Big Data, AI & Semantic Matching of Jobs and People

Jakub Zavrel

CEO **Textkernel**

textkernel

Semantic Recruitment Technology

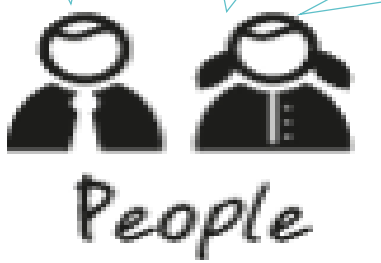
Language gap

I like programming, but I'm interested do take on more project management responsibility

Is there a job in our organisation that better fits my degree?

I'd like to work on our mobile strategy. I've helped a friend develop a mobile app.

I'd like to do more with my organisational talent.

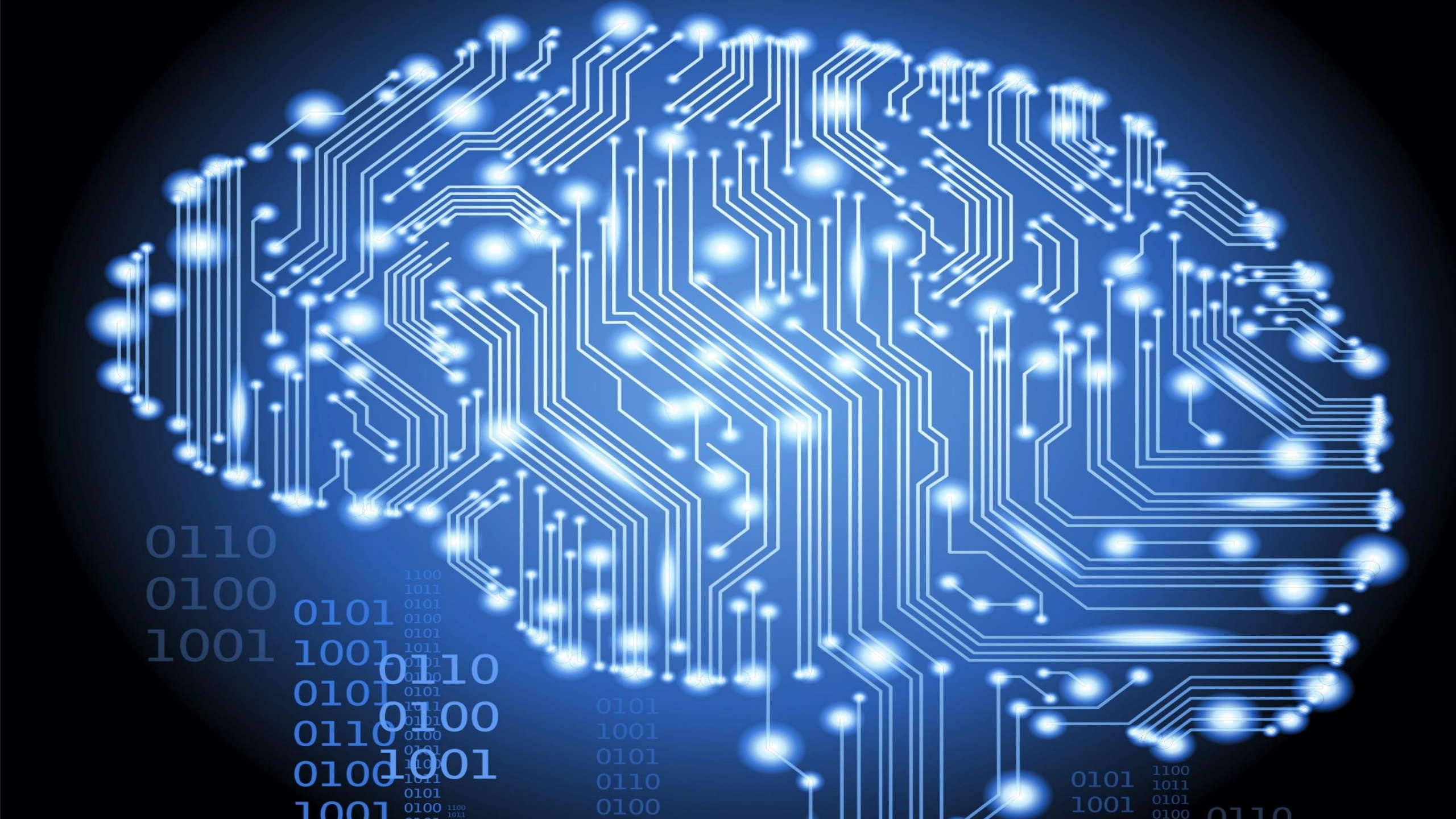


We are looking to hire:
An experienced tech team team lead

The ideal candidate has:
- min. 5yr of experience
- Certfied scrummaster
- Exp. w/iOS, Android

Completed academic studies
Computer Science or related

30% travel for customer presentations


0110
0100
1001
0101
1001
0101
0110
0100
1001
0101
1001
0101
0110
0100
0101
1001
1100
1011
0101
0100

Value for Recruitment and HR

Software to fill your jobs faster and better

Time savings and process optimisation

Insight and understanding

June 2nd 2016
Textkernel 15 years conference

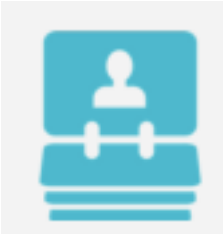
Impact of technology on recruitment…

Balacz Paroczay, Randstad:
https://youtu.be/7E4clMBgps0?list=PLBALD27enveIgCO8cBCieGn5rIiplbY3q&t=681

Glen Cathey, Kforce:
https://youtu.be/mHe9Sma1Y4w?list=PLBALD27enveIgCO8cBCieGn5rIiplbY3q&t=1340

https://youtu.be/mHe9Sma1Y4w?list=PLBALD27enveIgCO8cBCieGn5rIiplbY3q&t=1577

Textkernel products & value:



Extract!

Smooth Candidate Experience

One click mobile apply

Search & Match!

Smooth Recruiter Experience

Sourcing in your ATS database.

Embedded Semantic Search & Match

Jobfeed

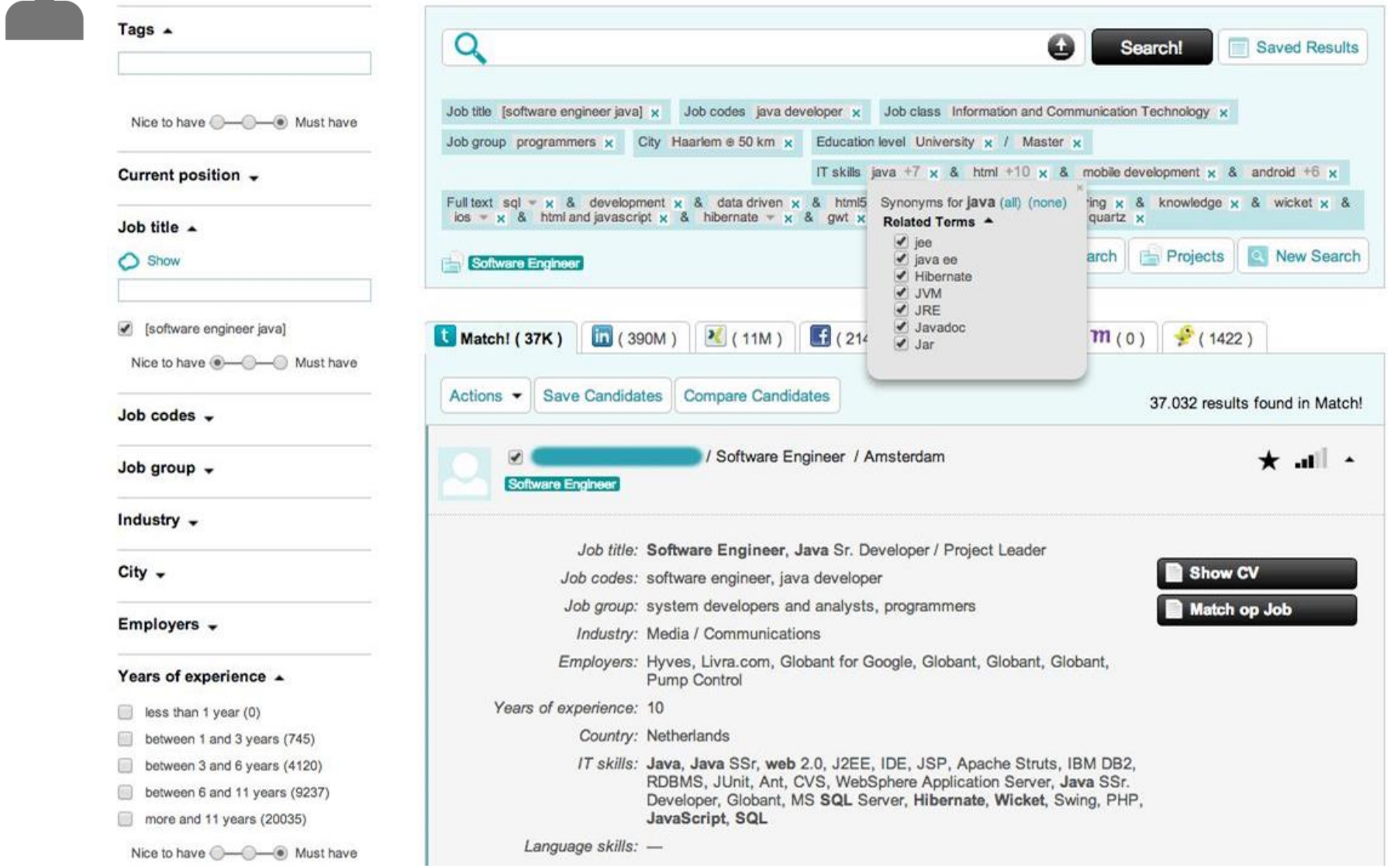
Lead Generation for Staffing

“Match your supply to the demand in the market”

Search & Match!

- **Semantic search:**
  Find what you mean, not what you type

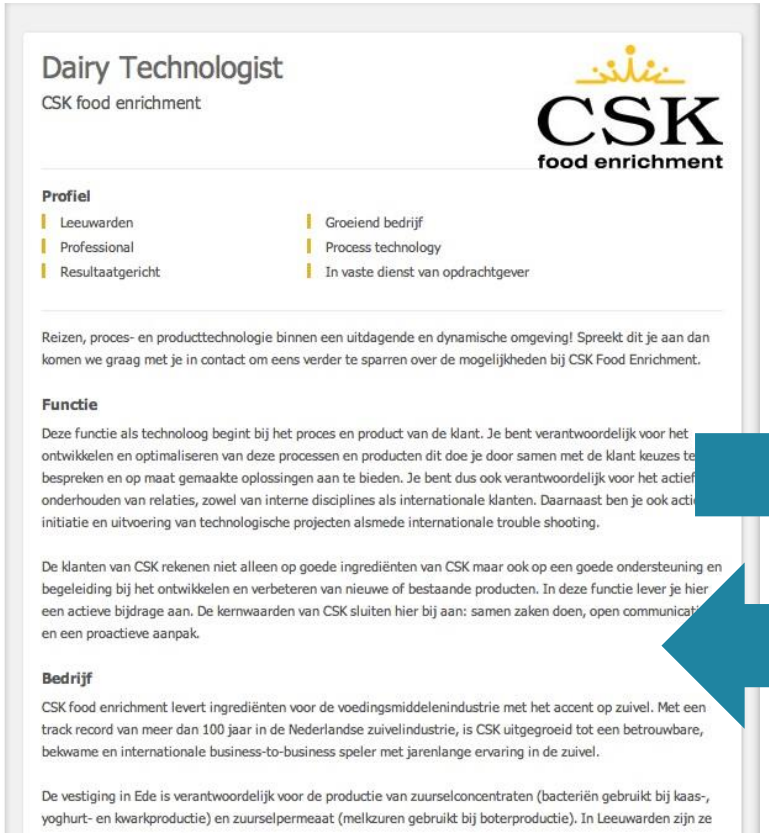
- **Match**:
  Upload your job and automatically get recommendations of relevant candidates and vice versa



Match!

