



CEDEFOP

European Centre for the Development
of Vocational Training

50
YEARS
SHAPING LEARNING AND
SKILLS FOR EUROPE

Cedefop advancements in OJA- driven Skills Intelligence

Using data from the web to shape next-generation
labour market and skills intelligence

Adam Tsakalidis
Cedefop Expert (Skills Intelligence & Foresight)

Brussels, Belgium | 01 Apr 2025

Introduction



50 years: Vocational Education & Training: Policies & Implementation

➤ Dept. VET & Skills

Introduction



SHAPING
LEARNING
AND SKILLS
FOR EUROPE

50 years: Vocational Education & Training: Policies & Implementation

➤ Dept. VET & Skills



30 years: Thessaloniki, Greece

By [CeeGee](#), licensed
under [CC BY-SA 4.0](#)

Introduction



50 years: Vocational Education & Training: Policies & Implementation

- Dept. VET & Skills



30 years: Thessaloniki, Greece



20 years: First steps towards Skills Intelligence:

- Quantitative & qualitative sources
- Identify trends, skill gaps/mismatches, provision of education/training...

Introduction



50 years: Vocational Education & Training: Policies & Implementation

- Dept. VET & Skills



30 years: Thessaloniki, Greece



20 years: First steps towards Skills Intelligence:

- Quantitative & qualitative sources
- Identify trends, skill gaps/mismatches, provision of education/training...



