leading Apache Hadoop distribution.

## Exploratory analysis on job titles

represent a set of skills

# JobApis Open Source

We offer PHP API clients for many job boards:


- [Adzuna Jobs](#)
- [Authentic Jobs](#)
- [CareerBuilder](#)
- [Careercast](#)
- [Careerjet](#)
- [CareersInGovernment](#)
- [Craigslist](#)
- [Dice](#)
- [Elance](#)
- [Github Jobs](#)
- [GoRemote](#)
- [Government Jobs](#)
- [IEEE](#)
- [Indeed](#)
- [Jobinventory](#)
- [Jobs2Careers](#)
- [JuJu](#)
- [Monster.com](#)
- [Muse](#)

### Software Engineer (m/f)

#### General Electric

[Tweet](#) [G+](#) [Like 0](#)

**Location:** Eysins  
**Date:** 05/03/2017  
**Job Code:** ge2-2882193  
**Categories:** Engineering  
Manufacturing / Production  
Healthcare

 **GE imagination at work**  
[General Electric](#)  
[Apply Now](#)  
Apply on the Company Site

[Save Ad](#) [Email Friend](#) [Print](#)

#### Job Details

**Job Title** Software Engineer (m/f)  
**Jobid** ge2-2882193  
**Location:** Eysins, CHE  
**Description** \*\*About Us:\*\*  
There is no other life sciences company quite like GE Healthcare Life Sciences. A \$5 billion business with more than 11,000 individuals in over 100 countries, we are at work enabling a new era in medicine. For more than 60 years, GE Healthcare Life Sciences has globally supported the development and use of new therapies, helping accelerate the journey from biological understanding to clinical therapy. GE Healthcare Life Sciences works with cutting edge technology enabling and advancing the discovery, diagnosis and medical treatments of the present and the future. We are accelerating biology to cell therapy in ways no other company can.

<http://expat.careers.adicio.com/jobs/software-engineer-m-f-eyins-95831815-d?rsite=careercast&rgroup=1&clientid=expat&widget=1&type=job&>

# JobApis

# Open Source

We offer PHP API clients for many job boards:

- [Adzuna Jobs](#)
- [Authentic Jobs](#)
- [CareerBuilder](#)
- [Careercast](#)
- [Careerjet](#)
- [CareersInGovernment](#)
- [Craigslist](#)
- [Dice](#)
- [Elance](#)
- [Github Jobs](#)
- [GoRemote](#)
- [Government Jobs](#)
- [IEEE](#)
- [Indeed](#)
- [Jobinventory](#)
- [Jobs2Careers](#)
- [JuJu](#)
- [Monster.com](#)
- [Muse](#)

```
Query : engineering
Source : Careercast
Title : Software Engineer (m/f)
Url : http://expat.careers.adicio.com/jobs/software-engineer-m-f-ey-sins-95831815-d?rsite=careercast&rgroup=1&clientid=expat&widget=1&type=job&
Industry : null
OccupationalCategory : Engineering, Manufacturing / Production, Healthcare
Country : CHE
Location : Eysins, VD
DatePosted : 2017-05-03T07:00:00+0000
▶ Company {6}
▶ Adress {5}
▶ Salary {3}
▶ Description {11}
▶ Other {2}
```

```
Description : Software Engineer (m/f) Jobid ge2-2882193 Location: Eysins, CHE
Description **About Us:** There is no other life sciences company quite like GE Healthcare Life Sciences. A $5 billion business with more than 11,000 individuals in over 100 countries, we are at work enabling a new era in medicine. For more than 60 years, GE Healthcare Life Sciences has globally supported the development and use of new therapies, helping accelerate the journey from biological understanding to clinical therapy. GE Healthcare Life Sciences works with cutting edge technology enabling and advancing the discovery, diagnosis and medical treatments of the present and the future. We are accelerating biology to cell therapy in ways no other
```

