A New Approach to Building a Skills Taxonomy

ESCoE Human Capital Workshop

22nd February 2022

Liz
Gallagher

India
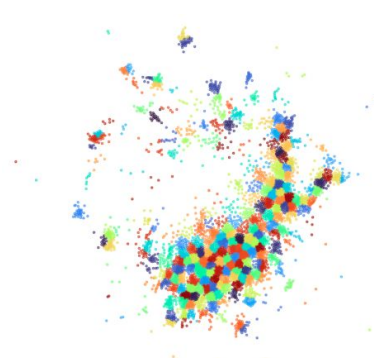Kerle

Cath
Sleeman

George
Richardson

Agenda

nesta

Introduction

We have introduced a **new, data-driven approach** to building a UK skills taxonomy.

This improves on our original approach in a **three** ways:

1. We **no longer rely on a proprietary list of skills** that had been extracted from job adverts and, as a result, we can detect previously unseen skills.

2. We can **publish the entire taxonomy and codebase**.

3. The previous taxonomy was developed almost four years ago, so we can now offer **a more up-to-date view** of UK skill demands.
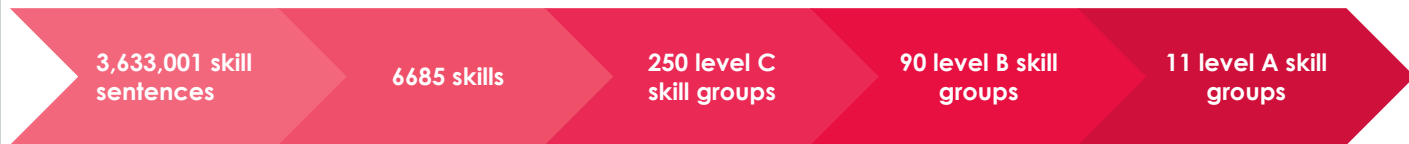
nesta

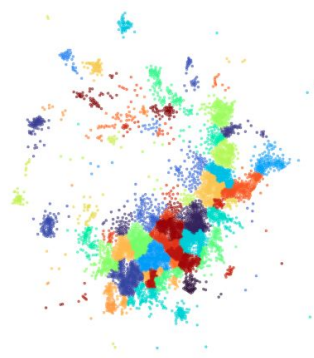# Introduction

| 3,633,001 skill sentences | 6685 skills | 250 level C skill groups | 90 level B skill groups | 11 level A skill groups |

Skills coloured by the 250 level C groups

Skills coloured by the 90 level B groups
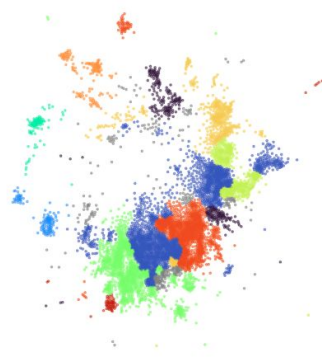
Skills coloured by the 11 level A groups

Methodology
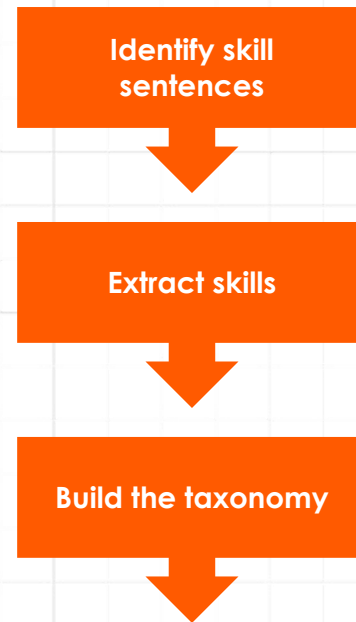
Identify skill sentences

Extract skills

Build the taxonomy

nesta

Identify skill sentences

Extract skills

Build the taxonomy

Why?

Remove 'noisy' text in job adverts unrelated to skills

nesta

Identify skill sentences

Extract skills

Build the taxonomy

Why?

Remove 'noisy' text in job adverts unrelated to skills

How?