What if:

…you can recommend the right candidate ***directly*** from a vacancy description without searching? And vice-versa?



Dairy Technologist

CSK food enrichment

CSK food enrichment

Profiel

- Leeuwarden
- Professional
- Resultaatgericht
- Groeiend bedrijf
- Process technology
- In vaste dienst van opdrachtgever

Reizen, proces- en producttechnologie binnen een uitdagende en dynamische omgeving! Spreekt dit je aan dan komen we graag met je in contact om eens verder te sparren over de mogelijkheden bij CSK Food Enrichment.

Functie

Deze functie als technoloog begint bij het proces en product van de klant. Je bent verantwoordelijk voor het ontwikkelen en optimaliseren van deze processen en producten dit doe je door samen met de klant keuzes te bespreken en op maat gemaakte oplossingen aan te bieden. Je bent dus ook verantwoordelijk voor het actief onderhouden van relaties, zowel van interne disciplines als internationale klanten. Daarnaast ben je ook acti initiatie en uitvoering van technologische projecten alsmede internationale trouble shooting.

De klanten van CSK rekenen niet alleen op goede ingrediënten van CSK maar ook op een goede ondersteuning en begeleiding bij het ontwikkelen en verbeteren van nieuwe of bestaande producten. In deze functie lever je hier een actieve bijdrage aan. De kernwaarden van CSK sluiten hier bij aan: samen zaken doen, open communicat en een proactieve aanpak.

Bedrijf

CSK food enrichment levert ingrediënten voor de voedingsmiddelenindustrie met het accent op zuivel. Met een track record van meer dan 100 jaar in de Nederlandse zuivelindustrie, is CSK uitgegroeid tot een betrouwbare, bekwame en internationale business-to-business speler met jarenlange ervaring in de zuivel.

De vestiging in Ede is verantwoordelijk voor de productie van zuurselconcentraten (bacteriën gebruikt bij kaas-, yoghurt- en kwarkproductie) en zuurselpermeaat (melkzuren gebruikt bij boterproductie). In Leeuwarden zijn ze

Currently available in:

- Netherlands
- Belgium
- Germany
- France
- United Kingdom
- Austria
- Italy*
- New: USA!!!

Lead Generation

Market Analysis

Match your supply to the demand in the market”

Recruitment questions, that can be answered with Jobfeed

- Which direct employers are looking for candidates that we have in our database?
- Which new jobs are my customers publishing and through which platforms?
- **What is the average salary and skills asked for this type of profile?**
- Who are the largest employers in my region? And which professions are they hiring in?

Labor market analysis questions

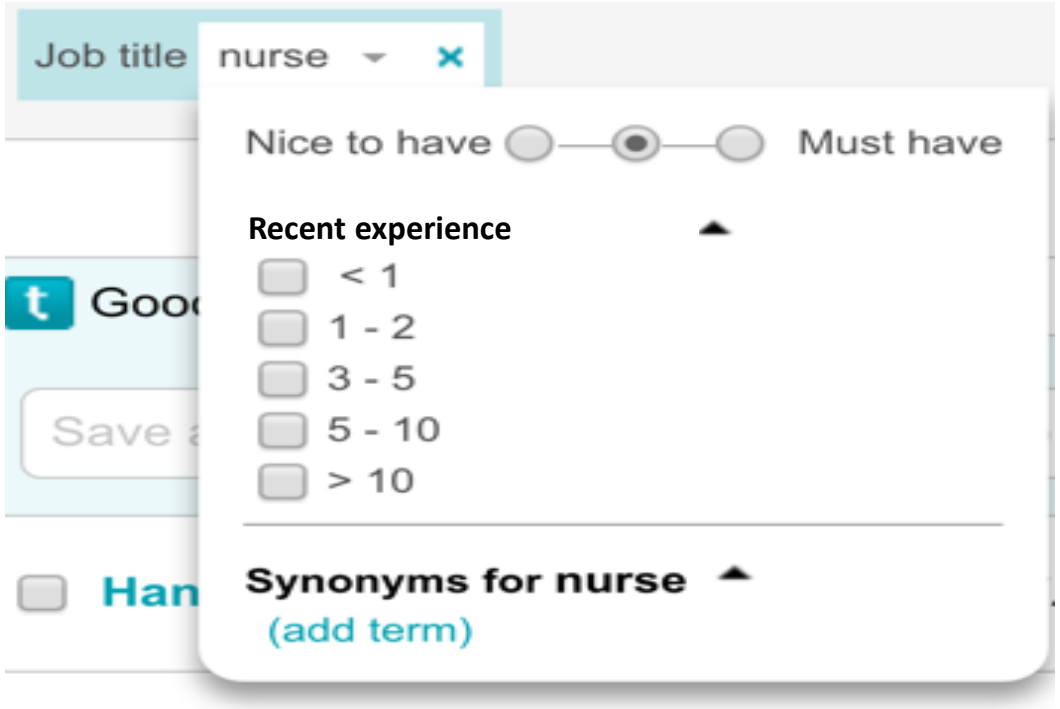
- What are the professions in highest demand? Which professions are growing most rapidly?
- Which requirements are there for e.g. Electrical Engineers and how do these change over time?
- Do new professions evolve? (cf. “data scientist” or “community manager”). What skills do these need?
- Is Microsoft or Linux a better technology to teach?

WHAT'S NEXT?