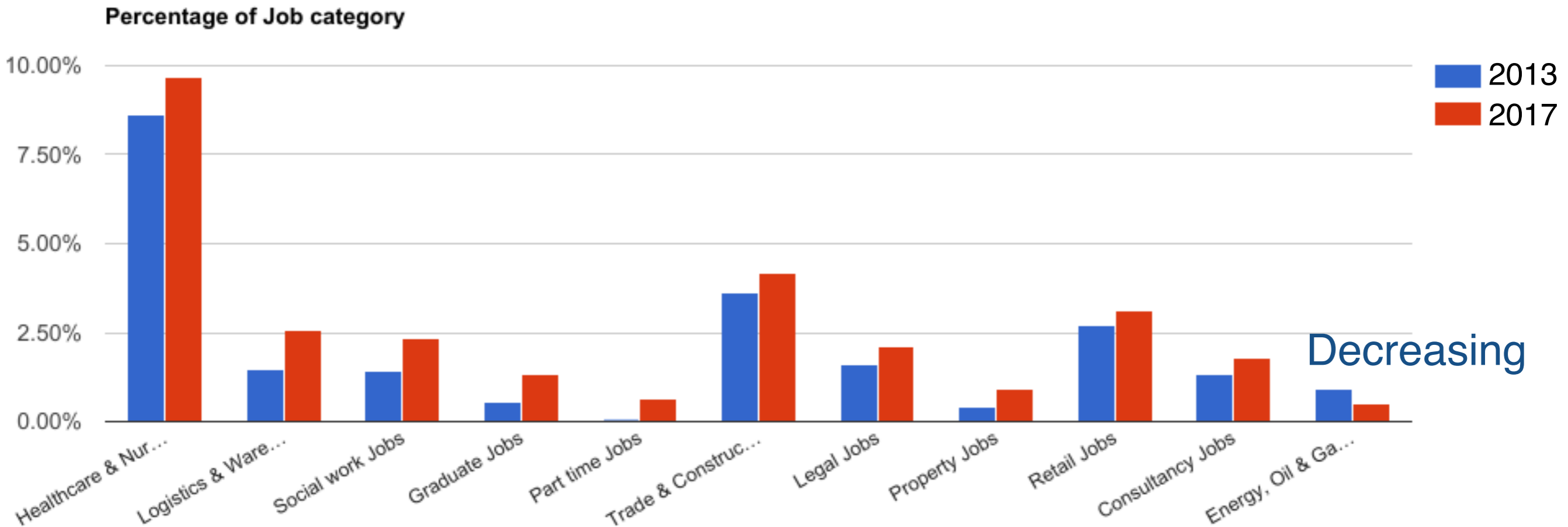
# Exploratory Analysis



Year	# Jobs	# Categories	# Job Title	Location
2013	244,768	29	135,436	UK
2017	357,750	29	168,143	UK

- How the job ads change between **2013 and 2017**
- Two datasets are in similar context