Trained a supervised machine learning model to classify 'skill' sentences

nesta

Identify skill sentences

→

Extract skills

↓

Build the taxonomy

↓

Why?

Remove 'noisy' text in job adverts unrelated to skills

How?

Trained a supervised machine learning model to classify 'skill' sentences

Results

On a held-out test set, the model was able to correctly identify sentences containing a skill on 93% of occasions

nesta

Labelled Job Advert Example

Identify skill sentences

Extract skills

Build the taxonomy

salt is a PR agency of 40 people based in Putney, South West London (soon to relocate to Oval) and Singapore. Put simply, we work to help our clients communicate better. Our approach is informed by the basic truth of PR, that you can't force someone to tell your story. We consciously combine the worlds of corporate and consumer communications into award-winning work for our clients. We love what we do and it shines through in the work we deliver and the atmosphere in our London office and abroad.

Due to our recent expansion and drive for new business, we would like to recruit an additional person for our Design Team. The Design Executive will assist with designing and creating innovative and eye-catching designs for presentations and client materials, including advertorials and magazine layouts [skill]. Being able to deliver successful new business presentations [skill] and package our existing work in a way that differentiates our offering within the industry is an essential part of our strategic plan and as part of that, you will provide an exceptional graphic design service [skill] for the company. You will support the Design Manager in maintaining a consistent house style of all material across the agency, in partnership with our client teams [skill].

This role would suit a Graphic Design graduate with two years industry experience, ideally gained in a media environment.

Essential skills for this role include:
* Advanced PowerPoint and InDesign skills [skill]
* Experience of using Adobe Suite [skill]
* Graphic Design Degree
* Two years industry experience, ideally gained in a media environment.
* An eye for detail [skill]
* Ability to distinguish "good" from "outstanding" in a presentation. [skill]
* Experience of both designing AND building graphic layouts. [skill]

Around 25% of sentences in job adverts are classified as skill sentences

nesta

Skills Extraction Pipeline

Identify skill sentences

Extract skills

Build the taxonomy

Skill sentences

| Develop Unit Tests. |
| Can prepare food for catering. |
| Experience with unit testing. |
| Good food hygiene standards. |
| Working in a kitchen. |
| Cooking fresh food. |
| Able to code up unit tests. |
| Is experienced with kitchen management. |
| Ability to develop unit tests full uk or equivalent driving licence good customer facing skills initiative. |
| … |

Mask and prune sentences

| Develop Unit Tests. |
| [MASK] prepare food [MASK] catering. |
| Experience [MASK] unit testing. |
| Good food hygiene standards. |
| Working [MASK] [MASK] kitchen. |
| Cooking fresh food. |
| [MASK] [MASK] code [MASK] unit tests. |
| [MASK] experienced [MASK] kitchen management. |
| … |

BERT sentence embeddings

| [0.81, 0.22, 0.43, . . . , 0.76] |
| [0.53, 0.67, 0.29, . . . , 0.55] |
| … |
| … |
| … |
| … |
| … |
| … |
| |

Skill name: Unit testing.
Sentences: Develop Unit Tests, Experience with unit testing, Able to code up unit tests.

Cluster to find skills

Reduced embeddings

| [0.68, 0.33] |
| [0.27, 0.91] |
| … |
| … |

Skill name: Kitchen management.
Sentences: Good food hygiene standards, Working in a kitchen, Is experienced with kitchen management.

Skill name: Meal preparation.
Sentences: Can prepare food for catering, Cooking fresh food.