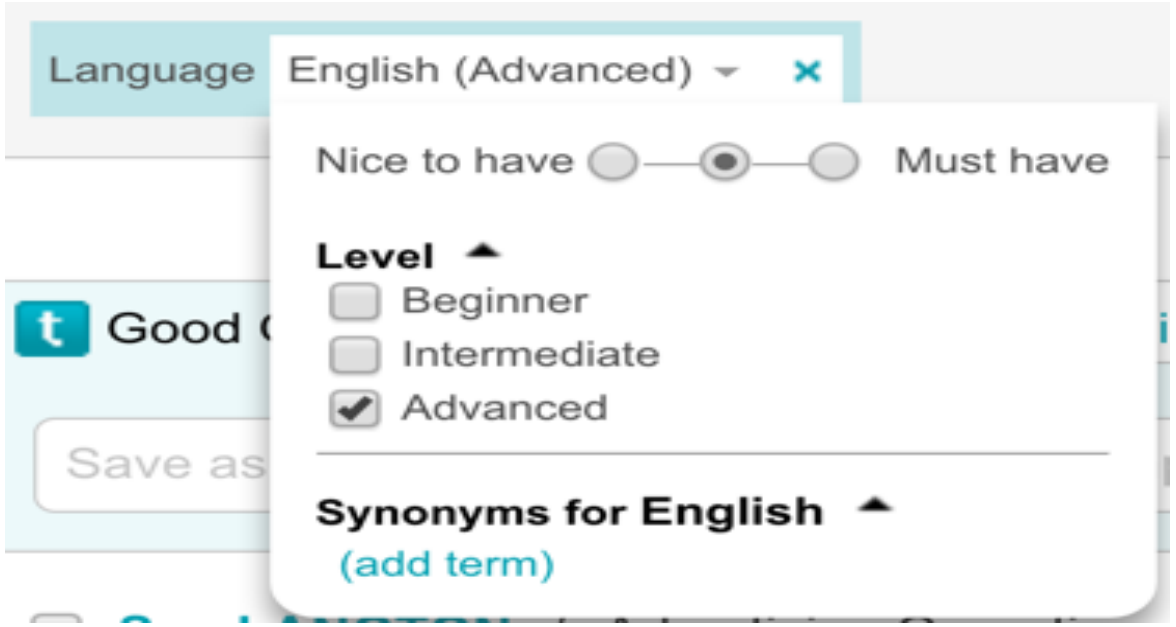
2016 Roadmap

Textkernel Search on subfields

Recent experience



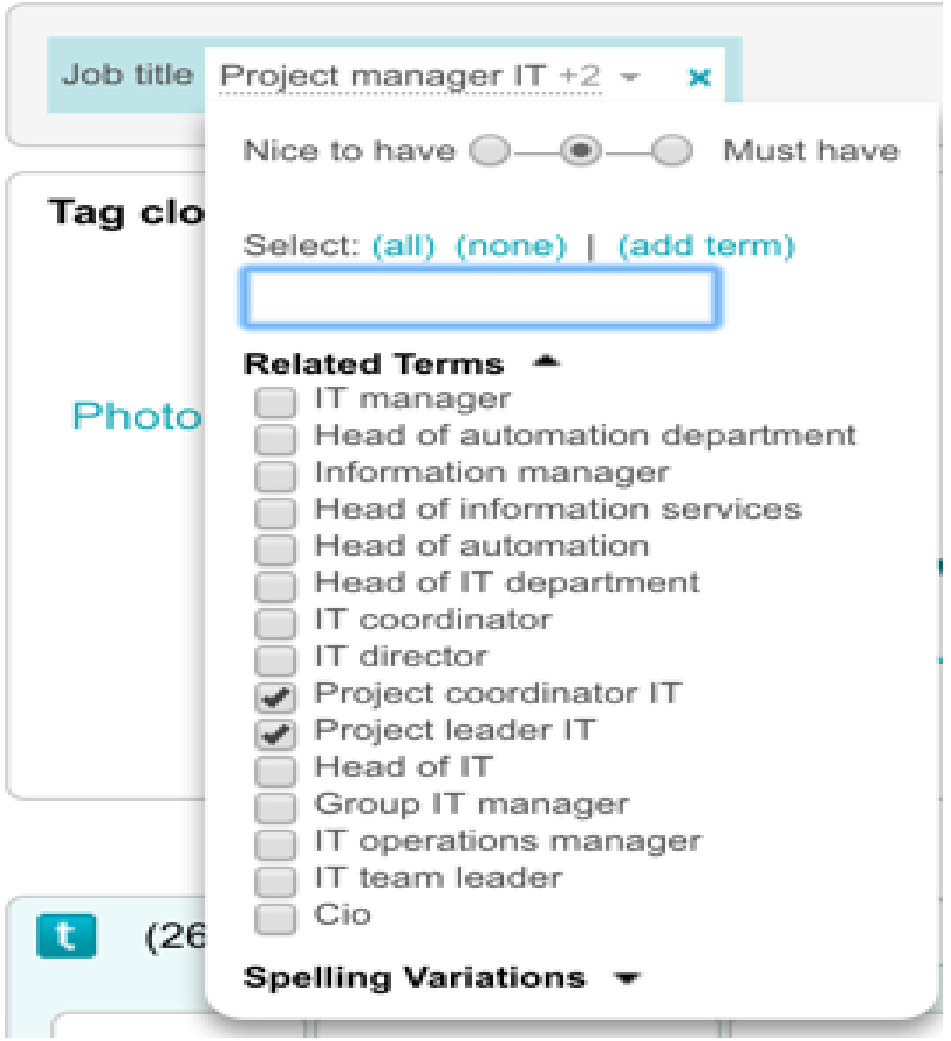
Skill level



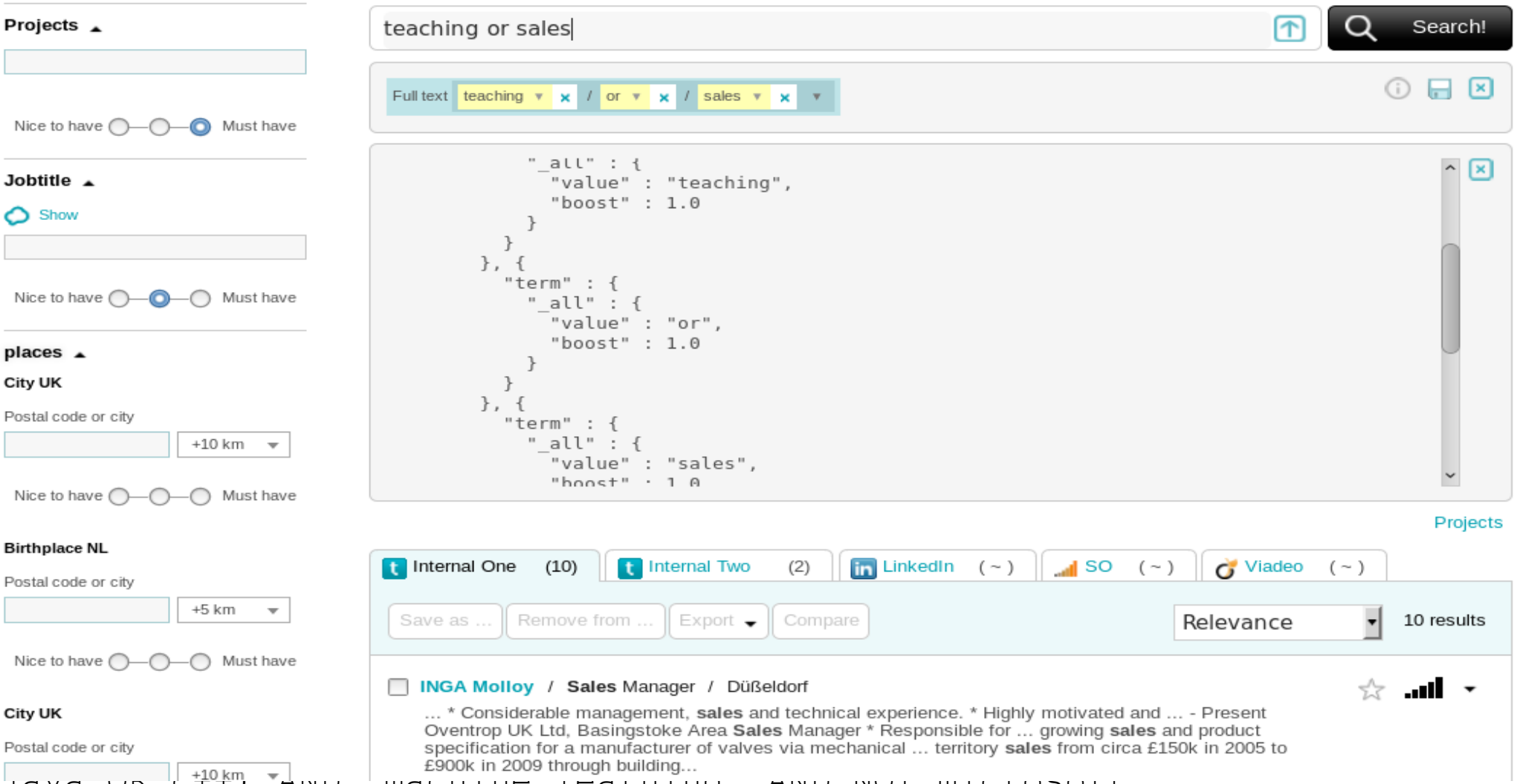
Crowdsourcing knowledge

Suggesting related terms added by co-workers

Improve the standard knowledge base

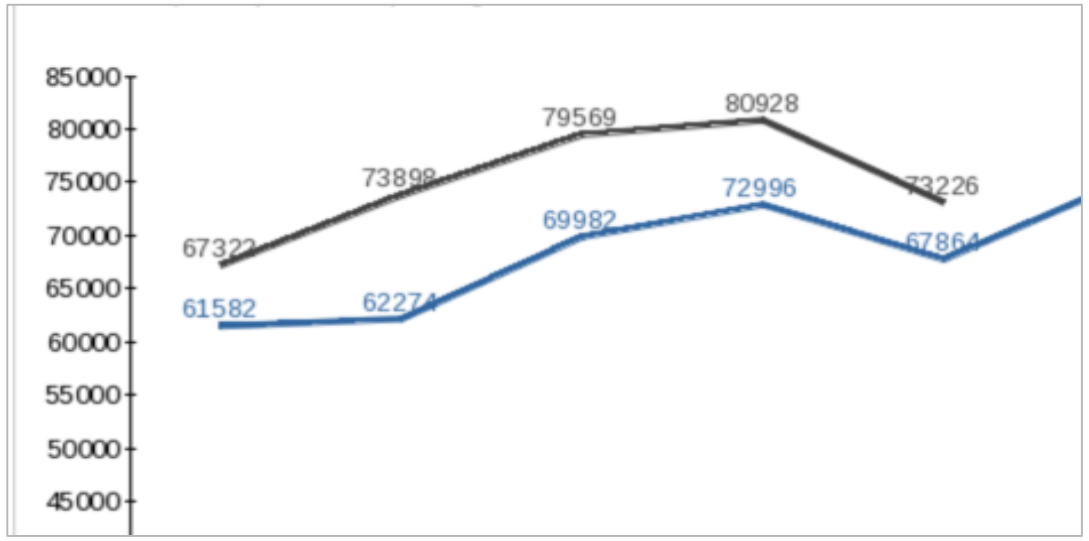


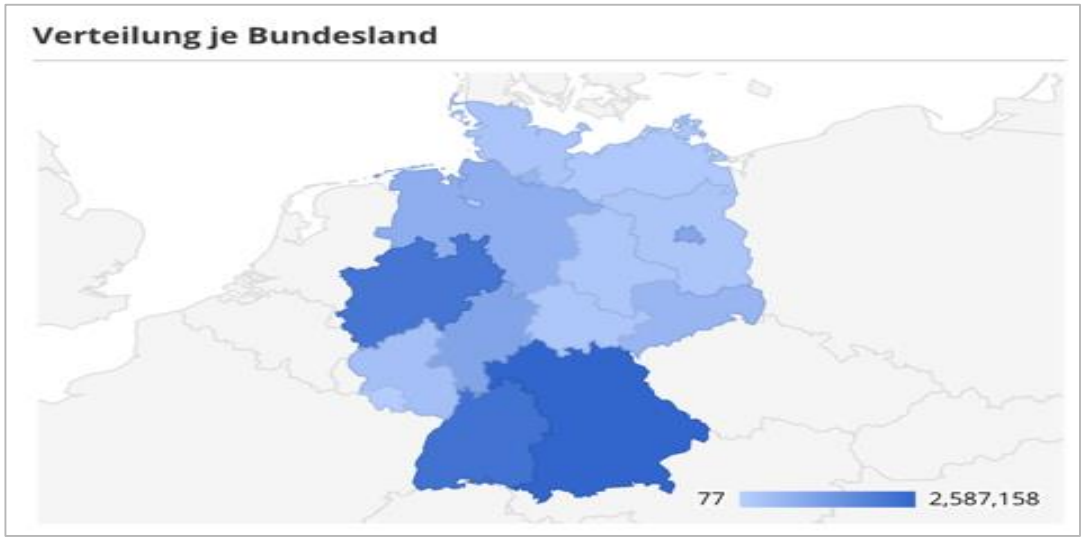
New: Boolean syntax in Textkernel Search!?

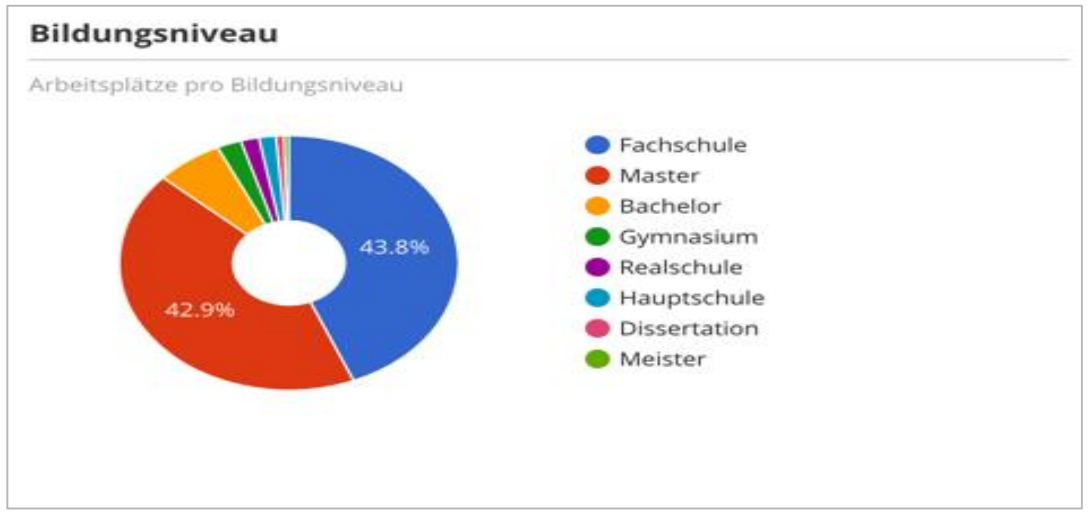


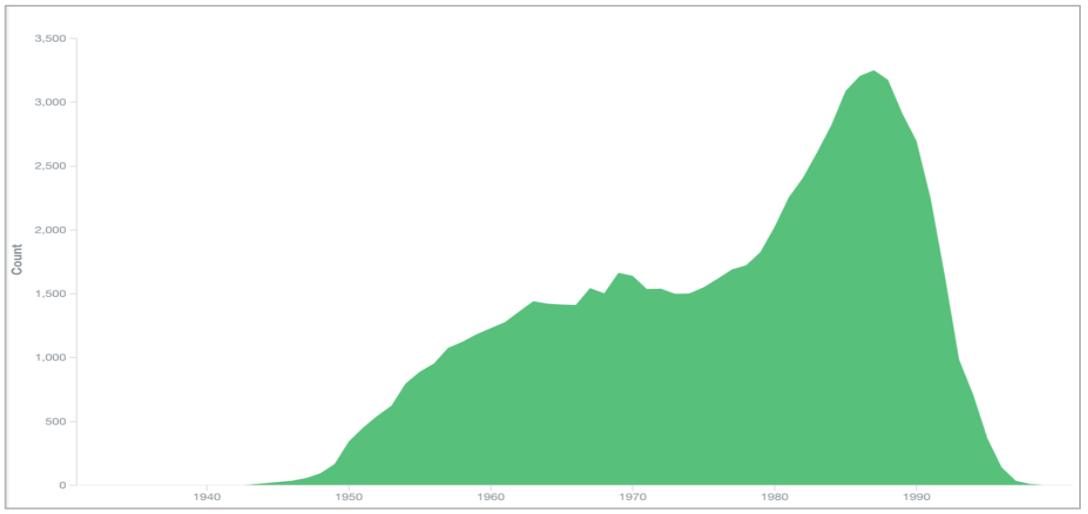
("java OR C++") AND "machine learning" AND NOT microsoft