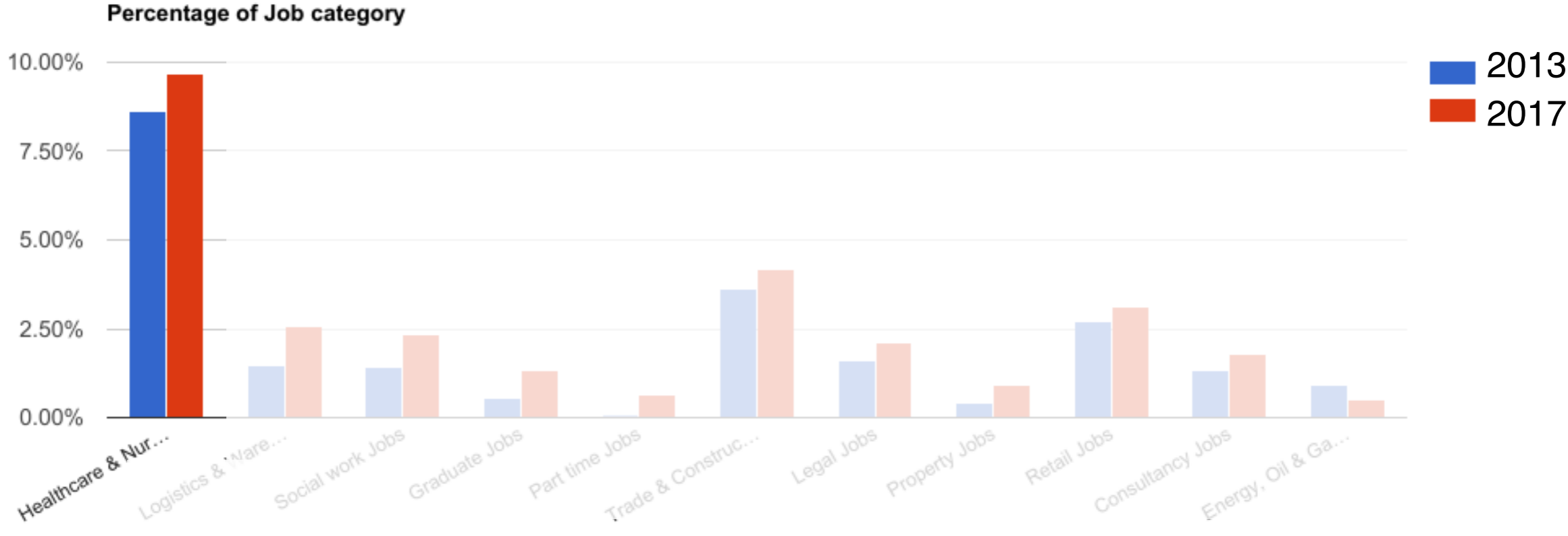
# Rank of top 10 increasing job categories