nesta

Skill Name Examples

Identify skill sentences

Extract skills

Build the taxonomy

microsoft office programme

Food and cleaning

culinary expertise

auditing accounting

hygiene standard

Financial

vehicle maintenance
risk management compliance

reporting hazard

repair faulty equipment

financial accounting

clinical research

compliance data protection

forecasting revenue

Care

office supply

provide regular reporting

autism challenging behaviour

global marketing

advise customer

Customer service

childrens safeguarding

product development

communication customer service

design development

data analysis
accurate record keeping

supporting pupil challenging behaviour

software testing

strong communicator

query language

team member

familiar crm

using html

meeting diary management

creativity innovation

good relationship client

Personal attributes

Language

fluency cantonese

time management

nesta

Building the Taxonomy Pipeline
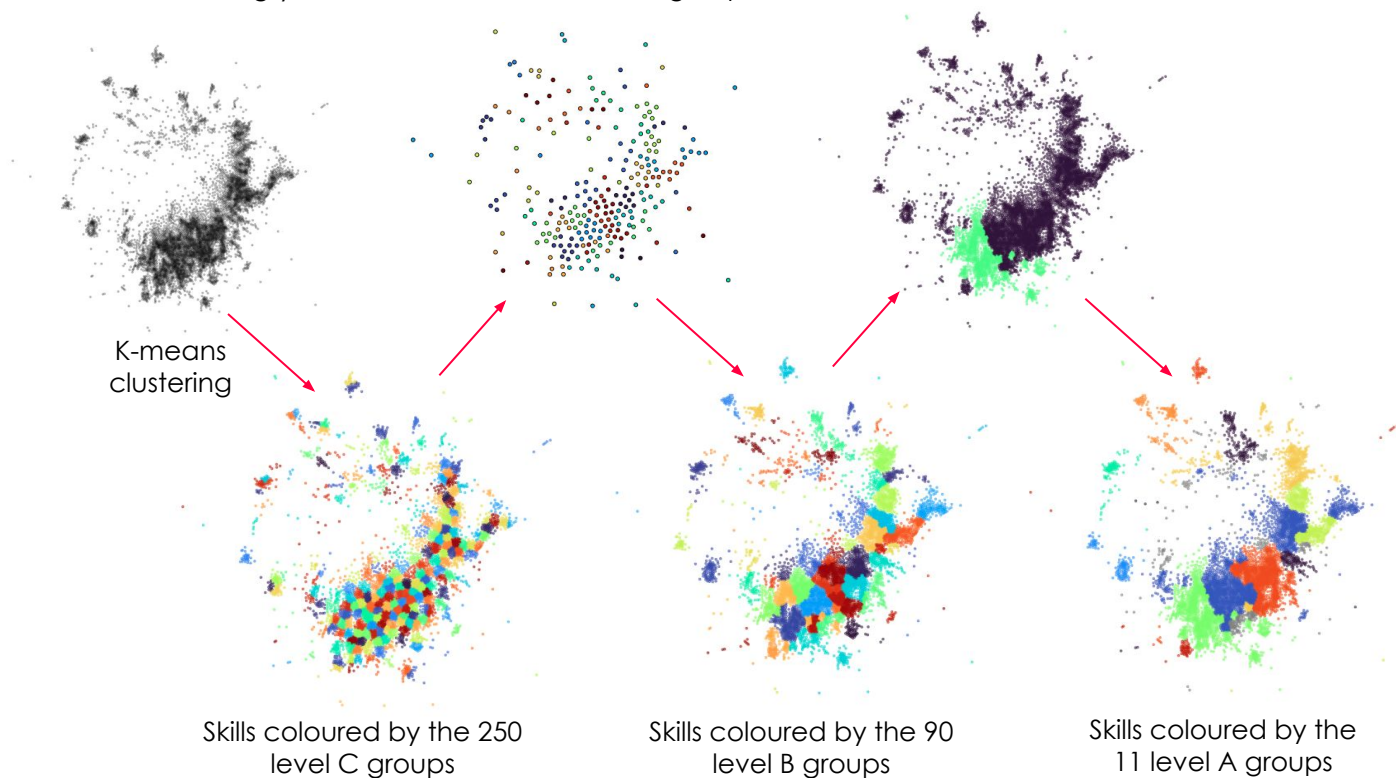
Identify skill sentences

Extract skills

Build the taxonomy

Skills by 2D coordinates (average reduced sentence embeddings)