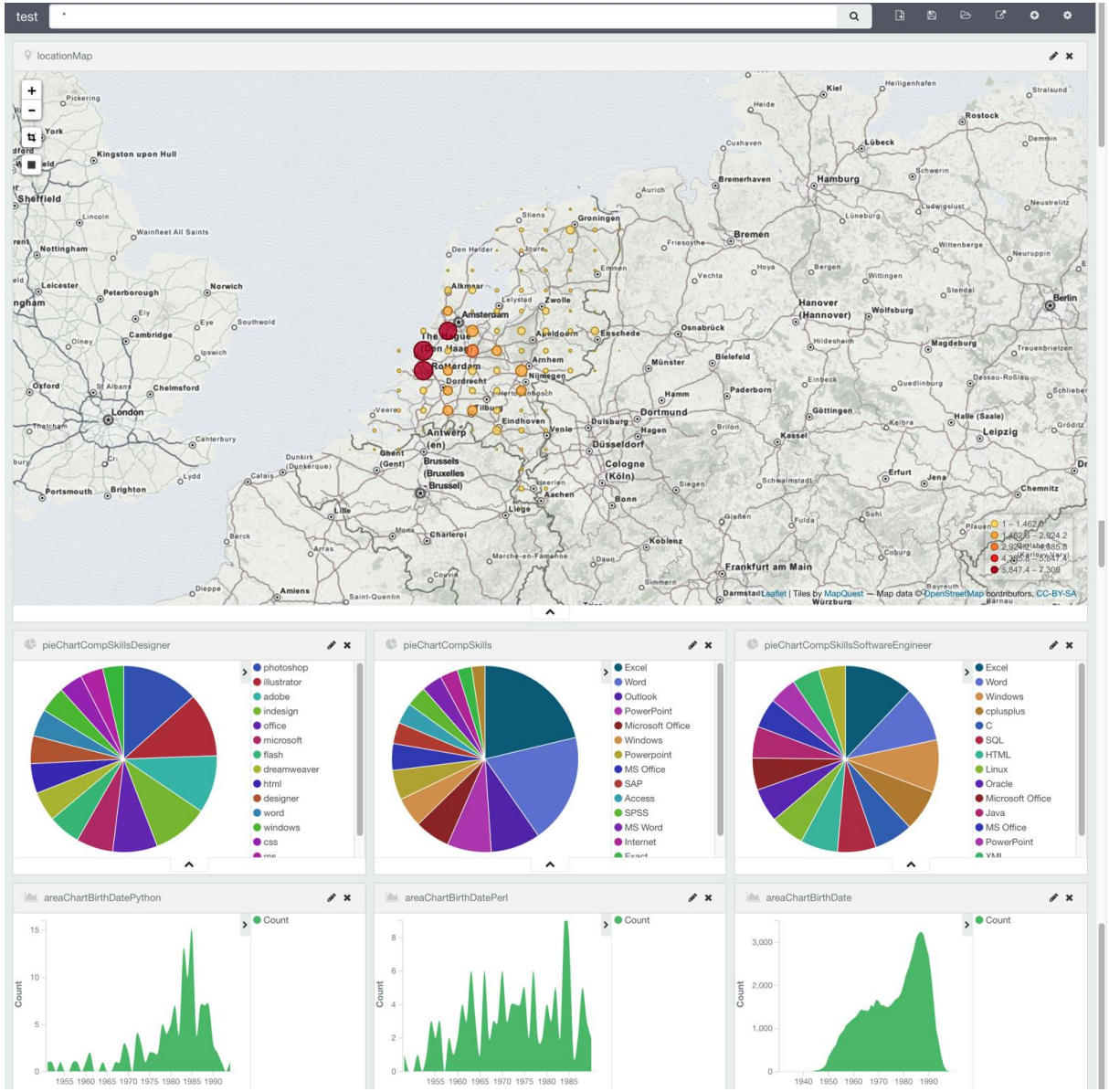
Roadmap: Visualizing the data



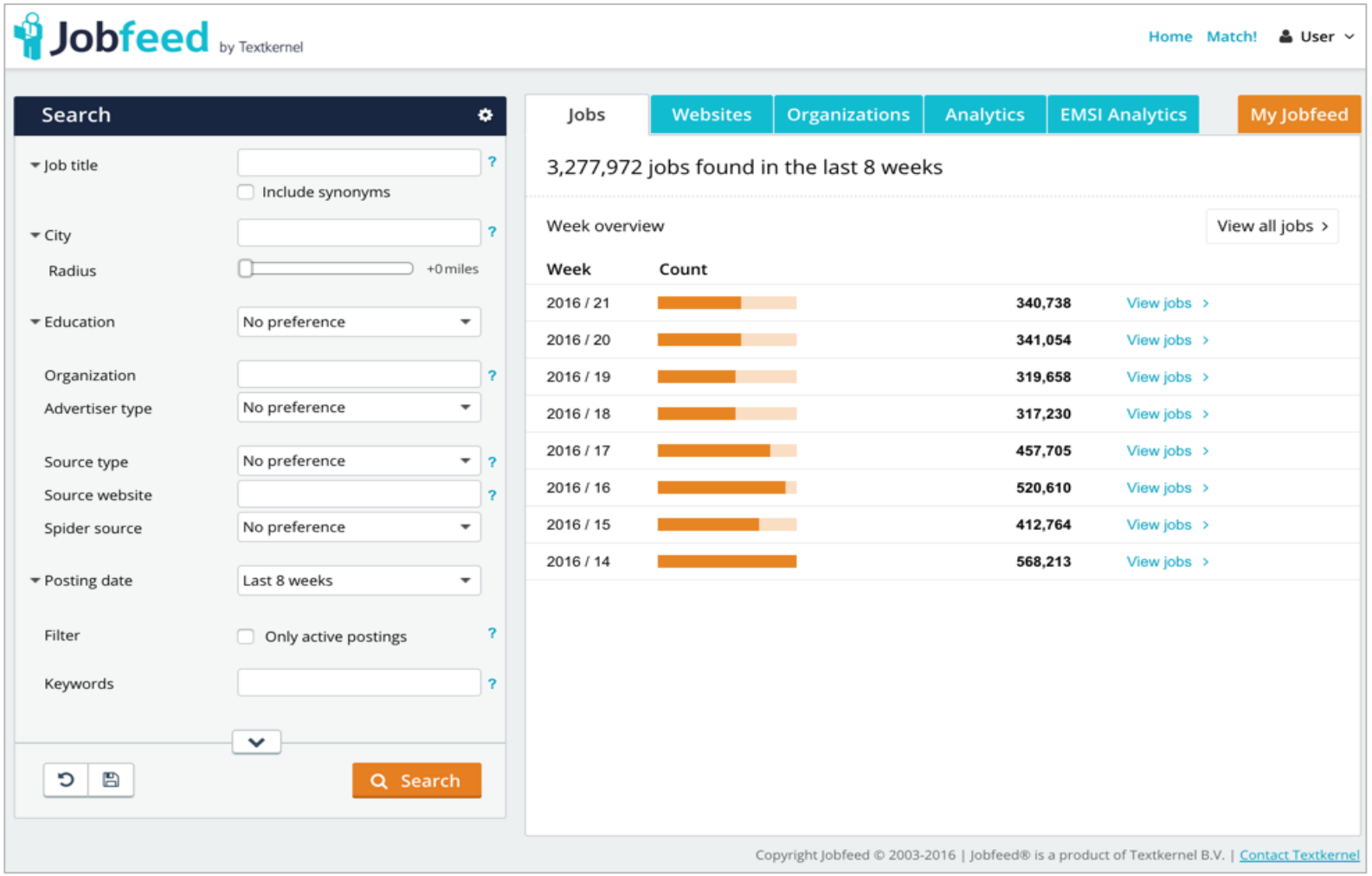







test
locationMap
pieChartCompSkillsDesigner
photoshop
illustrator
adobe
indesign
office
microsoft
flash
dreamweaver
html
designer
word
windows
css
pieChartCompSkills
Excel
Word
Outlook
PowerPoint
Microsoft Office
Windows
Powerpoint
MS Office
SAP
Access
SPSS
MS Word
Internet
pieChartCompSkillsSoftwareEngineer
Excel
Word
Windows
cplusplus
C
SQL
HTML
Linux
Oracle
Microsoft Office
Java
MS Office
PowerPoint
areaChartBirthDatePython
Count
areaChartBirthDatePerl
Count
areaChartBirthDate
Count

Jobfeed - United States



New fields in Jobfeed

- Filter: salary, contact details, posting language
- Reporting: ISCO, vacancy language, skills

IT skills / Year+Month	2016 - 04	2016 - 03
MS-Excel	20370	10986
C/C++	11094	6534
Data Warehouse	9248	5245
MS-Office	8750	5056
Network	8305	5134
mySQL	8599	4833
HTML	8043	4003
MS-Word	7558	4400
Enterprise Resource Planning	7391	4122
Agile	6931	4148

ar+Month	2016 - 04	2016 - 03
English	11665	6593
German	1647	871
French	1561	854
Spanish	833	445
Chinese	497	320
Italian	464	261
Welsh	270	269
Dutch	322	194
Fijian	235	140
Japanese	194	103

Supply and Demand

Web Developer
March 2014 to February 2016

Supply
0
CareerBuilder.com - INTL Test Account Candidates
3,113
Active Candidates
View Details

Demand
26,522
Job Postings

Hiring Indicator
31

40,153 Total Available Workforce

Job Posters and Compensation Data

Top Job Posters

Cyber Coders	1,201
Oracle Corporation	891
Robert Half International Inc.	778
Randstad Holding Nv	513
Liberty Mutual Holding Company Inc.	466
International Business Machines Corporation	421