Google Trends



# Increase: Healthcare & Nursing Jobs



6 Jan. 2013

25 May 2014

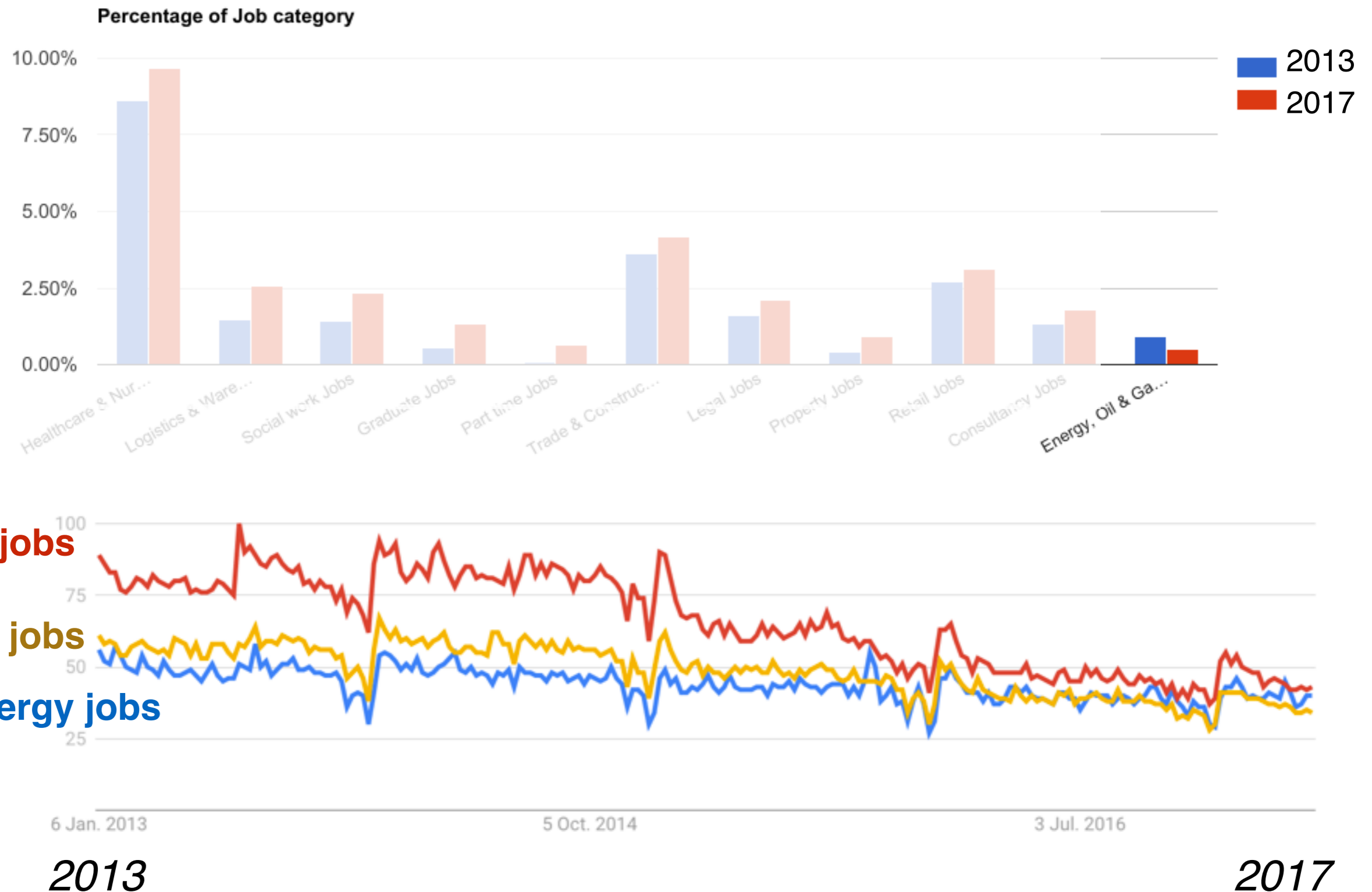
11 Oct. 2015

26 Feb. 2017