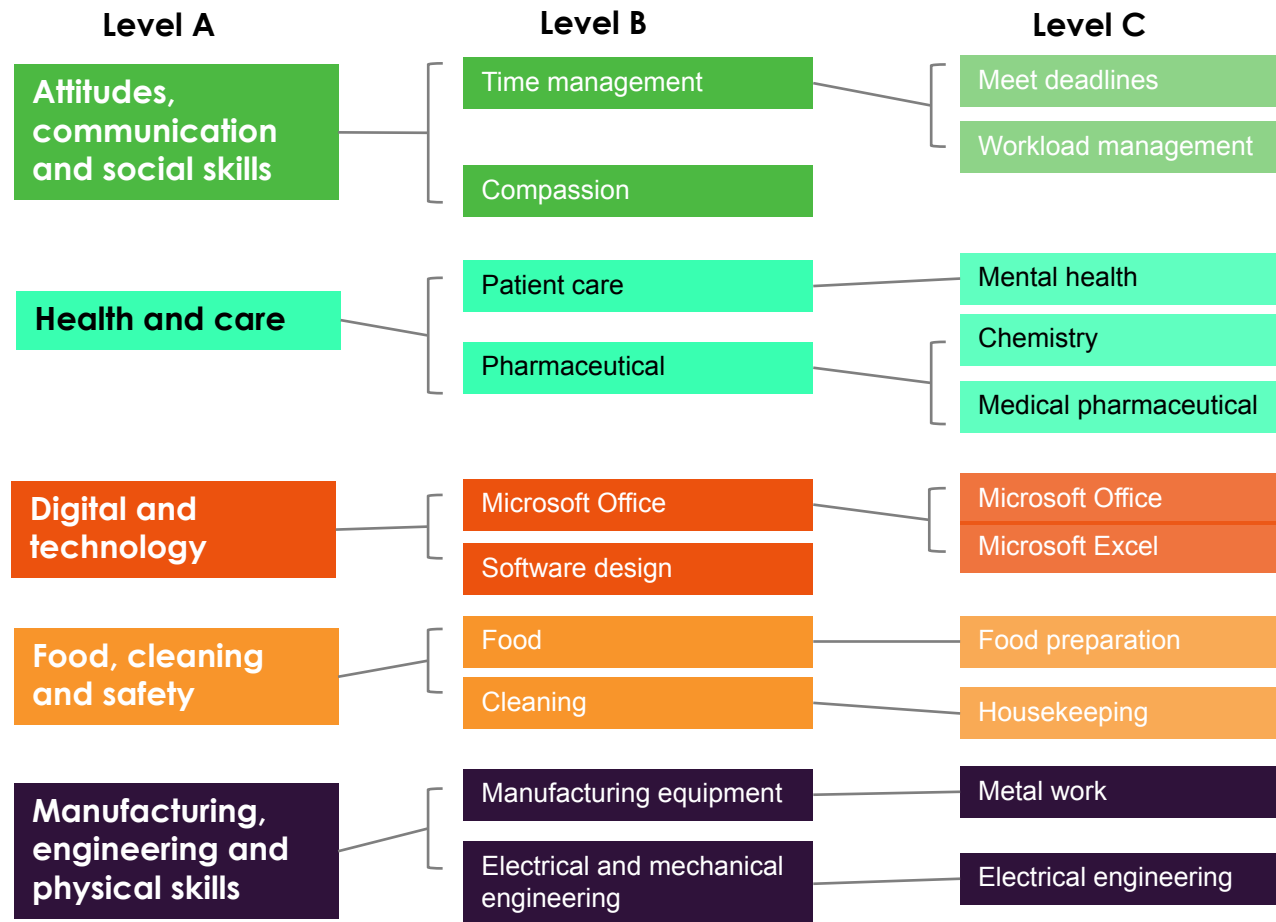K-means clustering of average coordinates of each level C group