10 years: Skills Intelligence via analysing Online Job Vacancies

- Big data
- Real-time

Timeline

Cedefop Skills Intelligence kicks off

2005



Timeline

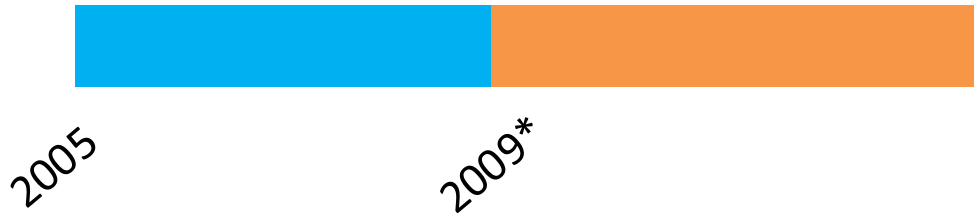
Cedefop Skills Intelligence kicks off



Timeline

Cedefop Skills Intelligence kicks off

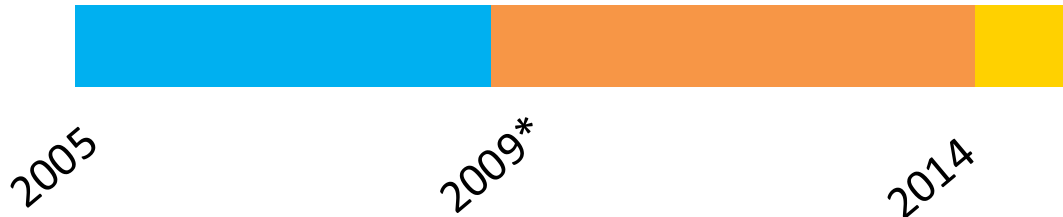
On the background: Big data



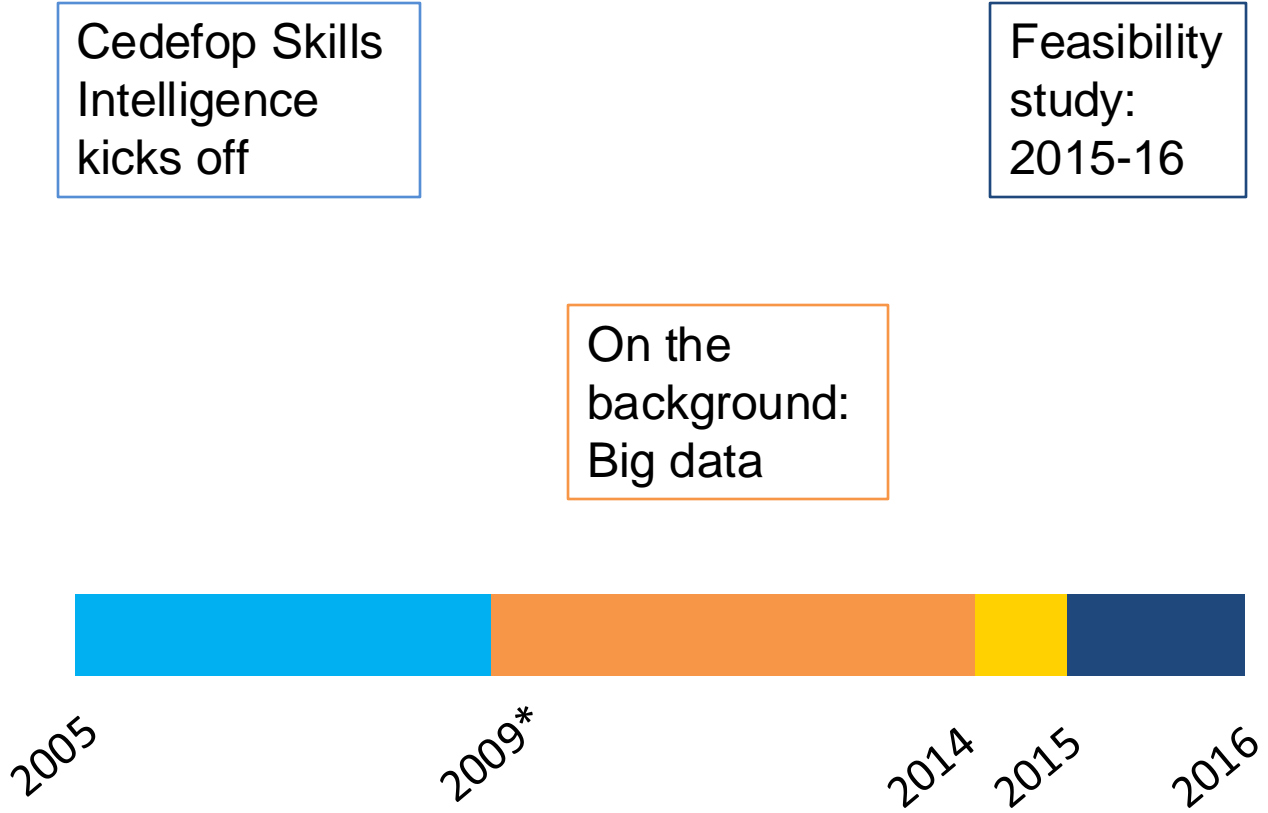
Timeline

Cedefop Skills Intelligence kicks off