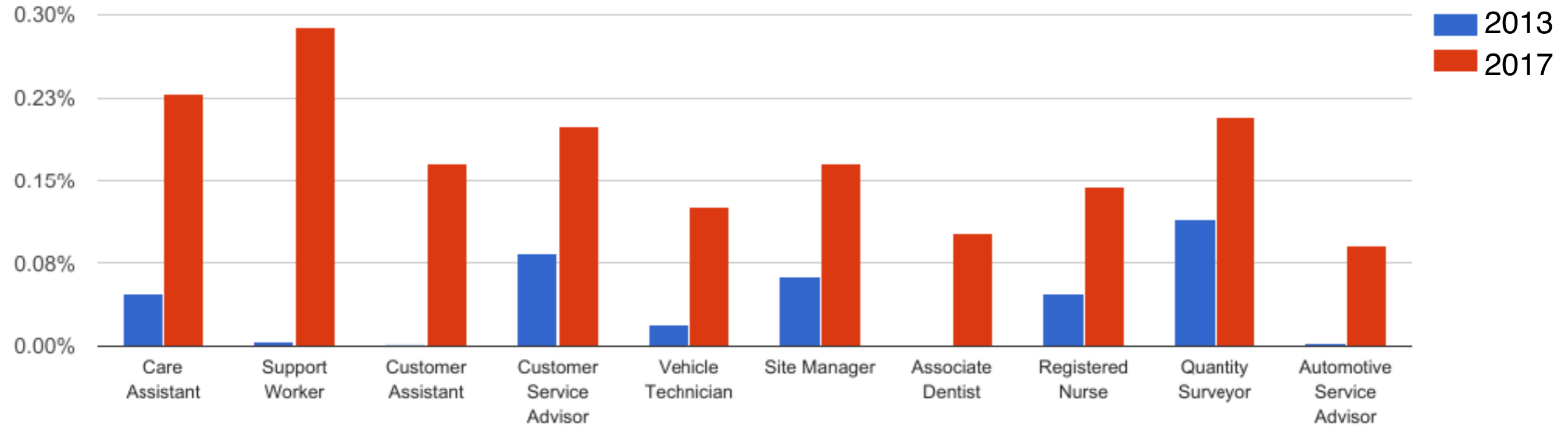
**2013**

**2017**

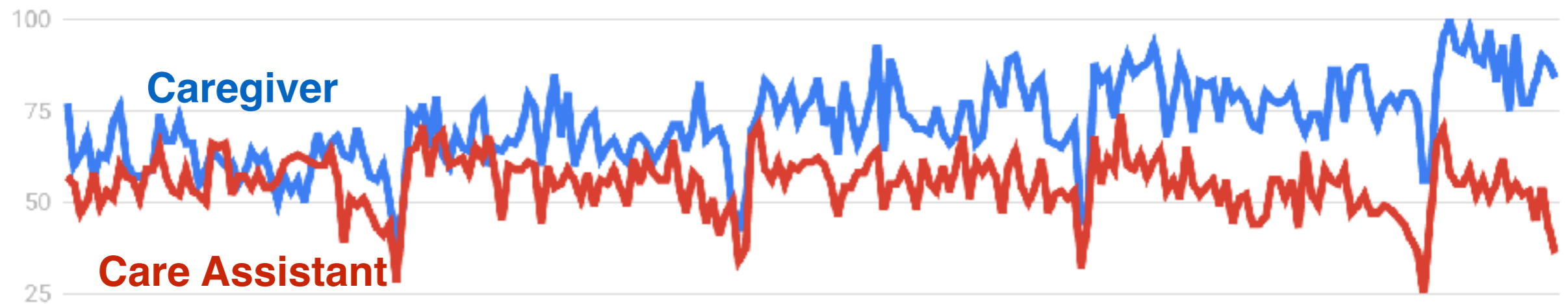
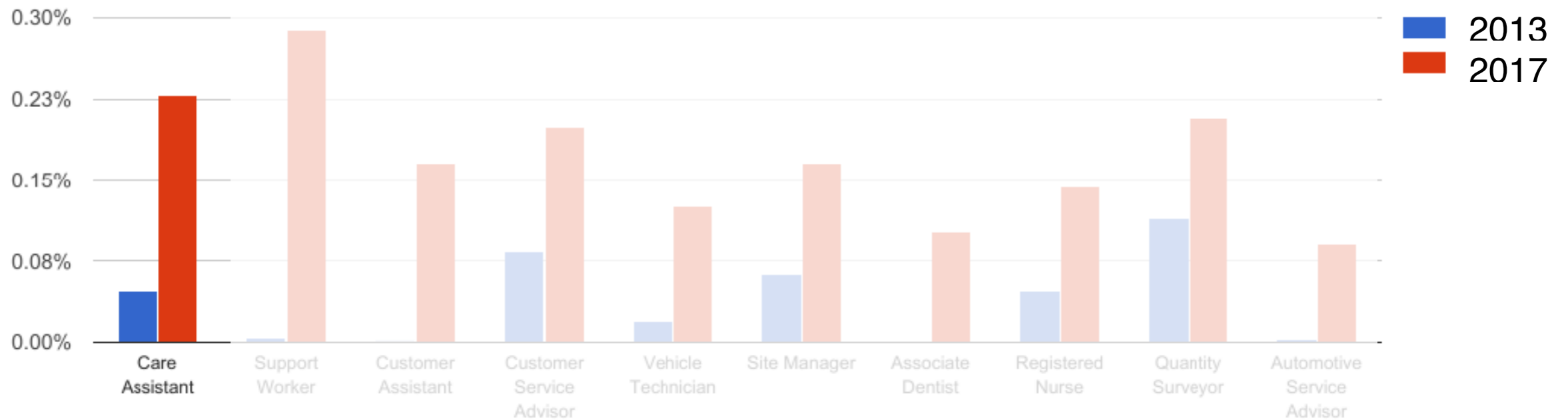
# Decrease: Energy, Oil & Gas Jobs