Manual clustering of each level B group

K-means clustering

Skills coloured by the 250 level C groups

Skills coloured by the 90 level B groups

Skills coloured by the 11 level A groups

nesta

Taxonomy examples

Identify skill sentences

↓

Extract skills

↓

Build the taxonomy

↓

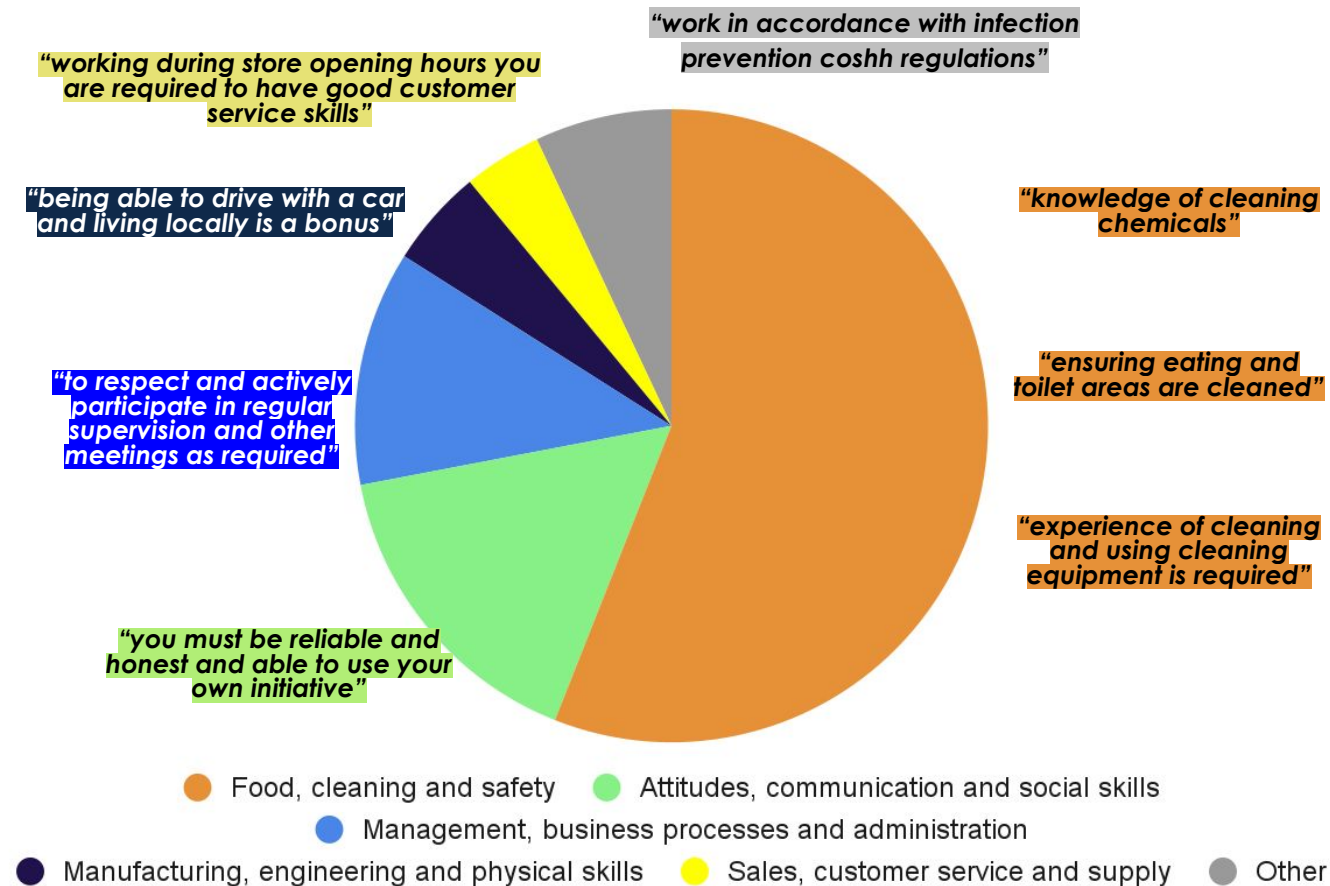| Level A	Level B	Level C
Attitudes, communication and social skills	Time management	Meet deadlines
		Workload management
	Compassion	
Health and care	Patient care	Mental health
	Pharmaceutical	Chemistry
		Medical pharmaceutical
Digital and technology	Microsoft Office	Microsoft Office
		Microsoft Excel
	Software design	
Food, cleaning and safety	Food	Food preparation
	Cleaning	Housekeeping
Manufacturing, engineering and physical skills	Manufacturing equipment	Metal work
	Electrical and mechanical engineering	Electrical engineering