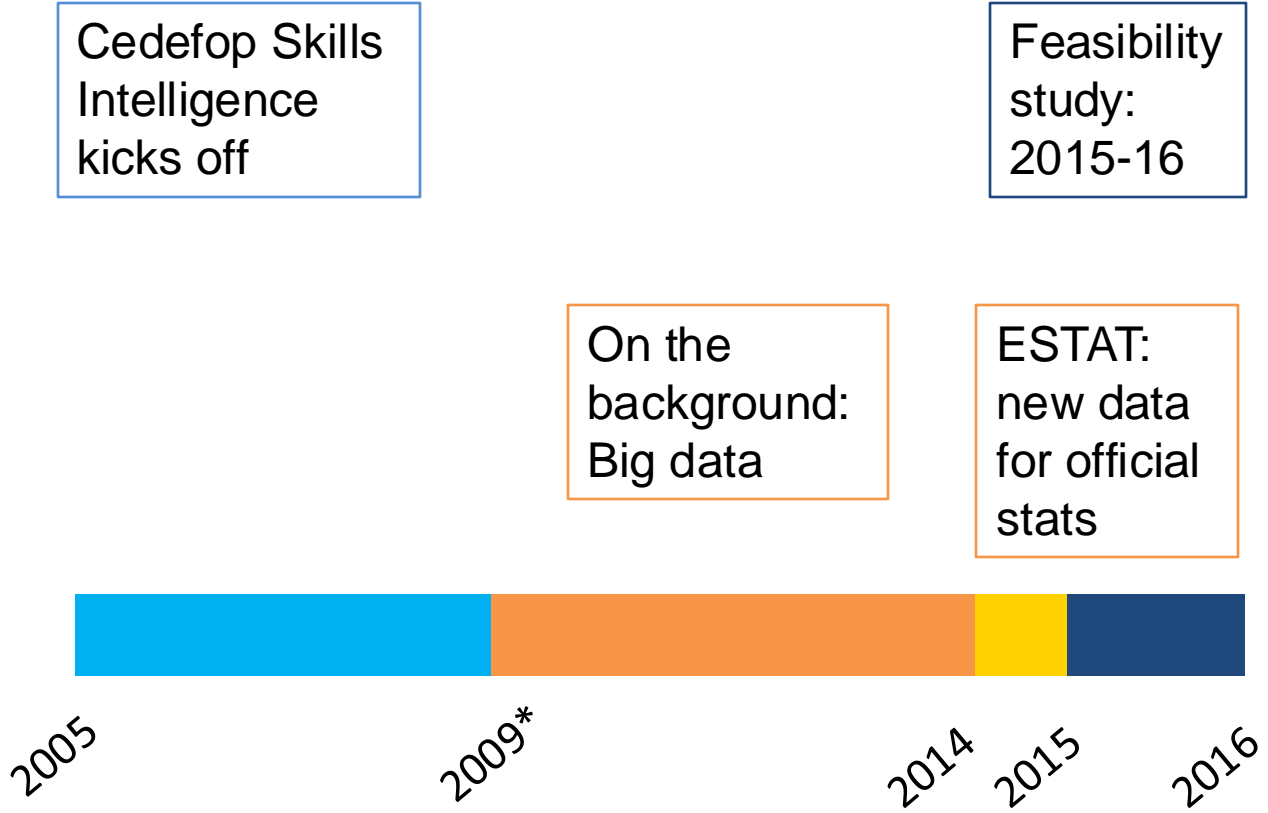
On the background: Big data



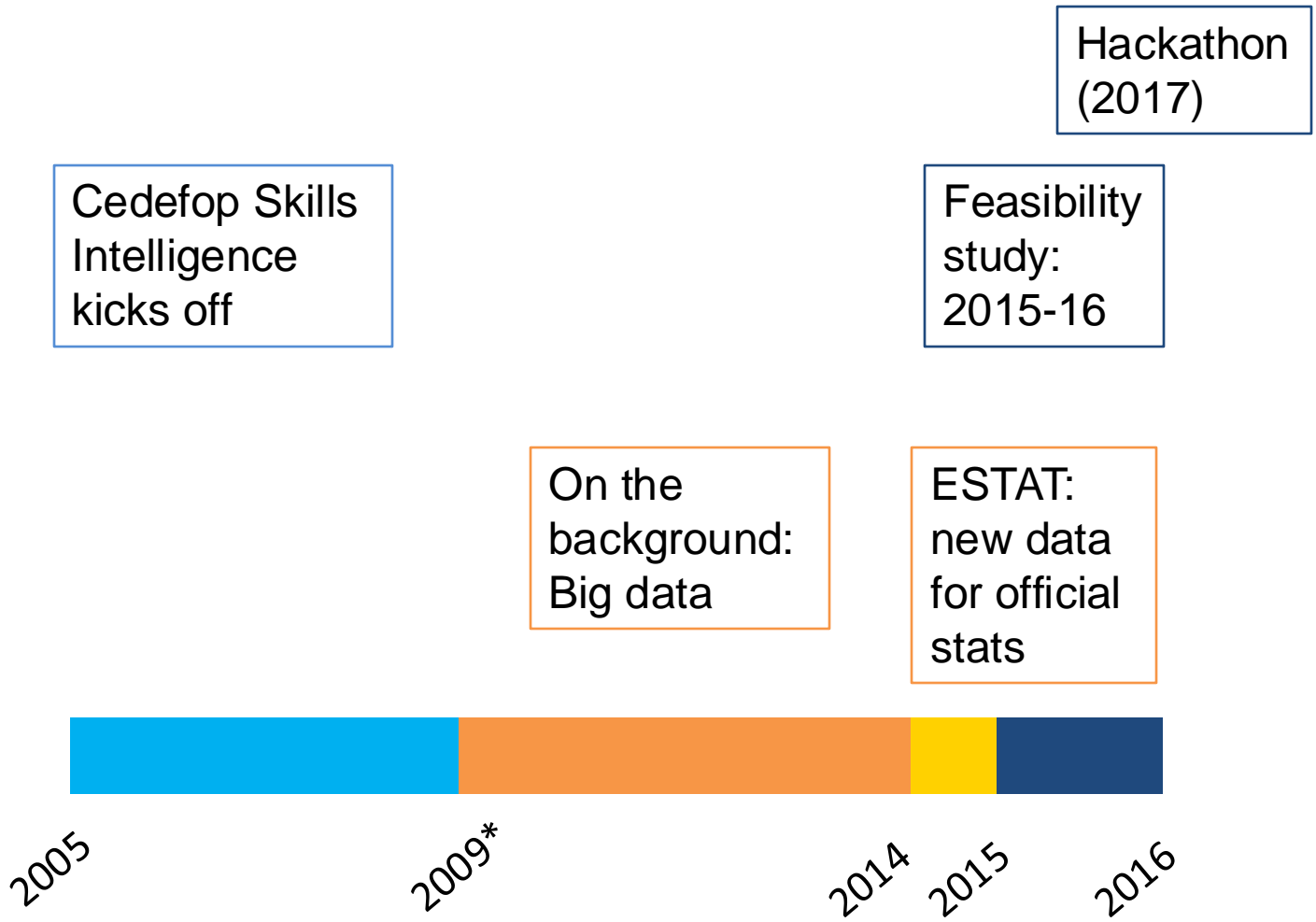
Timeline



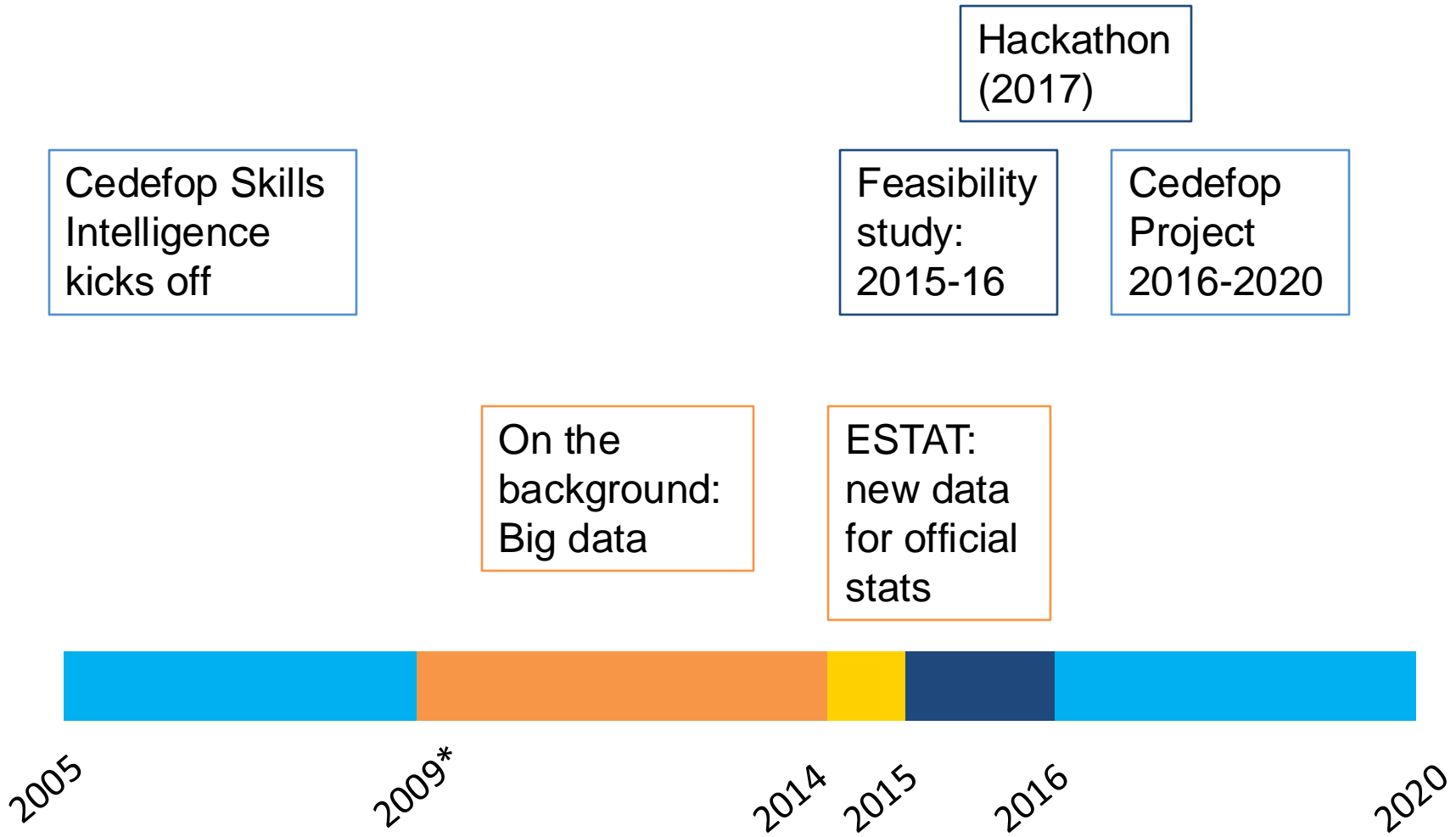
Timeline



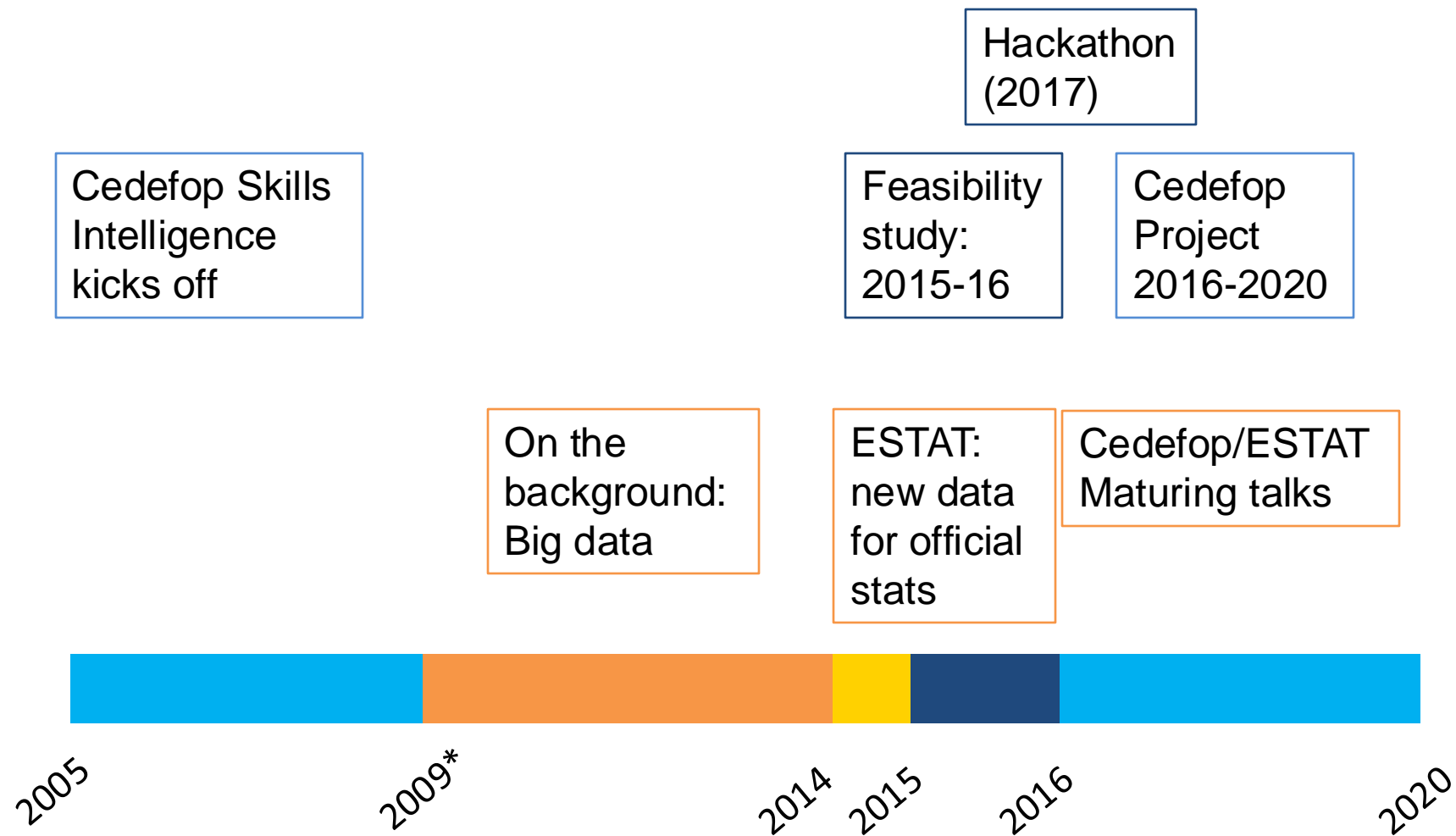
Timeline



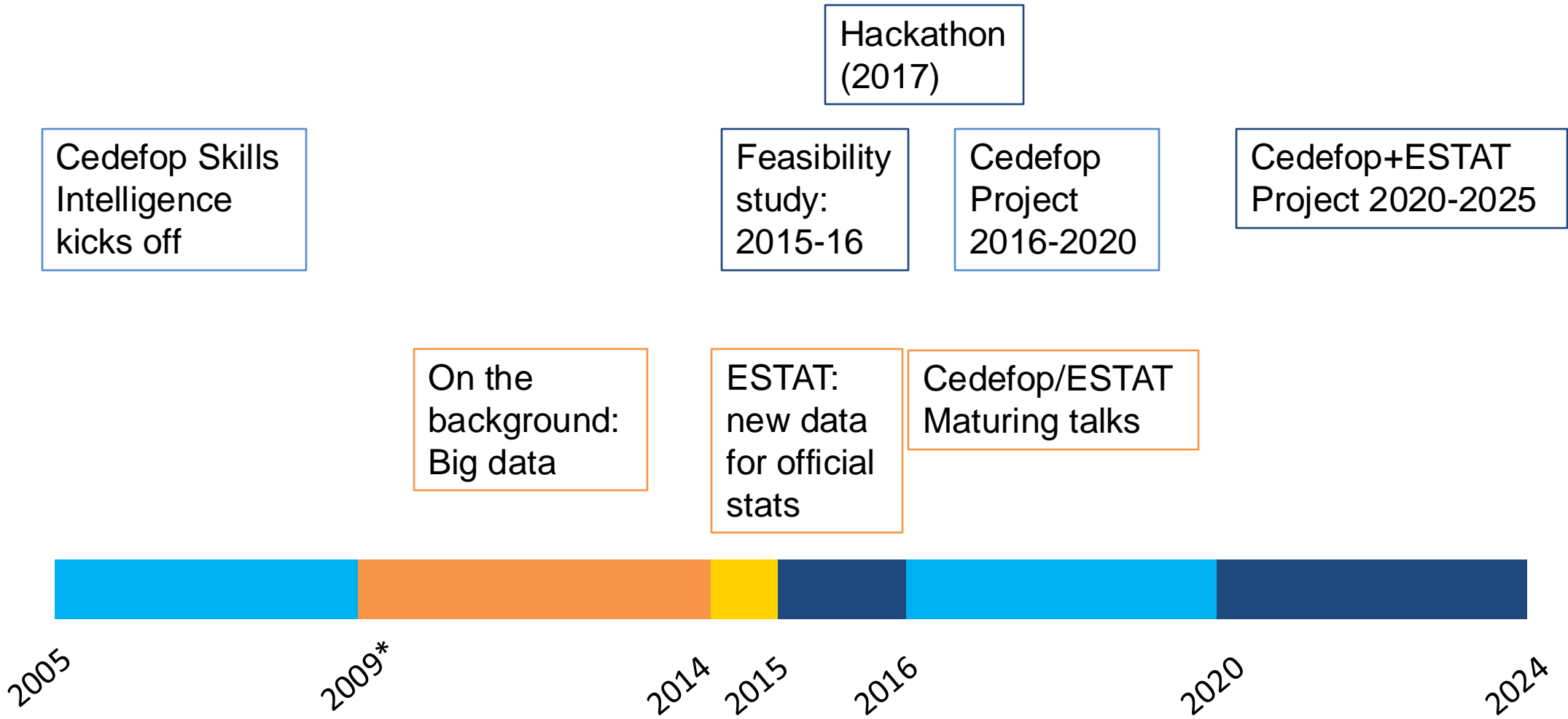
Timeline



Timeline



Timeline



Classifications, Taxonomies & Data

- **Time coverage:** 2018Q3-now
- **Countries*:** 32 countries
- **Number of OJAs*:** >260M (EU27+UK)