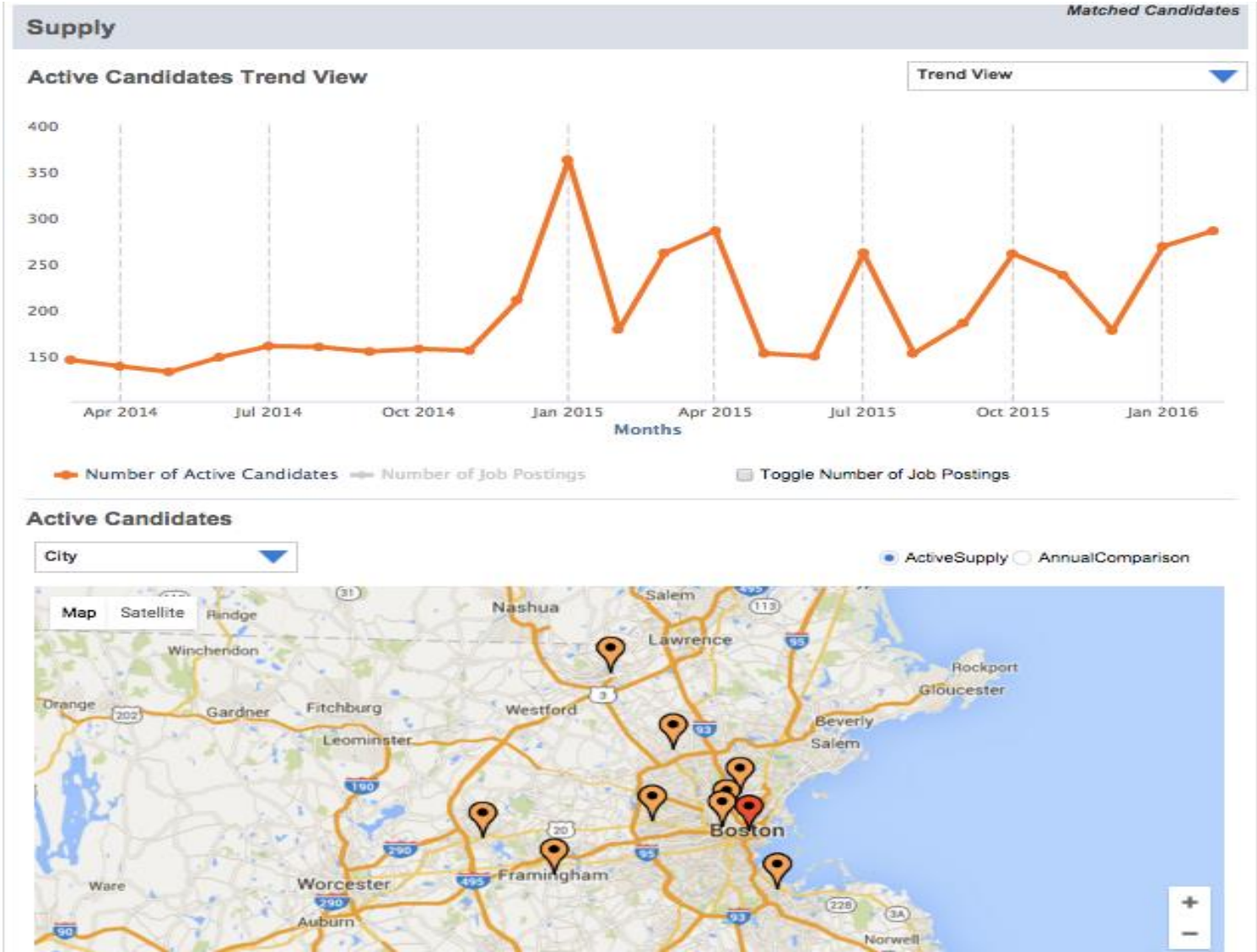
Compensation

$117,635	$100,000	$118,000	$138,000
AVG	25th	50th	75th

Your Markets

$108,723	$85,000	$103,000	$125,000

Supply and Demand



Supply and Demand

https://www.arbeidsmarktdashboard.nl/index.php/page/brainport/command/dashboards/bb/1/#/dashboard06

Nieuws

Overheid

- ECONOMIE
- DEMOGRAFIE
- BEROEPSBEVOLKING
- WERKGELEGENHEID
- VESTIGINGEN
- WERKLOOSHEID EN UITKERINGEN
- VACATURES
- ONDERWIJS
- AANSLUITING ONDERWIJS EN ARBEIDSMARKT
- MOBILITEIT

FILTER RESULTATEN

Welkom terug, Jakub Zavrel

Werkloosheid en uitkeringen

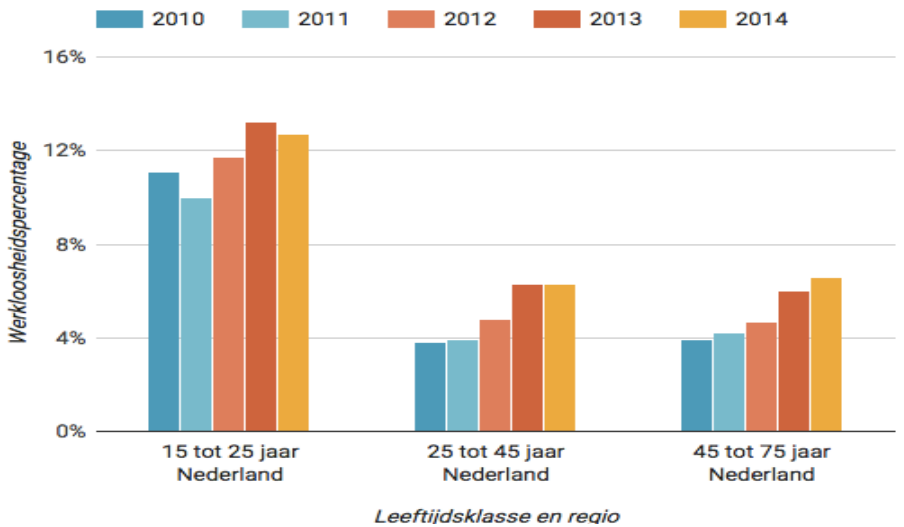
Op de arbeidsmarkt draait het om de vraag naar arbeidskrachten (werkgelegenheid) en het aanbod van arbeidskrachten (beroepsbevolking). In de praktijk sluiten vraag en aanbod op de arbeidsmarkt vaak niet op elkaar aan en zo ontstaan er vacatures of werkloosheid. Om de (verwachte) ontwikkeling van de werkloosheid te beschrijven, wordt het werkloosheidspercentage volgens de CBS-definitie gebruikt: de werkloze beroepsbevolking als percentage van de totale beroepsbevolking.

Werkloosheid

Wat is de gerealiseerde en verwachte ontwikkeling van het werkloosheidspercentage?

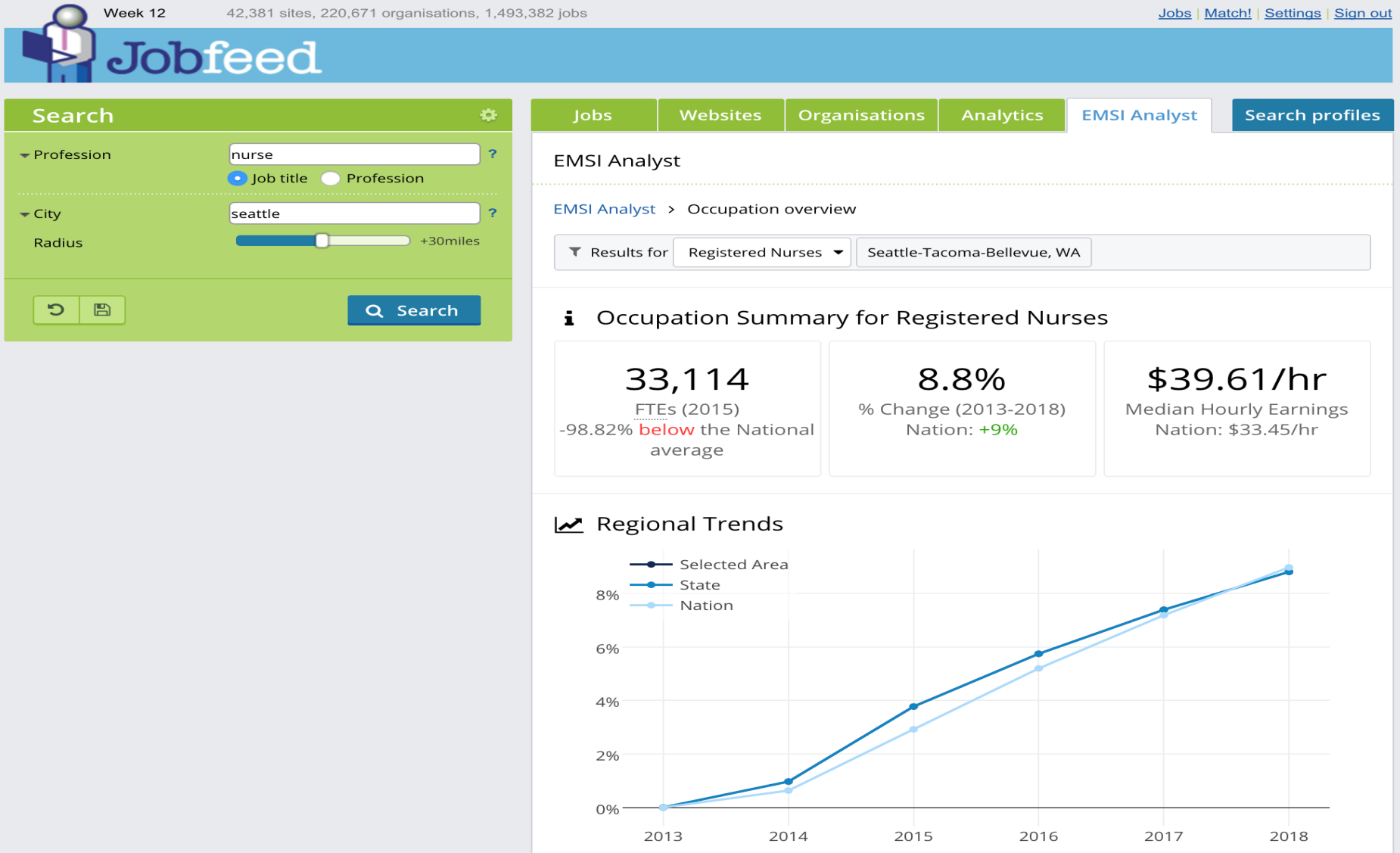
Werkloosheid onder jongeren en ouderen

In welke leeftijdsgroep is de werkloosheid het hoogst? Hoe heeft zich de werkloosheid in deze groepen ontwikkeld?