nesta

Analysis

- Evaluation
- Job skills
- Transversal skills
- Which regions specialise in which skills?
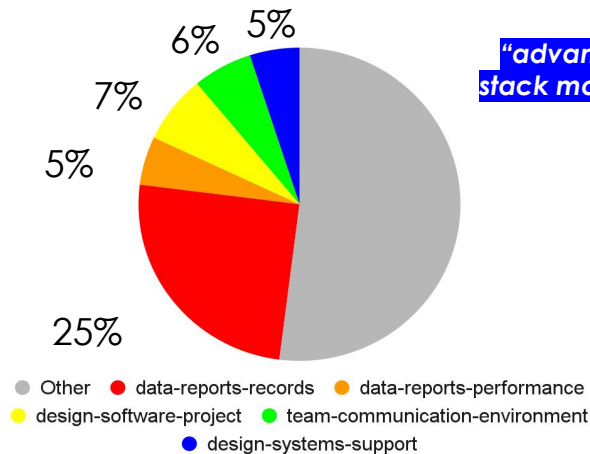- How did skill demands change after the onset of COVID-19?

nesta

Analysis

Does the taxonomy make sense?

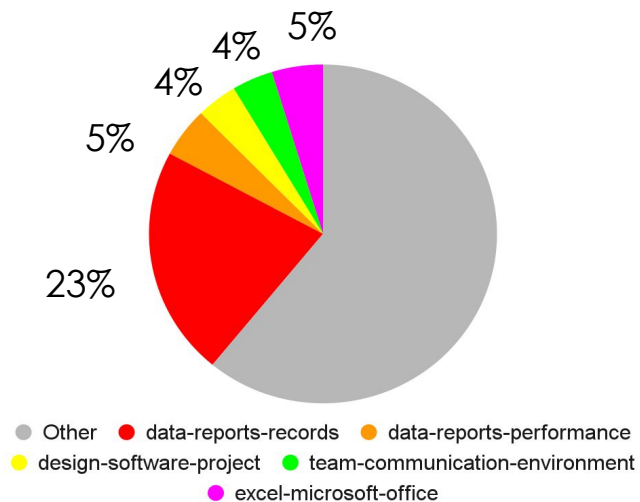| Level A name	Jobs with the highest proportion of skills in this group
Attitudes, communication and social skills	'Trainee Recruitment Consultant (47.21%)', 'Graduate Recruitment Consultant (45.87%)', 'Hairdressing Apprentice (45.56%)'
Cognitive skills and languages	'Maths Teacher (36.95%)', 'Teacher of Maths (29.78%)', 'Teacher of Mathematics (12.02%)'
Childcare and Education	'**History Teacher** (79.07%)', 'Music Teacher (75.29%)', 'Geography Teacher (72.4%)'
Health and care	'Practice Nurse (44.3%)', 'Advanced Nurse Practitioner (37.63%)', 'Nurse Practitioner (34.47%)'
Finance, auditing, legal and compliance	'Senior Financial Accountant (51.92%)', 'Finance Officer (50.7%)', 'Financial Accountant (50.11%)'
Sales, customer service and supply	'Sales Order Processor (33.33%)', 'Retail Assistant (32.26%)', 'Procurement Specialist (29.85%)'
Management, business processes and administration	'Head of Marketing (54.42%)', 'Marketing Manager (53.44%)', 'Product Marketing Manager (53.23%)'
Digital and technology	'Automation Tester (67.27%)', 'DevOps Engineer (66.6%)', 'Software Test Engineer (65.55%)'
Food, cleaning and safety	'**Part Time Cleaner** (63.79%)', 'Cleaner (56.4%)', 'Kitchen Assistant (52.98%)'
Manufacturing, engineering and physical skills	'Welder Fabricator (57.6%)', 'Fabricator Welder (55.86%)', 'HGV Mechanic (48.87%)'

nesta

Analysis

Job skills for 'Cleaner' job adverts

"work in accordance with infection prevention coshh regulations"