Classifications, Taxonomies & Data

- **Time coverage:** 2018Q3-now
- **Countries*:** 32 countries
- **Number of OJAs*:** >260M (EU27+UK)

- **Sectors:** NACE
- **Occupations:** ISCO-08
- **Skills:** ESCO
- **Regions:** NUTS

Classifications, Taxonomies & Data

- **Time coverage:** 2018Q3-now
- **Countries*:** 32 countries
- **Number of OJAs*:** >260M (EU27+UK)

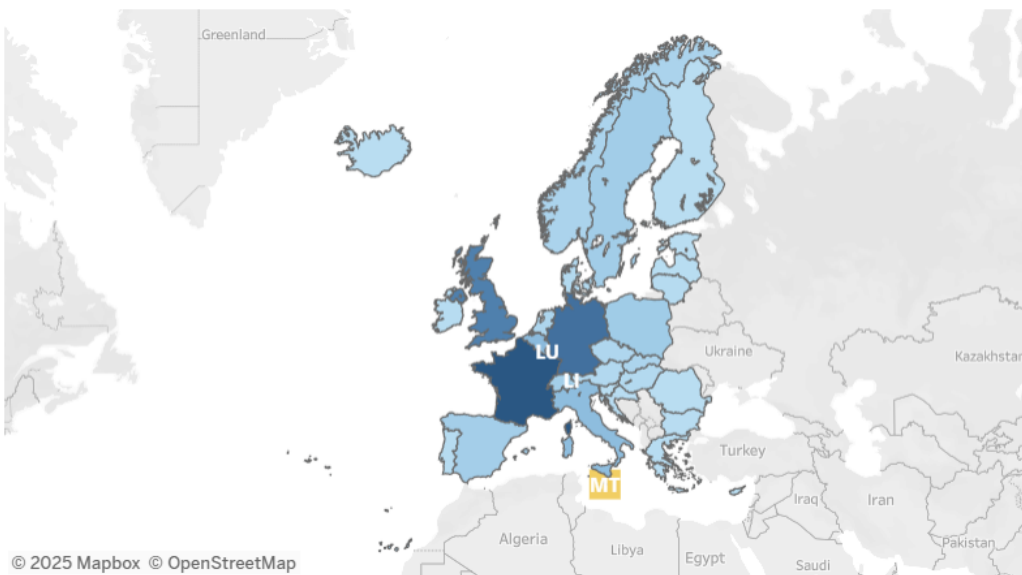
- **Sectors:** NACE
- **Occupations:** ISCO-08
- **Skills:** ESCO
- **Regions:** NUTS

Other classifications:

- Experience (#years)
- Salary
- Education
- Working time (full/part-time)
- ...

Quarterly Updates

OJAs by country

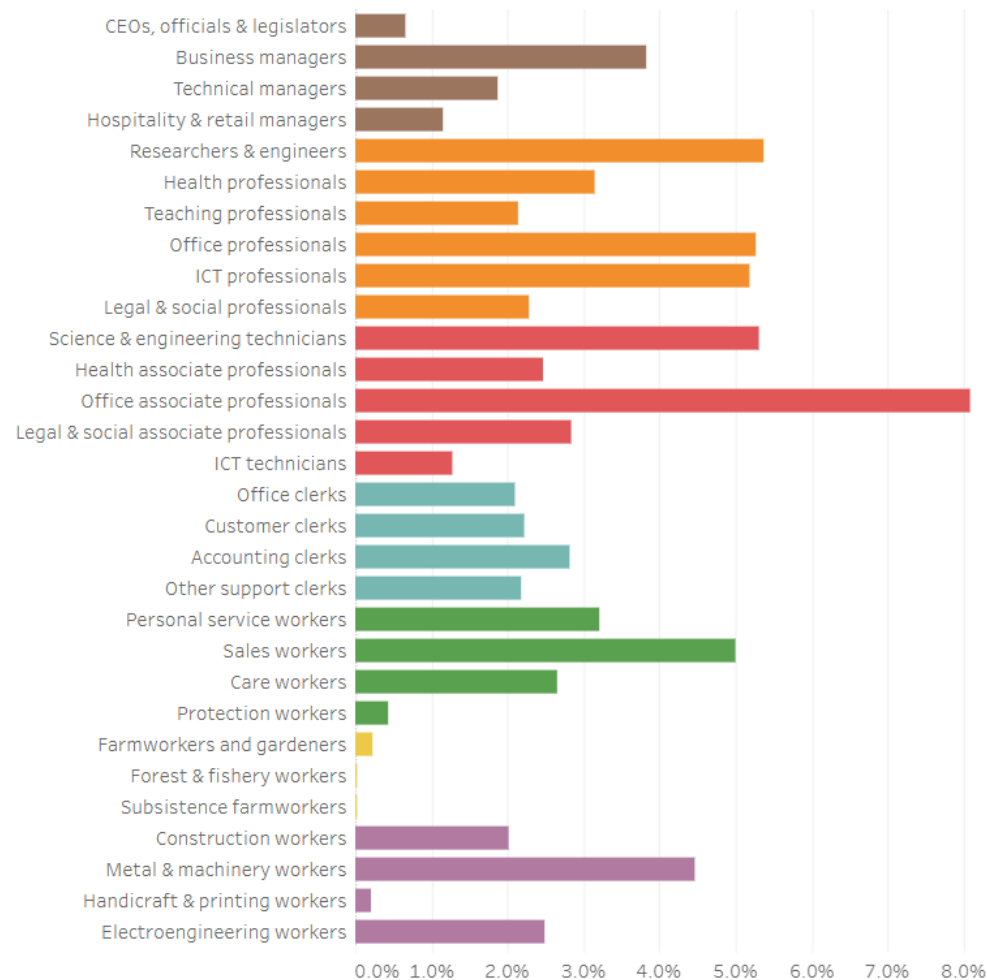


© 2025 Mapbox © OpenStreetMap

OJAs by occupation (1-digit ISCO)

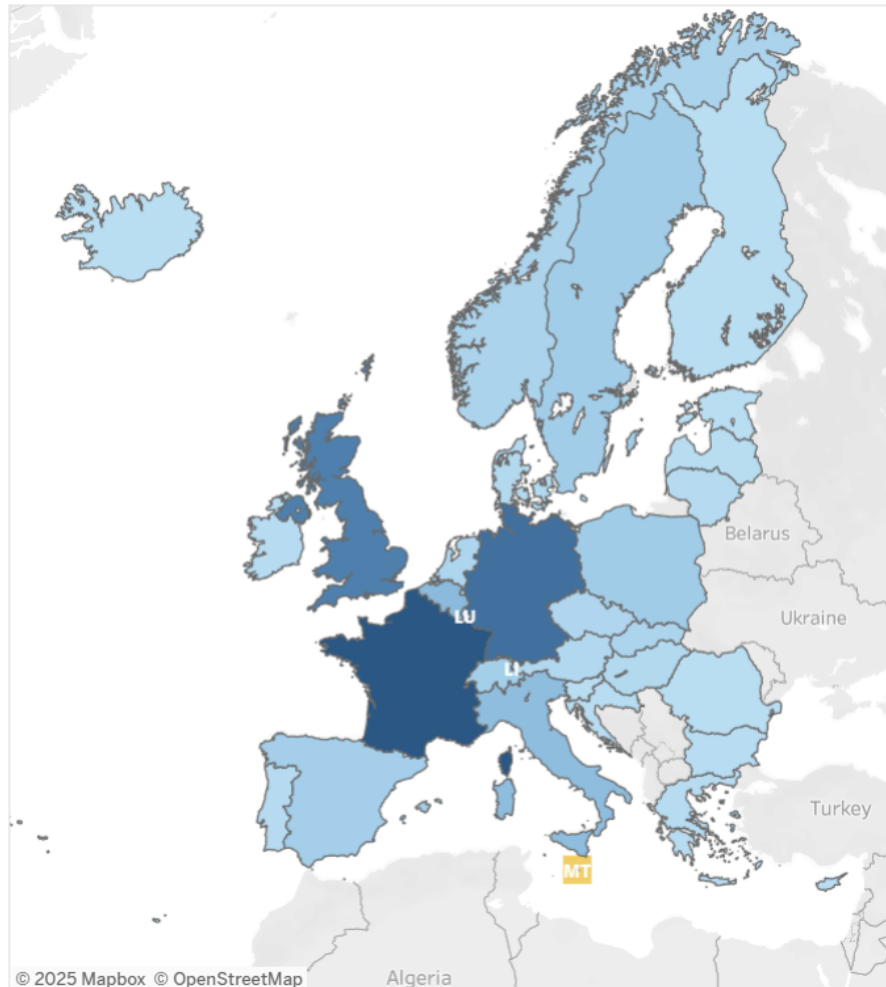


OJAs by occupation (2-digit ISCO)