FEEDBACK

Supply and Demand



Long term roadmap (Textkernel R&D)

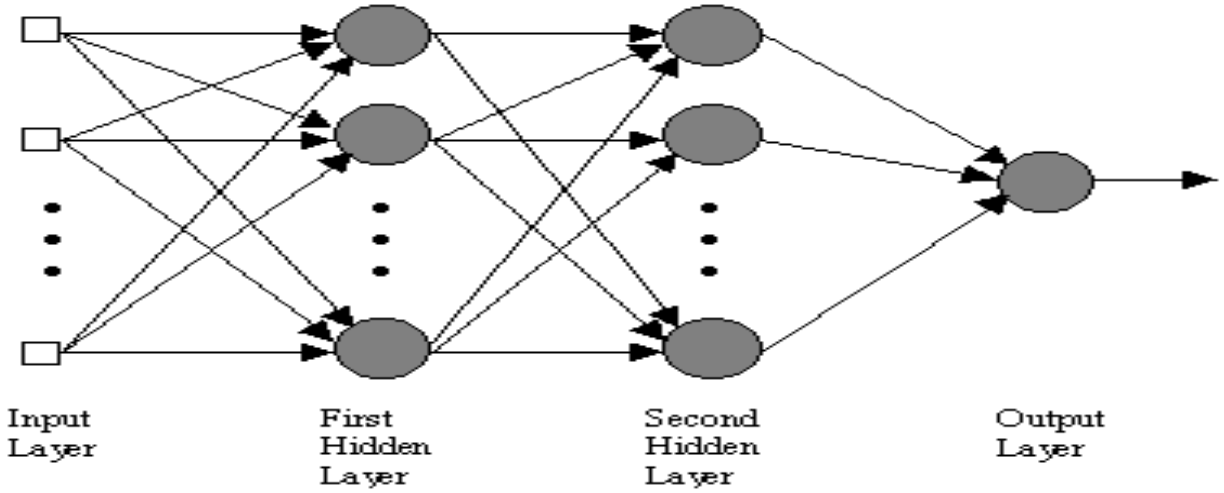
- Opportunities for HR domain
  - Deep Learning
  - Big Data
  - Learning from feedback & behavior
  - Conversational agents

Problems with classical Machine Learning

- Need to annotate a lot of data
- Need to design relevant features
  - Black art, intuition, trial & error
  - Time consuming

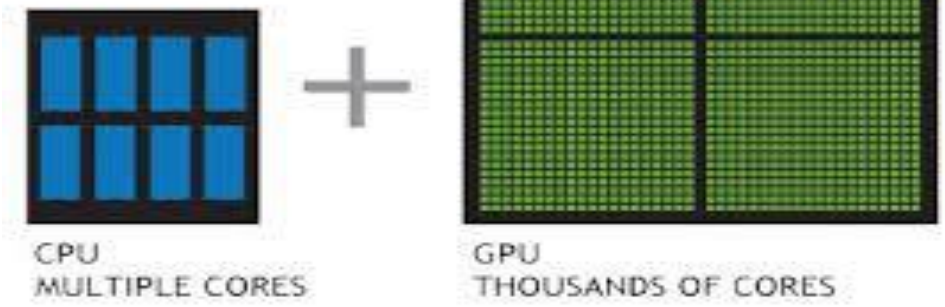
Neural networks

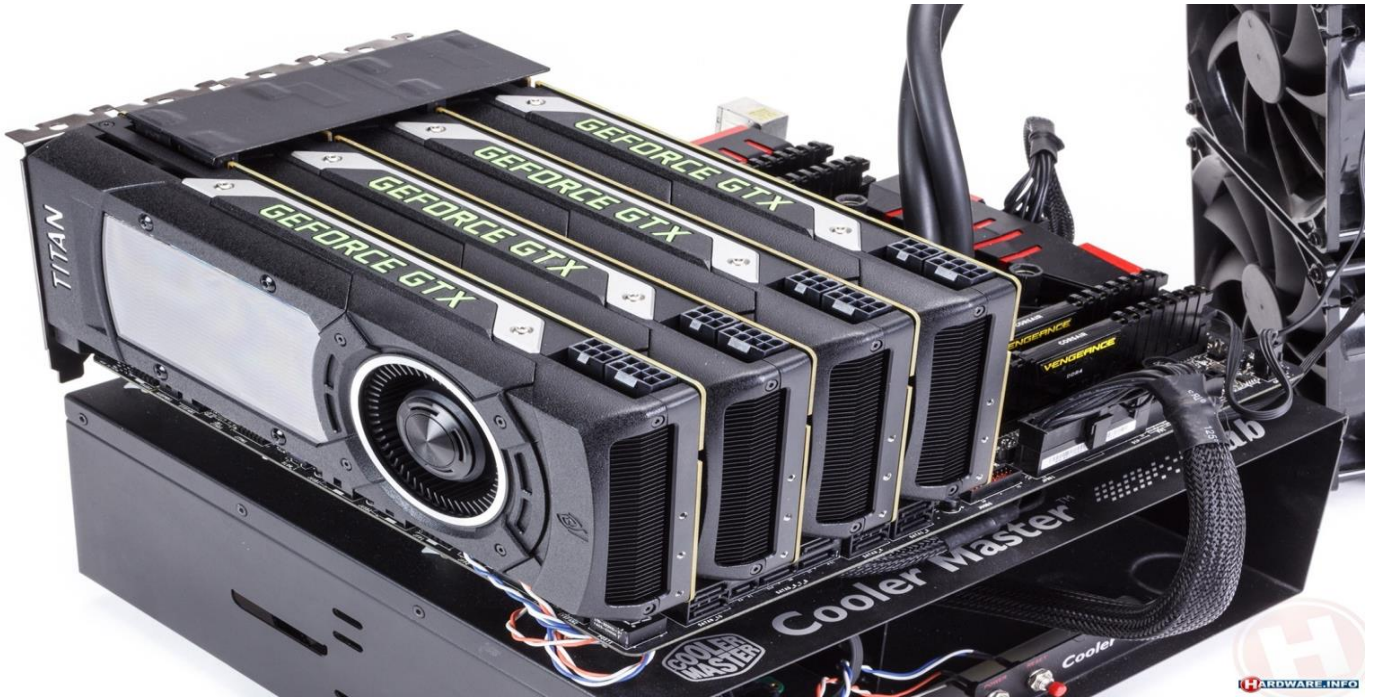
Inspired by the network of neurons in the brain



Marvin Minsky and Seymour Papert, 1969, Perceptrons: An Introduction to Computational Geometry
David Rumelhart, James L. McClelland, et al., 1986, Parallel Distributed Processing: explorations in the microstructure of cognition