"working during store opening hours you are required to have good customer service skills"

"being able to drive with a car and living locally is a bonus"

"knowledge of cleaning chemicals"

"to respect and actively participate in regular supervision and other meetings as required"

"ensuring eating and toilet areas are cleaned"

"experience of cleaning and using cleaning equipment is required"

"you must be reliable and honest and able to use your own initiative"

- Food, cleaning and safety
- Attitudes, communication and social skills
- Management, business processes and administration
- Manufacturing, engineering and physical skills
- Sales, customer service and supply
- Other

nesta

Analysis

Job skills - Data Scientist vs Data Analyst

"advance the continuous validation of our stack models used for product development"

6% 5%
7%
5%
25%

- Other
- data-reports-records
- data-reports-performance
- design-software-project
- team-communication-environment
- design-systems-support

"core skills for this data analyst role advanced excel skills essential"

4% 5%
4%
5%
23%

- Other
- data-reports-records
- data-reports-performance
- design-software-project
- team-communication-environment
- excel-microsoft-office

nesta

Transversal skills

Transversal skills were identified using a co-occurrence network approach. These skills relate to communication and interpersonal skills.
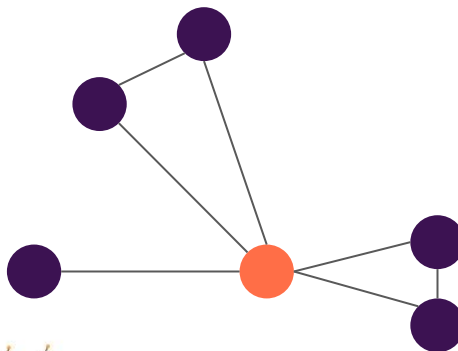
"liaison with client employees customers agencies and subcontractors"

"maintain and develop successful client relationships"

"comfortable with managing meetings and dealing with the client and their representatives"

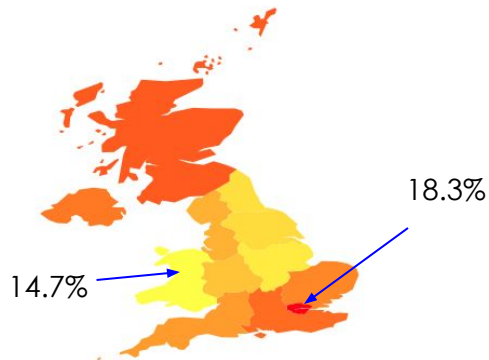*"the ability to work within a team is critical to this position"*

"support with the recruitment and training of the team"

nesta

Analysis
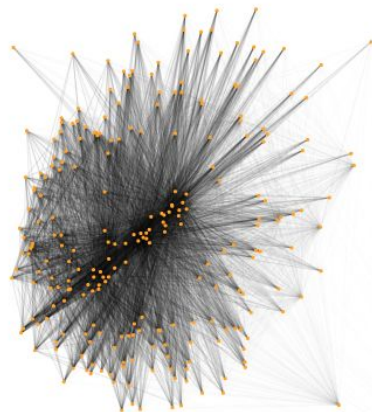
Which regions specialise in which skills?

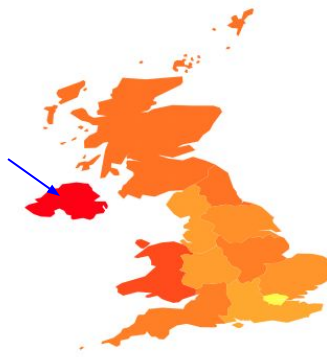### Digital and technology

14.7%

18.3%

Manufacturing, engineering and physical skills
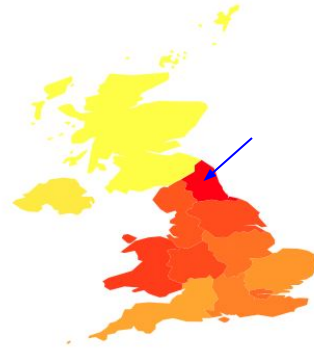