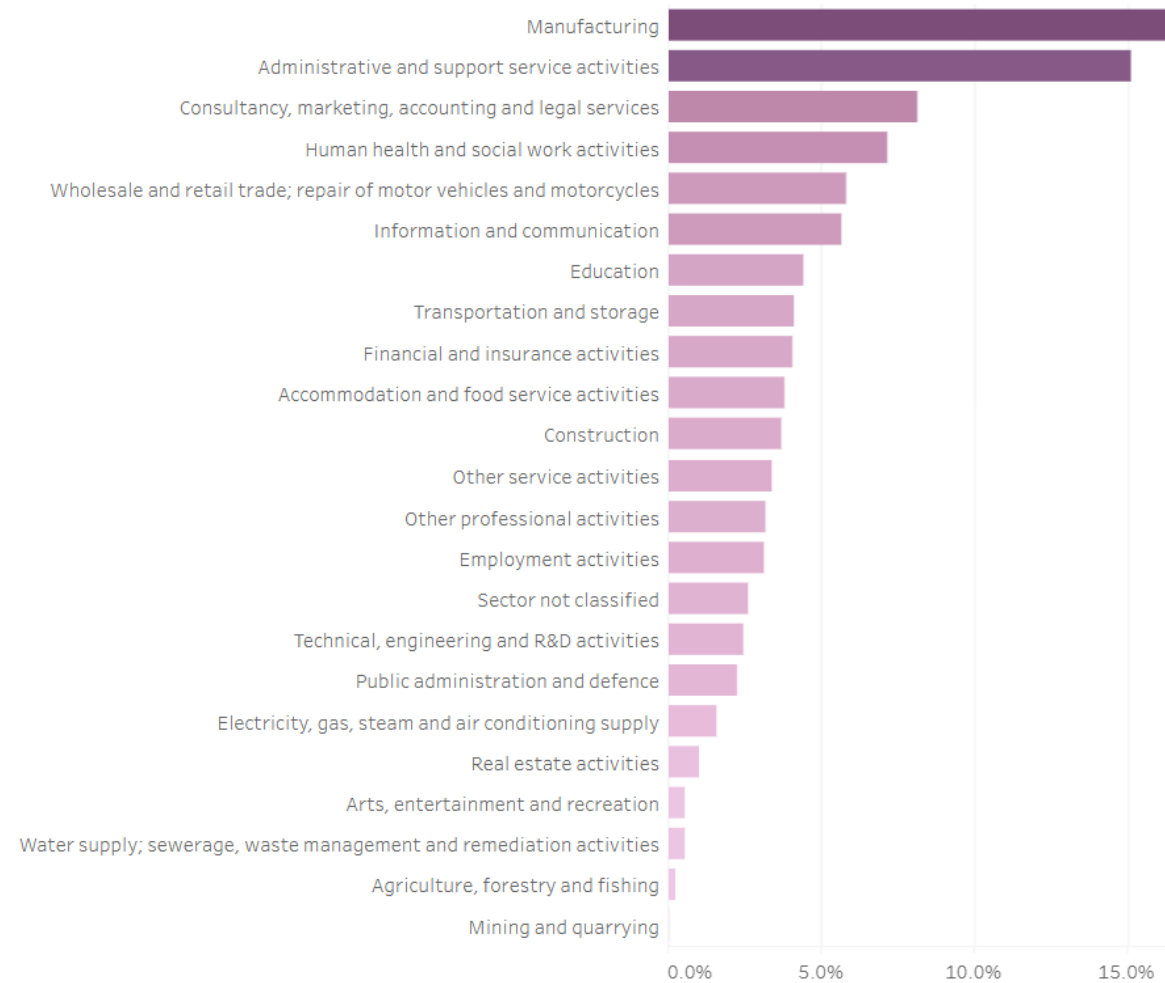


Quarterly Updates

OJAs by country

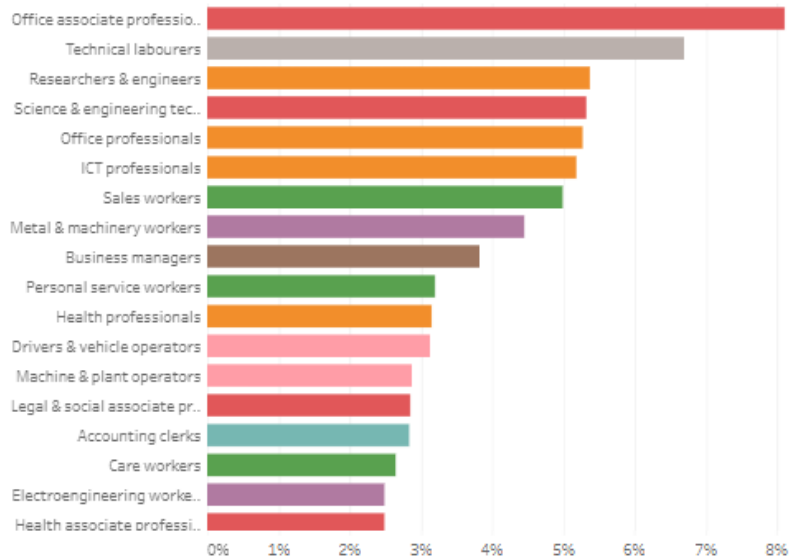


OJAs by sector

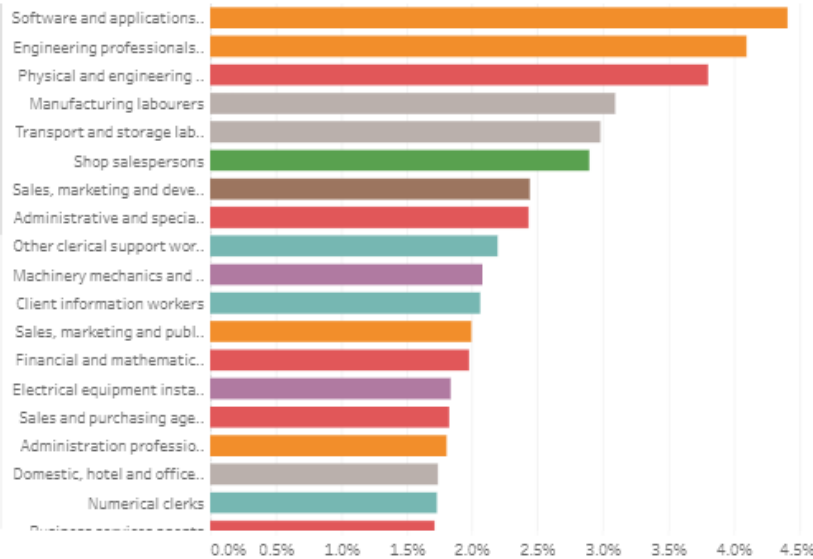


Quarterly Updates

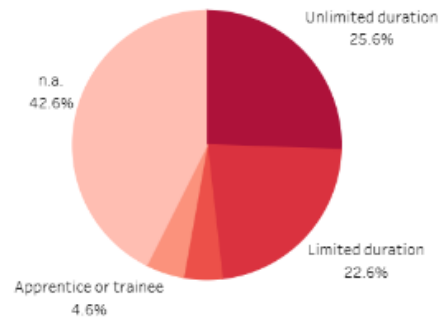
OJAs by occupation (2-digit ISCO)



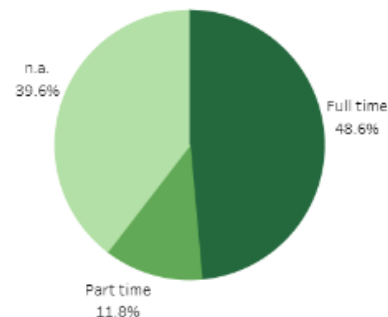
OJAs by occupation (3-digit ISCO)



OJAs by contract type



OJAs by working hours

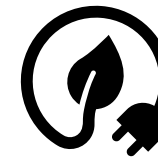


Twin Transition

(1/5)



Digital transition



Green transition

Labour market monitoring?