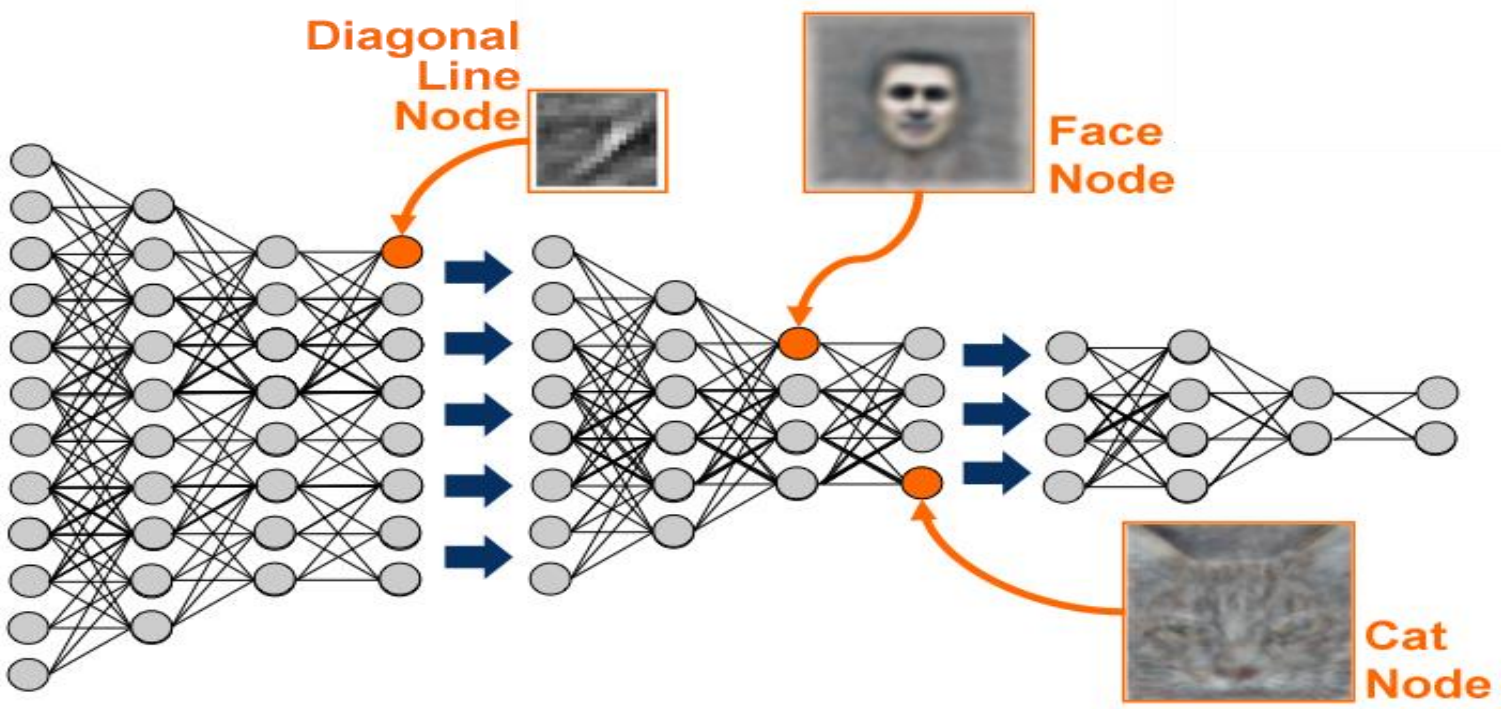
Better training, more power



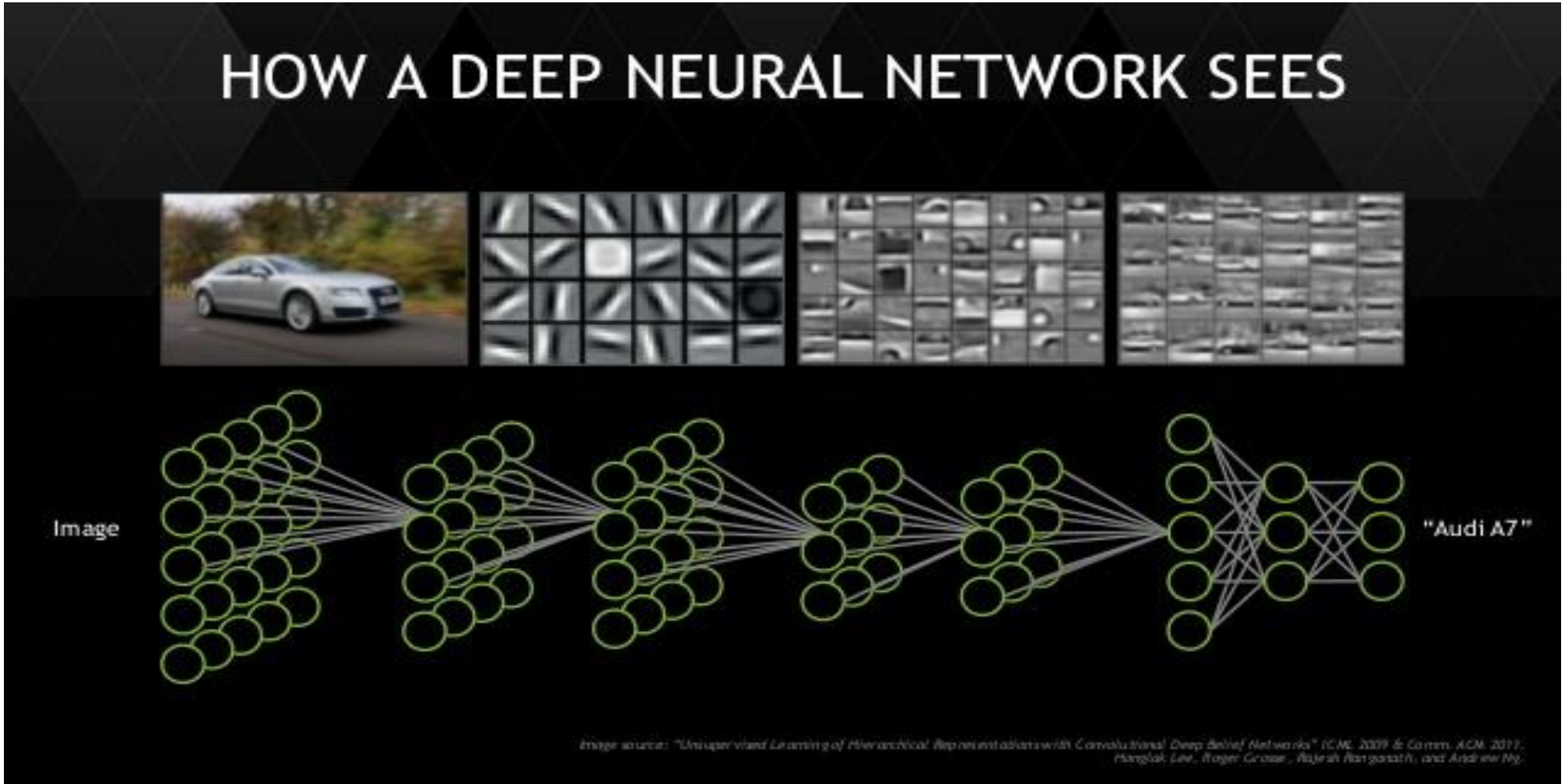


Deep Learning

Use lots of unannotated data
Automatically finds relevant features

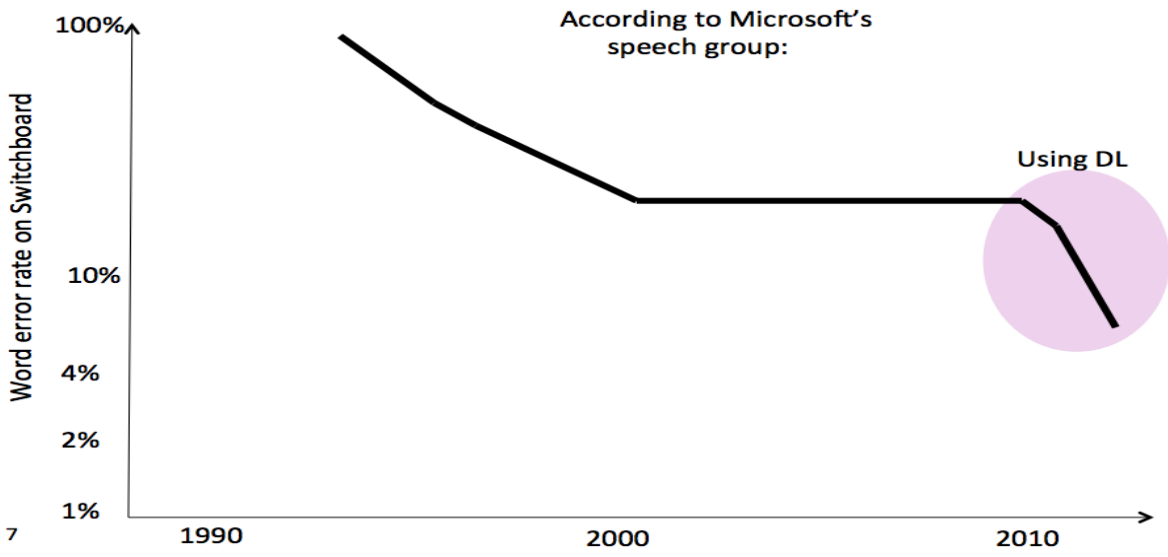


Hierarchically abstract information



Deep Learning Revolution

- State of the art
  - Computer vision
  - Speech recognition
  - Natural language processing
- How about HR domain?



Deep Learning to discover semantic relationships between words in HR

CEO

COO	0.84
CFO	0.83
SVP	0.76
CIO	0.76
CEO.	0.75
EVP	0.73
VP	0.72
CTO	
PRES	
DIRE	
CHA	
C.E.C	

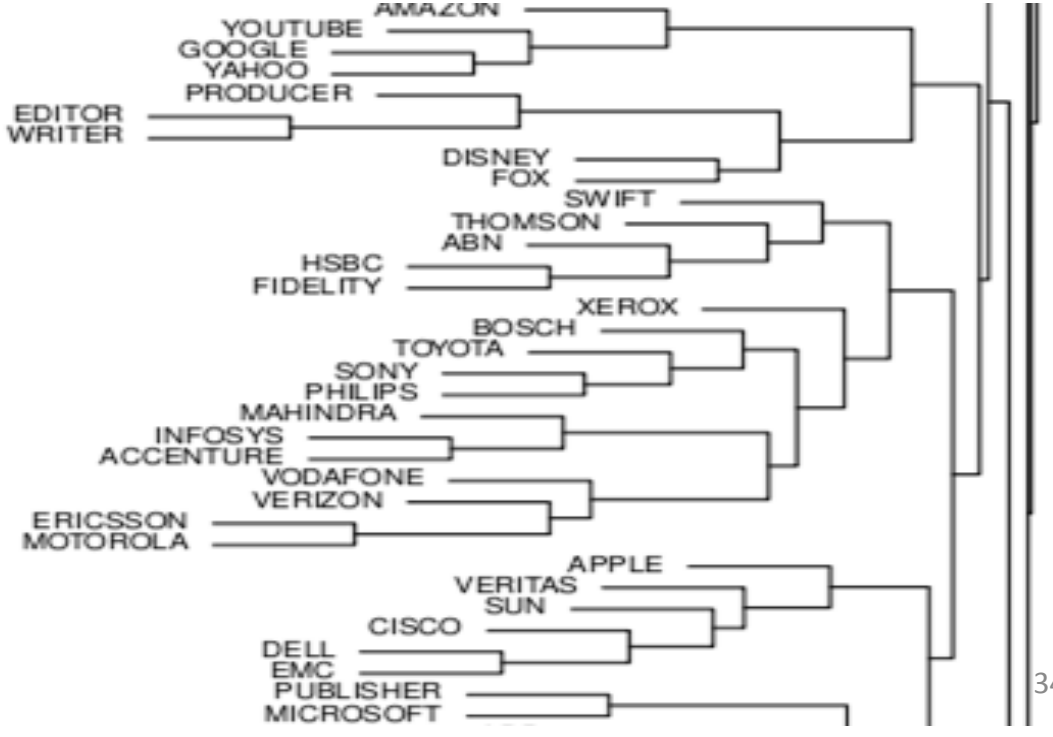
PLUMBER

CARPENTER	0.79
ROOFER	0.77
ELECTRICIAN	0.77
FITTER	0.77
BRICKLAYER	0.77
PLASTERER	0.75
LOCKSMITH	0.74

- Help parsing of unknow job titles
- Personalize parsing to a customer database
- Mine taxonomies of jobs, skills, names

Deep Learning

- Extract semantics from large amounts of CVs and jobs
  - Word similarities
  - Build taxonomies
  - Better parsing
  - Better matching



Big data

- Large amounts of data
  - Recruitment databases
  - Online jobs: 2M-32M per year
- Analytics
  - Average salary for a job or skill
  - Largest employers in my area
  - Professions/skills in highest demand
  - New jobs

Knowledge mining (ontologies)

- Taxonomies of jobs/skills
- Knowledge graph
  - Job <-> skill
  - Skill <-> skill
  - Job <-> Job
- Useful for searching

Network specialist CISCO
- troubleshooting
- CCNA
- OSPF

- audio-video production
- Final Cut Pro
- GoAnimate
- Haiku Deck
- Panopto

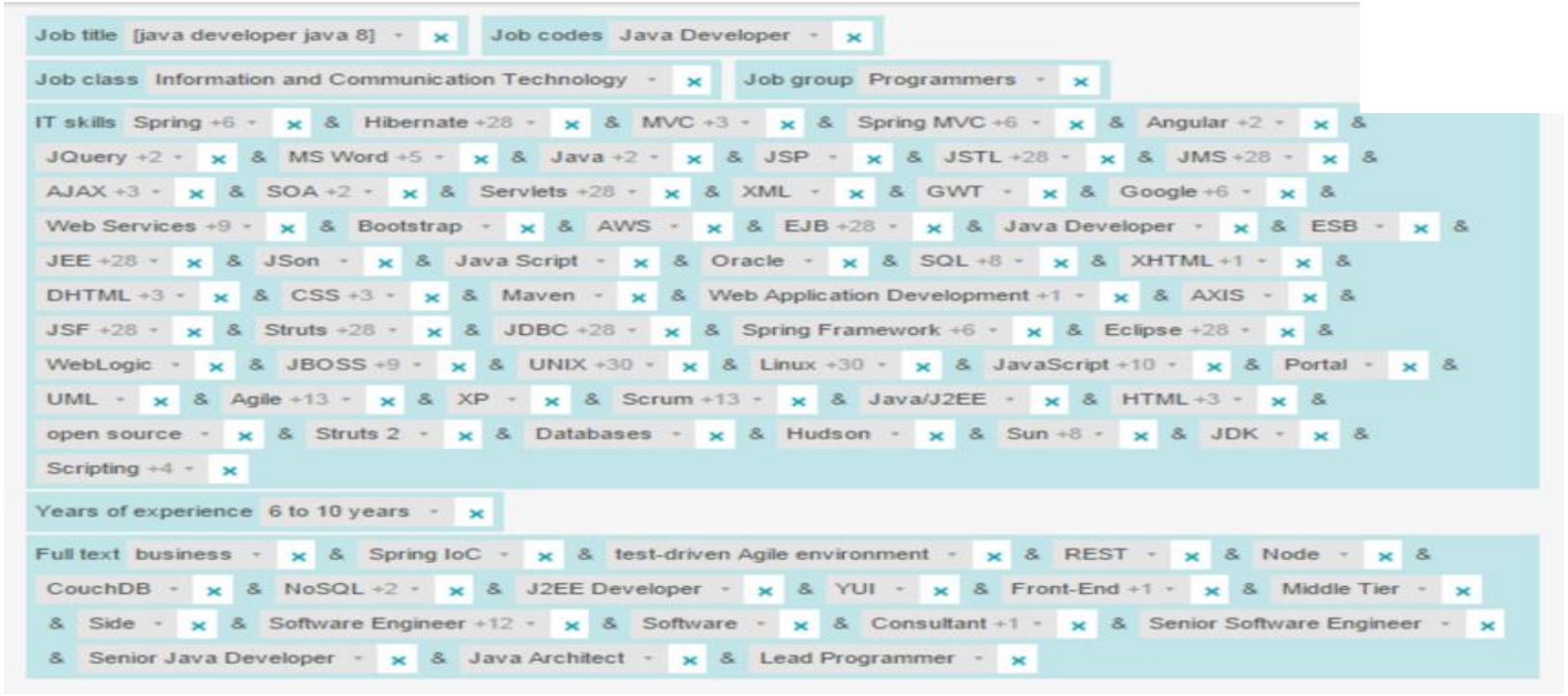
Example: Top 10 Keywords in German IT-jobs:

(per ISCO code)

All ISCO IT-related (~28 ISCO codes)	ISCO 1330 Information and communications technology service managers	ISCO 2511 Systems analysts	ISCO 2513 Web and multimedia developers	ISCO 2522 Systems administrators
informatik	informatik	sap	illustrator	windows
c	management	abap	photoshop	server
c++	service	informatik	css	administration
programmierung	it	customizing	design	vmware
software	englisch	entwicklung	html	support
embedded	entwicklung	programmierung	axure	systemintegration
entwicklung	automatisierungstechnik	net	usability	directory
softwareentwicklung	verantwortung	c	webdesign	linux
englisch	sap	wirtschaftsinformatik	adobe	active
systems	kommunikationsfähigkeit	java	kommunikationsdesign	systemadministration