7%

2.6%
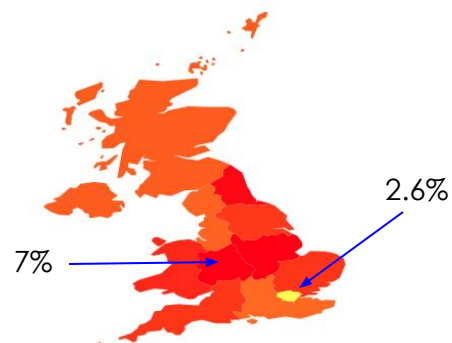
Cognitive skills and languages
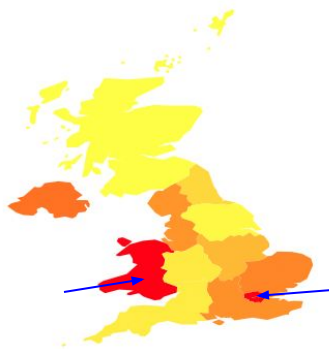

Food, cleaning and safety

Childcare and Education

nesta

Which regions specialise in which skills?

Our taxonomy reflects the **strong logistics sector in the Midlands**; the region contains most of the 'golden triangle' for logistics, from which it is possible to reach approximately 90% of the country within four hours.

http://www.rcslogistics.co.uk/blog-and-news/golden-triangle/2935
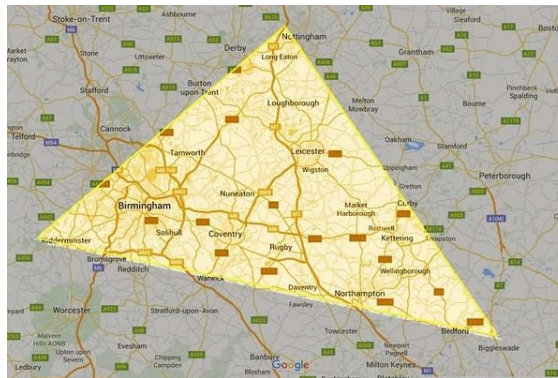
We found that the East Midlands has a higher proportion of skills to do with driving, engineering and stock management as compared to the rest of the country.

"ideal candidate delivery driving experience and full clean uk driving license is essential"

"knowledge and experience of interpreting engineering drawings"

"organise and maintain stock codes descriptions and storage bins"

nesta

How did skill demands change after the onset of COVID-19?

"the post holder will be expected to fully participate in and support the general surgery oncall rota"

"Health-clinical-medical"

231%

"Chef-food-kitchen"

58%

"the ability to oversee the day to day running of the kitchen and one other part time chef"

"Customer-service-customers"

79%

"most importantly ensure customer service standards are maintained to a high level always"

Conclusions

The key strength of our method is that it does not make presumptions about either the underlying list of skills or the grouping of these skills. This allows us to detect new skills (and new skill groups) as soon as they emerge within job adverts.

We hope our taxonomy could:

1. Be used to inform revisions to existing taxonomies, such as ESCO, by identifying new skills and new skill groups.

2. Act as the base (or starting point) for a UK-specific skills taxonomy. Experts could add branches to the taxonomy that capture skills which are not typically mentioned in online adverts.

3. Be used to provide rapid analysis of shocks, such as COVID-19.

4. Be used to enrich career advice and training in certain occupations, by providing very detailed descriptions of the skills required in each role.

nesta

Next steps

Compare the skills extracted in this taxonomy with the expert-derived list of ESCO skills

Examine the differences in skill demands over time, allowing us to see emerging and declining skills

Use the skill sentence classifier in our Open Jobs Observatory project

nesta

Thank you

nesta

Appendix

nesta

Sentence Classifier
Classification Report

	precision	recall	f1-score	support
non-skill	0.90	0.97	0.93	694
skill	0.93	0.78	0.85	344
accuracy			0.90	1038
macro avg	0.90	0.90	0.90	1038
weighted avg	0.90	0.90	0.90	1038

nesta