Twin Transition: Green Skills

(2/5)

Manual extraction: Classification of environmental protection activities (CEPA 2000); Classification of resource management activities (CReMA); IRENA Global Renewables Outlook 2020; LinkedIn 'green' skills; SGG Singapore green economy skills; Joint Research Centre's GreenComp framework; O*NET and ESCO green classifications.

Twin Transition: Green Skills

(2/5)

Manual extraction: Classification of environmental protection activities (CEPA 2000); Classification of resource management activities (CReMA); IRENA Global Renewables Outlook 2020; LinkedIn 'green' skills; SGG Singapore green economy skills; Joint Research Centre's GreenComp framework; O*NET and ESCO green classifications.

OJA-based extraction: lexical variations based on 6M OJAs in UK (English)

Twin Transition: Green Skills

(2/5)

Manual extraction: Classification of environmental protection activities (CEPA 2000); Classification of resource management activities (CReMA); IRENA Global Renewables Outlook 2020; LinkedIn 'green' skills; SGG Singapore green economy skills; Joint Research Centre's GreenComp framework; O*NET and ESCO green classifications.

OJA-based extraction: lexical variations based on 6M OJAs in UK (English)

Expert tasks:

- Validation
- Translation

Twin Transition: Green Skills

(2/5)

Manual extraction: Classification of environmental protection activities (CEPA 2000); Classification of resource management activities (CReMA); IRENA Global Renewables Outlook 2020; LinkedIn 'green' skills; SGG Singapore green economy skills; Joint Research Centre's GreenComp framework; O*NET and ESCO green classifications.

OJA-based extraction: lexical variations based on 6M OJAs in UK (English)

Expert tasks:

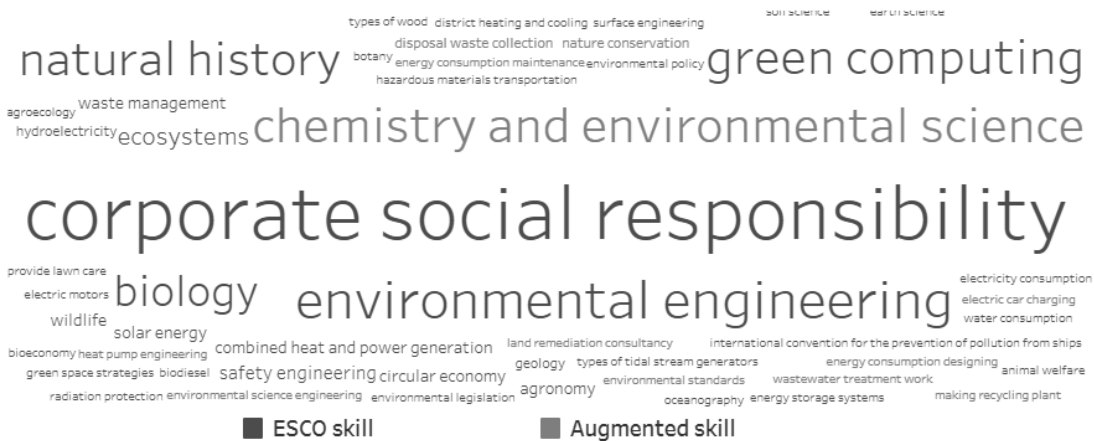
- Validation
- Translation

Map to ESCO green skills

Twin Transition: Green Skills

(3/5)

Top 50 green skills Q1-Q3 2024

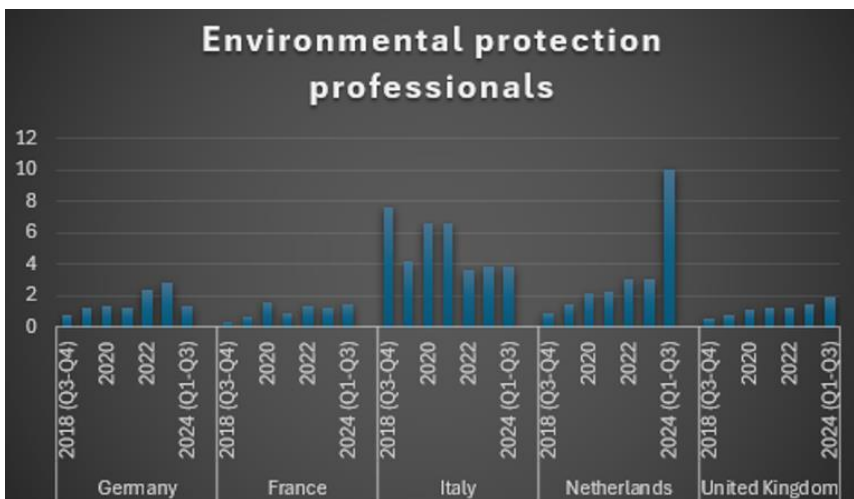
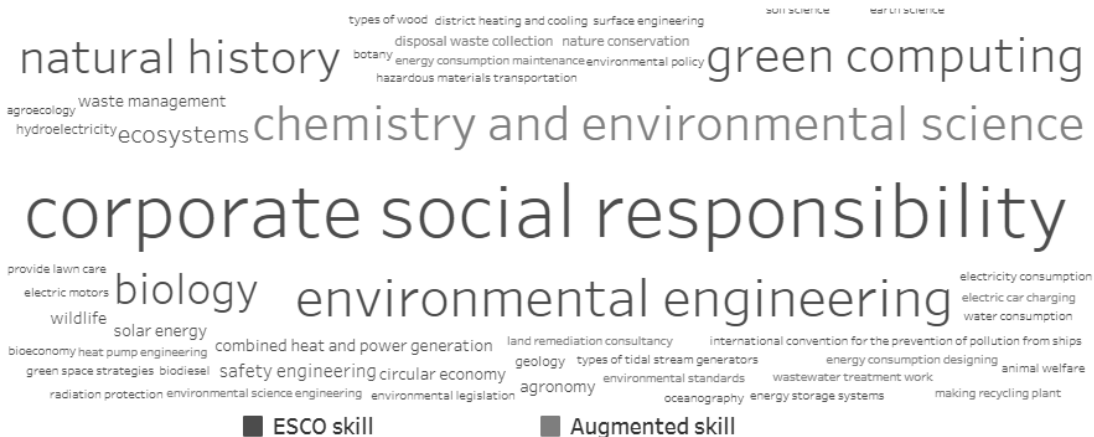


Green Pervasiveness: $(\#OJAs \text{ with green skills}) / (\#OJAs)$
Greenness: $(\#green \text{ skills}) / (\#skills)$

Twin Transition: Green Skills

(3/5)