Matching candidates to a job

Query produced by Match! from a job ad



Search! & learning from feedback

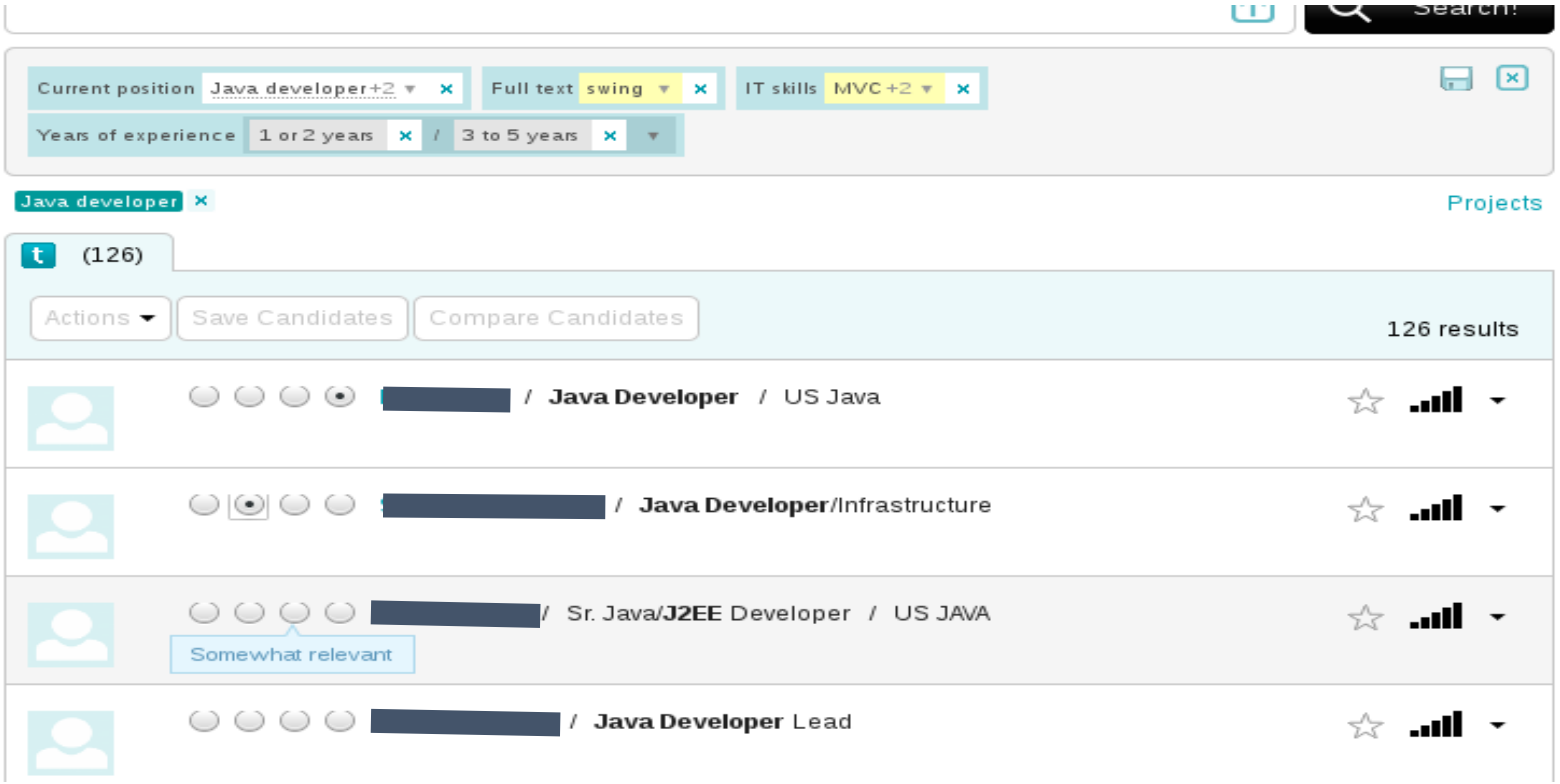
- Which candidate is the best?
  - Which criterion is more important?
  - How about combination of criteria?
  - How about domain biases?
- What if we can learn this automatically?
- Learning to Rank
  - Reorder a pool of candidates based on feedback

Feedback: hiring decisions

- Candidate hired for a job
- Need more info:
  - Applicants
  - Short listed applicants
  - Interviewed applicants

Feedback: explicit feedback

Recruiter marks good/bad candidates



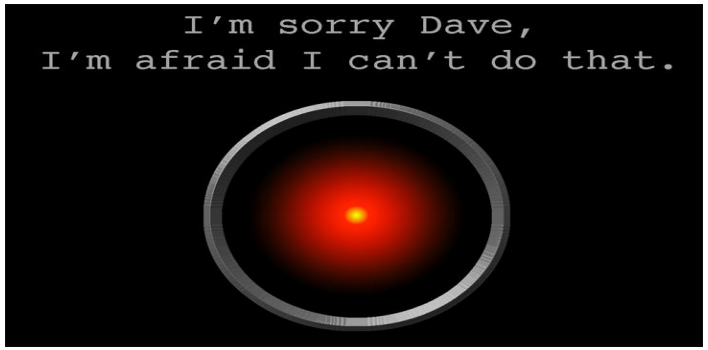
Implicit feedback: click data

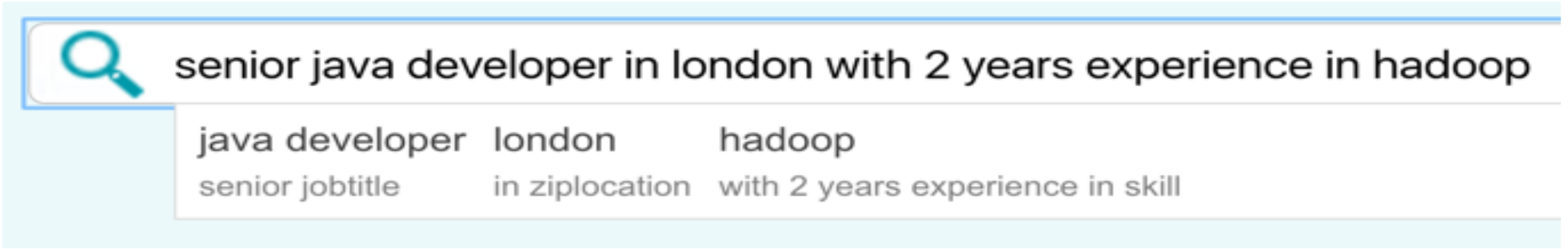
- Click/skip log
  - expand metadata, open the CV, or export/save
  - time to click, after query execution
  - dwell time: time spent before clicking next document
- Combine with explicit feedback

- Personalize search: beta stage
- 9-21% increase in performance

Dialog systems / Recruitment assistants?

- Success in other industries
  - Customer support
  - Siri, Google Now
- Natural language search
- Textio
- Wade & Wendy

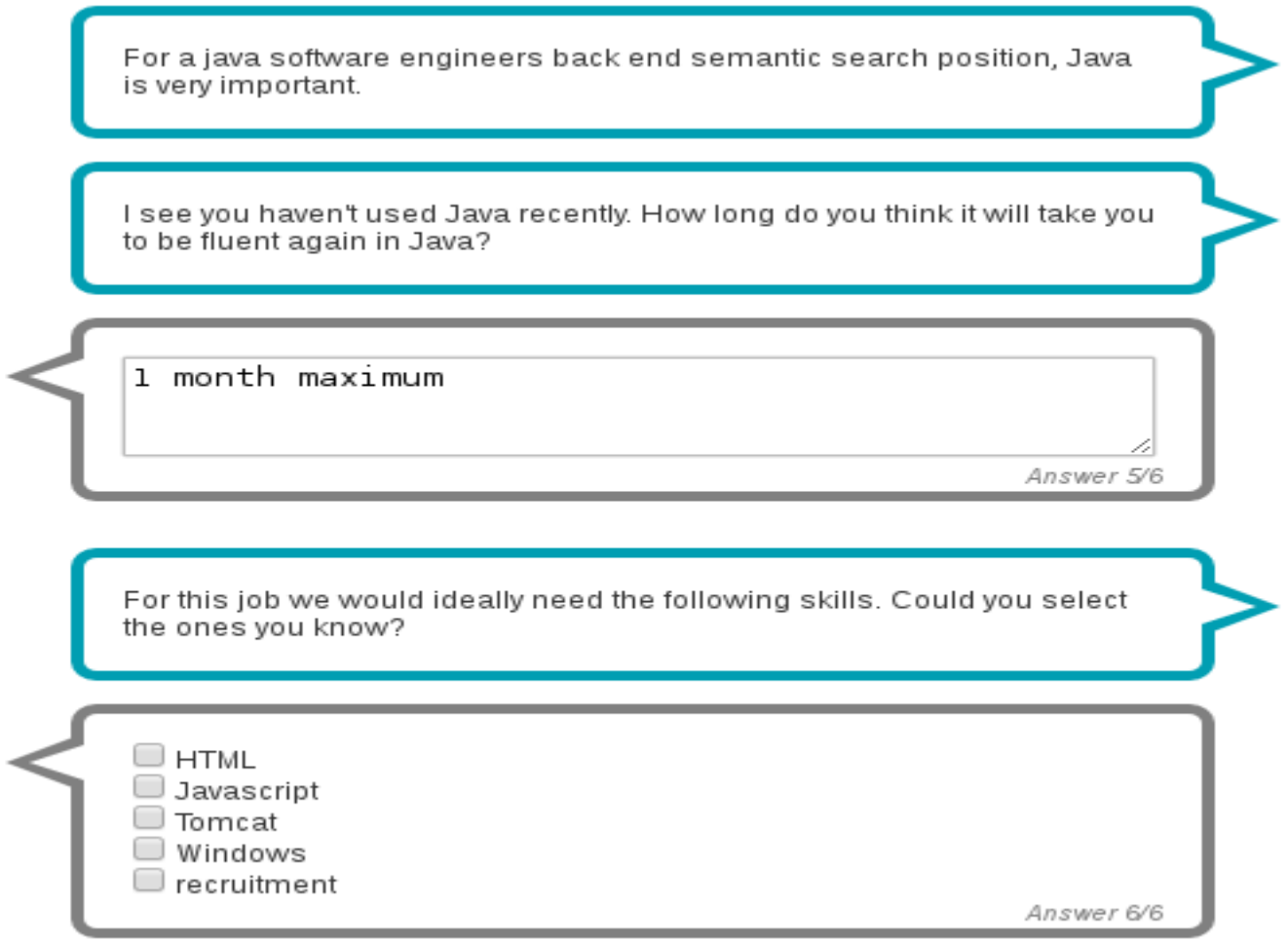




Talking to the candidate?

- Discuss with the candidate about the match with the job
  - Gap in experience, skills, education
  - Relocation
  - Engage the candidate

Example: Gap in skills



Makita
BOSCH
AVT
SKIL
FRAGILE

Wrap up

- Accelaration of research & technology in recruitment
- The data we process contain important knowledge of recruitment (unstructured CV data, job ads)
- For recruitment to become more efficient and more strategic, we need power tools
- Exciting times to be in the recruitment business