# Rank of top 10 increasing job titles



# Increase: Care Assistant



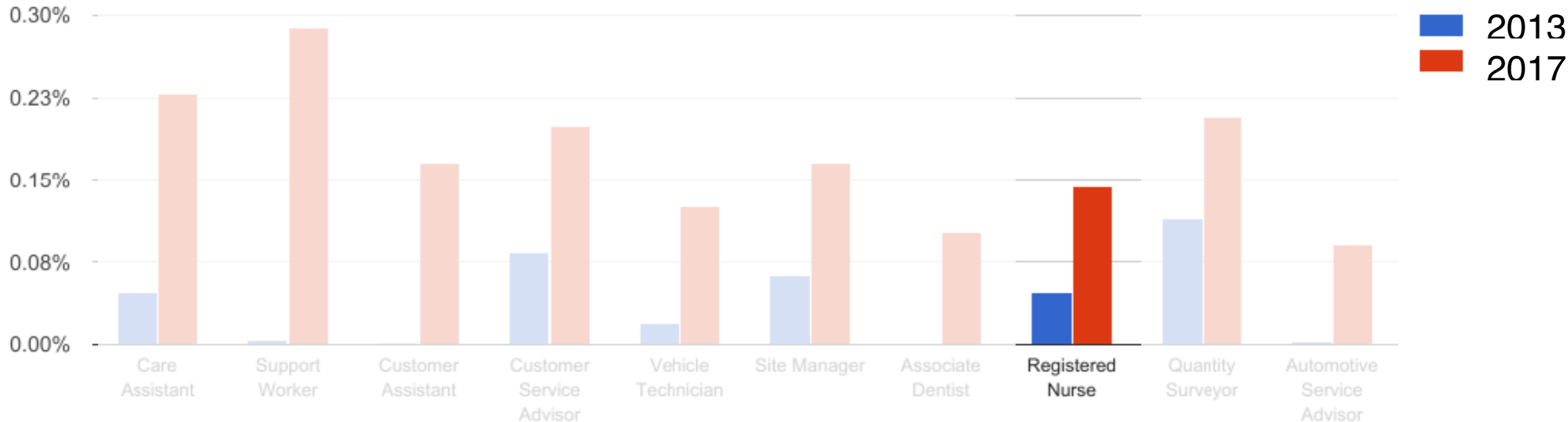
6 Jan. 2013  
**2013**

16 Nov. 2014

25 Sep. 2016

**2017**

# Increase: Registered Nurse



6 Jan. 2013

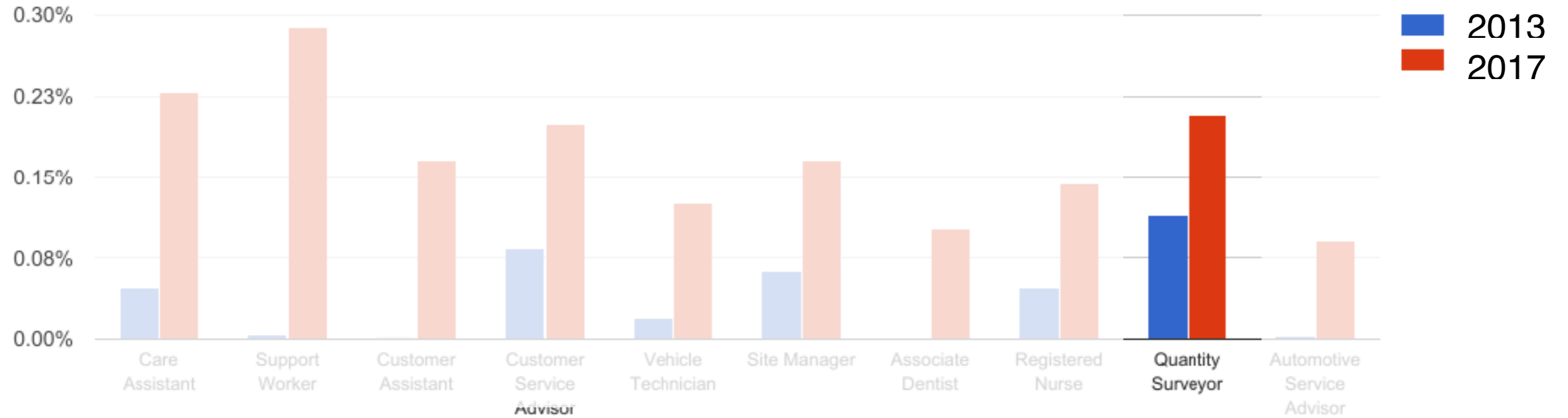
3 Aug. 2014

28 Feb. 2016

*2013*

*2017*

# Increase: Quantity Surveyor



6 Jan. 2013  
**2013**

5 Oct. 2014

3 Jul. 2016

**2017**



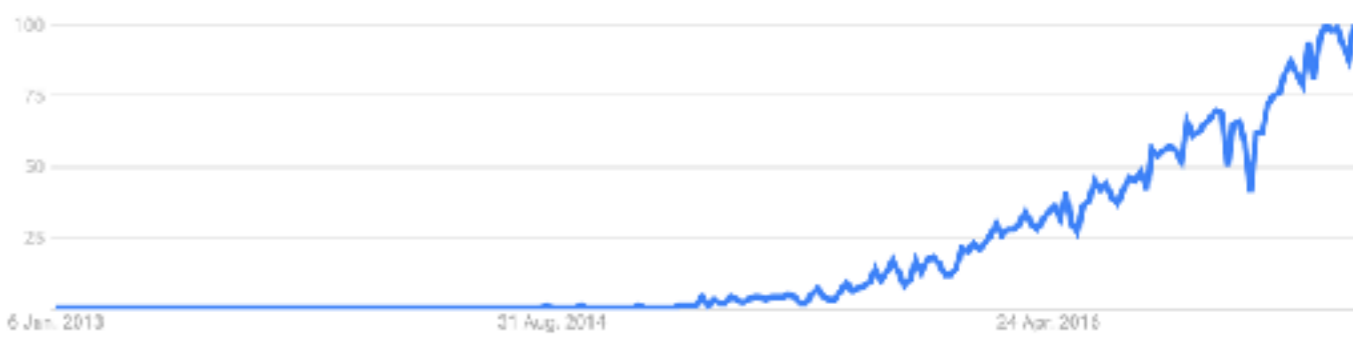
Functional Programming



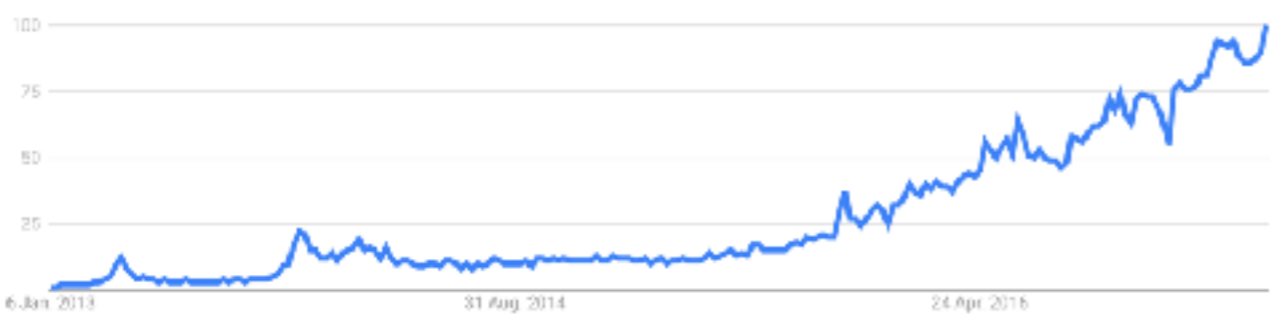
Data Science



Deep Learning



Jupyter Notebook



Blockchain



Social Media Manager

# Future Works

- Skill terms extraction
- Detecting emerging skills
- Skill gaps between education curricula and industrial requirements

Have a MSc in **Computer Science**, Electrical Engineering or a Ph.D. degree is a plus Have advanced **Java, Scala or Python** programming skills Experience with real-world data is a plus Have advanced **data science** skills including knowledge of data curation, data governance and data security hands-on experience with data manipulation using **SparkSQL, Phoenix, Solr** or similar using a leading **Apache Hadoop** distribution



Thank you