Top 50 green skills Q1-Q3 2024

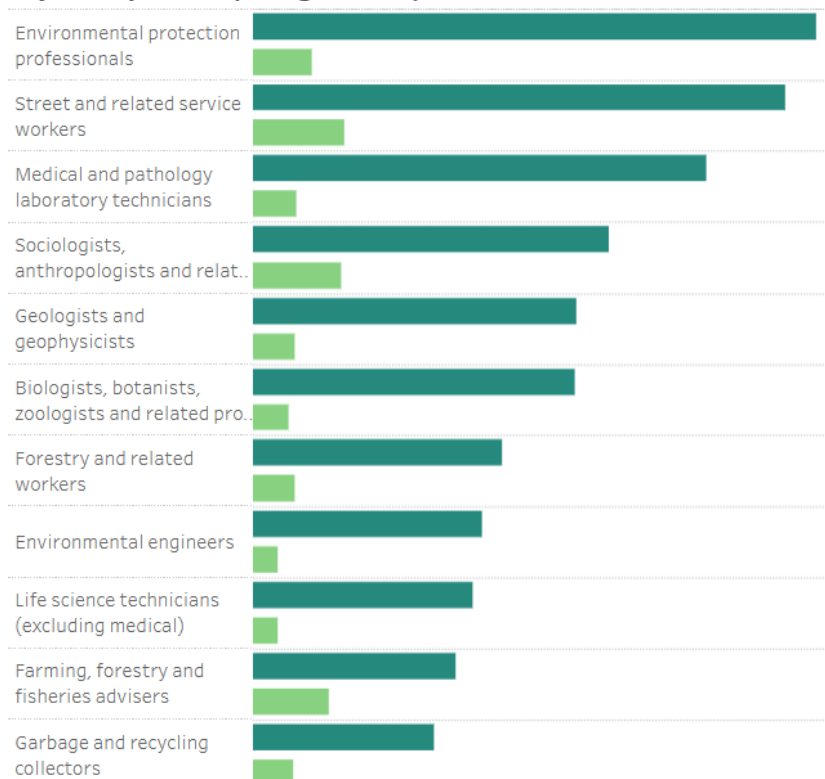


[Cedefop & UNESCO-UNEVOC \(2025\). Meeting skill needs for the green transition. Skills anticipation and VET for a greener future. Cedefop practical guide 4. Publications Office of the European Union. <http://data.europa.eu/doi/10.2801/6833866>](#)

Green Pervasiveness: (#OJAs with green skills) / (#OJAs)
Greenness: (#green skills) / (#skills)

Greenness and green pervasiveness by occupation (4-digit ISCO) Q1-Q3

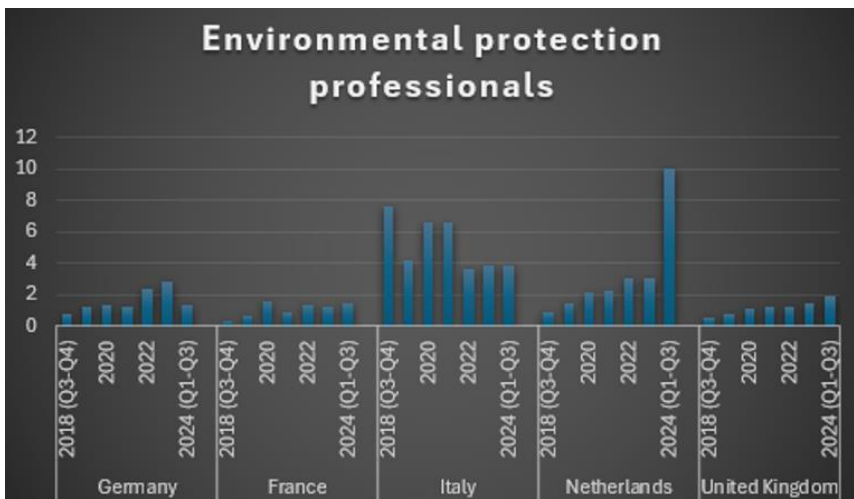
Sort by: Green pervasiveness



Twin Transition: Green Skills

(3/5)

Top 50 green skills Q1-Q3 2024

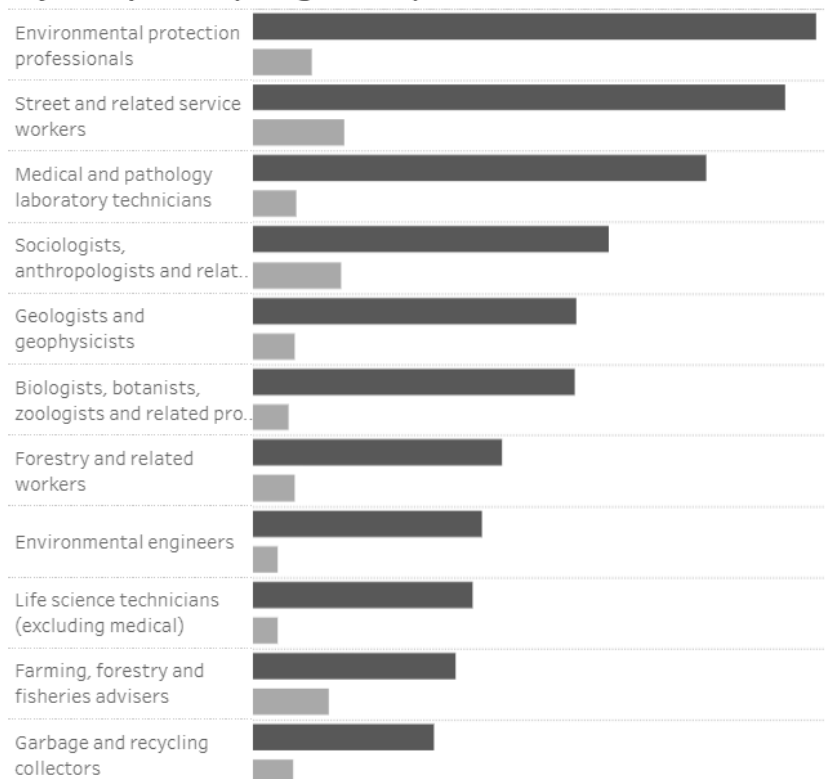


[Cedefop & UNESCO-UNEVOC. \(2025\). Meeting skill needs for the green transition. Skills anticipation and VET for a greener future. Cedefop practical guide 4. Publications Office of the European Union. <http://data.europa.eu/doi/10.2801/6833866>](#)

Green Pervasiveness: (#OJAs with green skills) / (#OJAs)
Greenness: (#green skills) / (#skills)

Greenness and green pervasiveness by occupation (4-digit ISCO) Q1-Q3

Sort by:

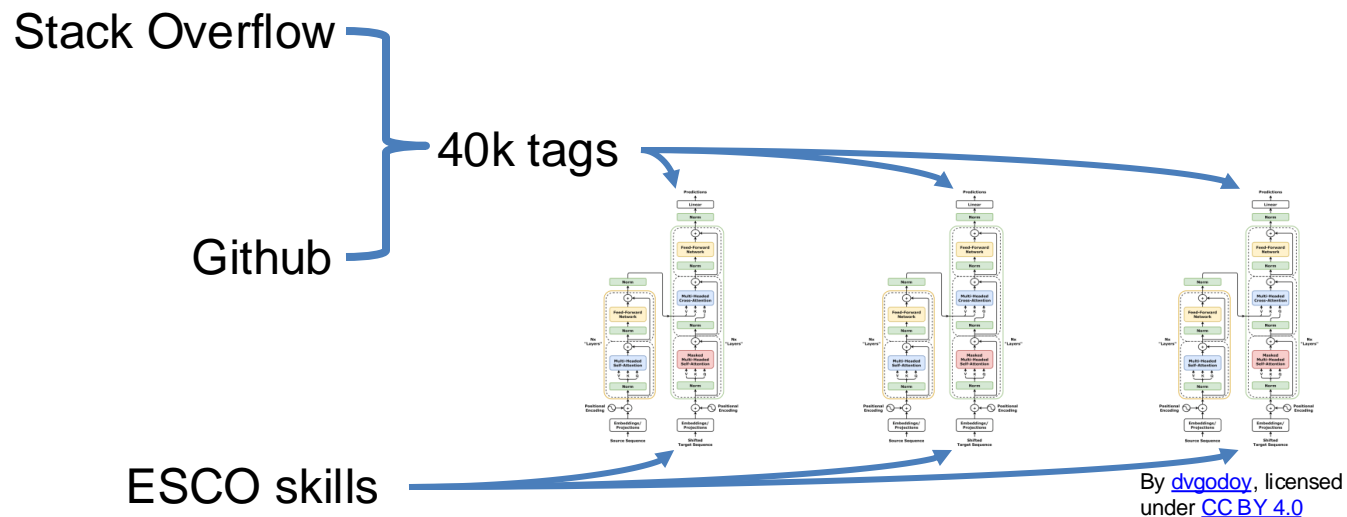


[Cedefop. \(2024\). Tracking the green transition in labour markets: using big data to identify the skills that make jobs greener. Publications Office of the European Union. Cedefop policy brief.](#)



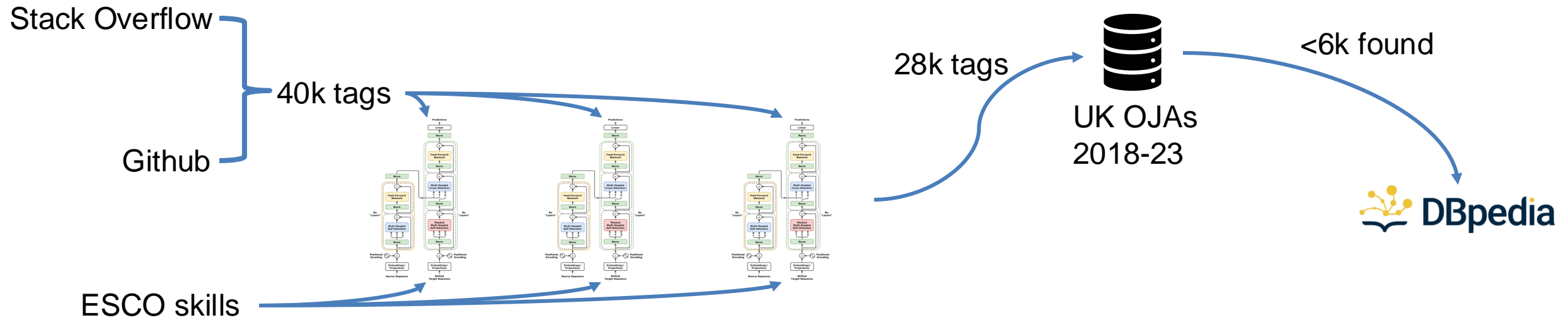
Twin Transition: Digital Skills

(4/5)



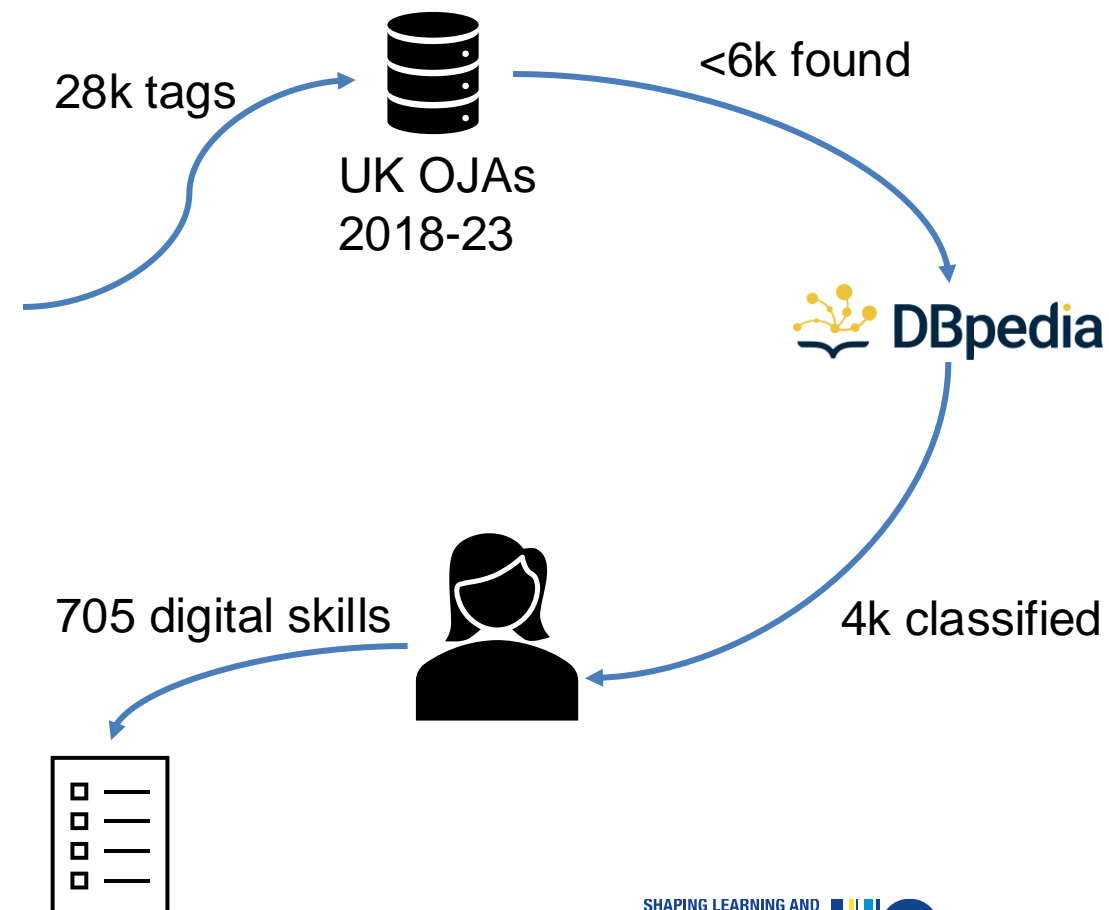
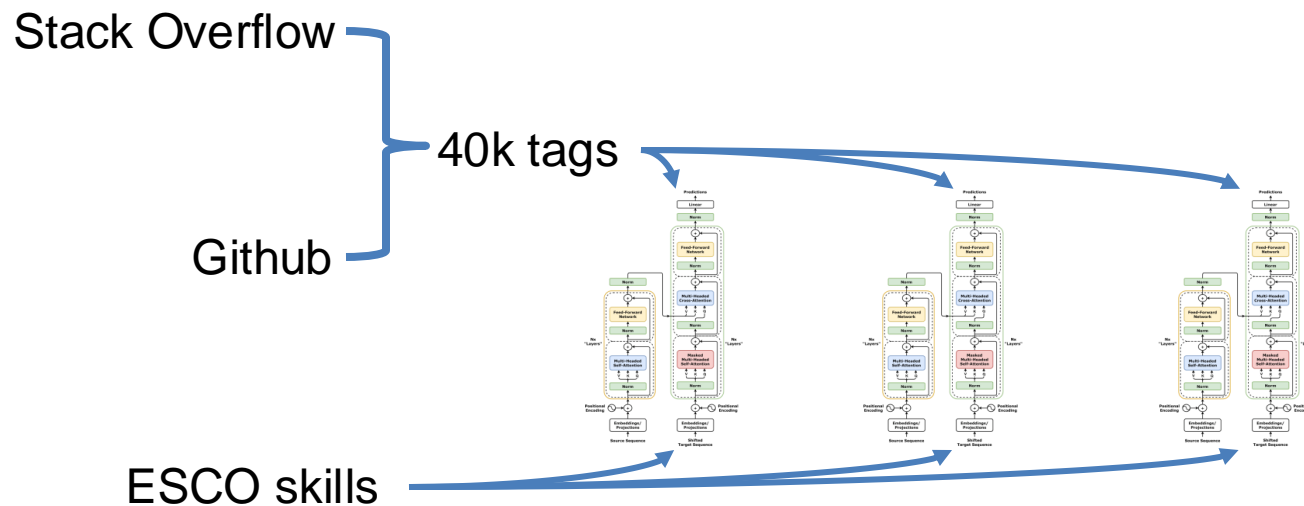
Twin Transition: Digital Skills

(4/5)



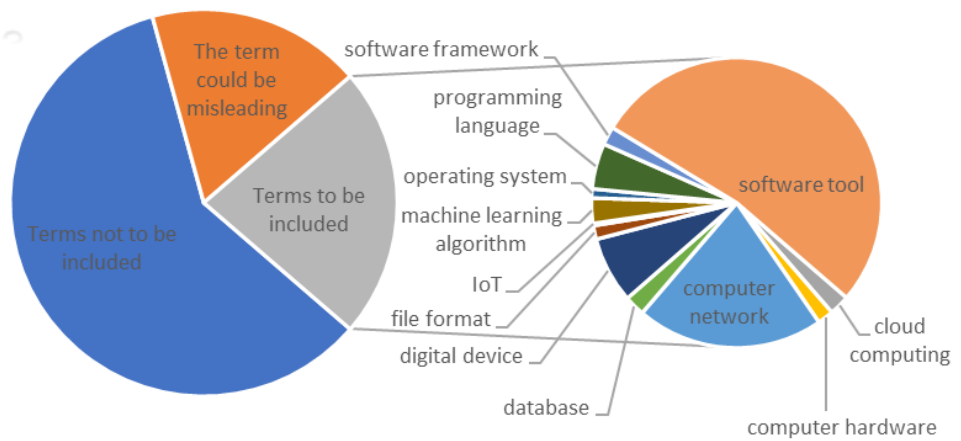
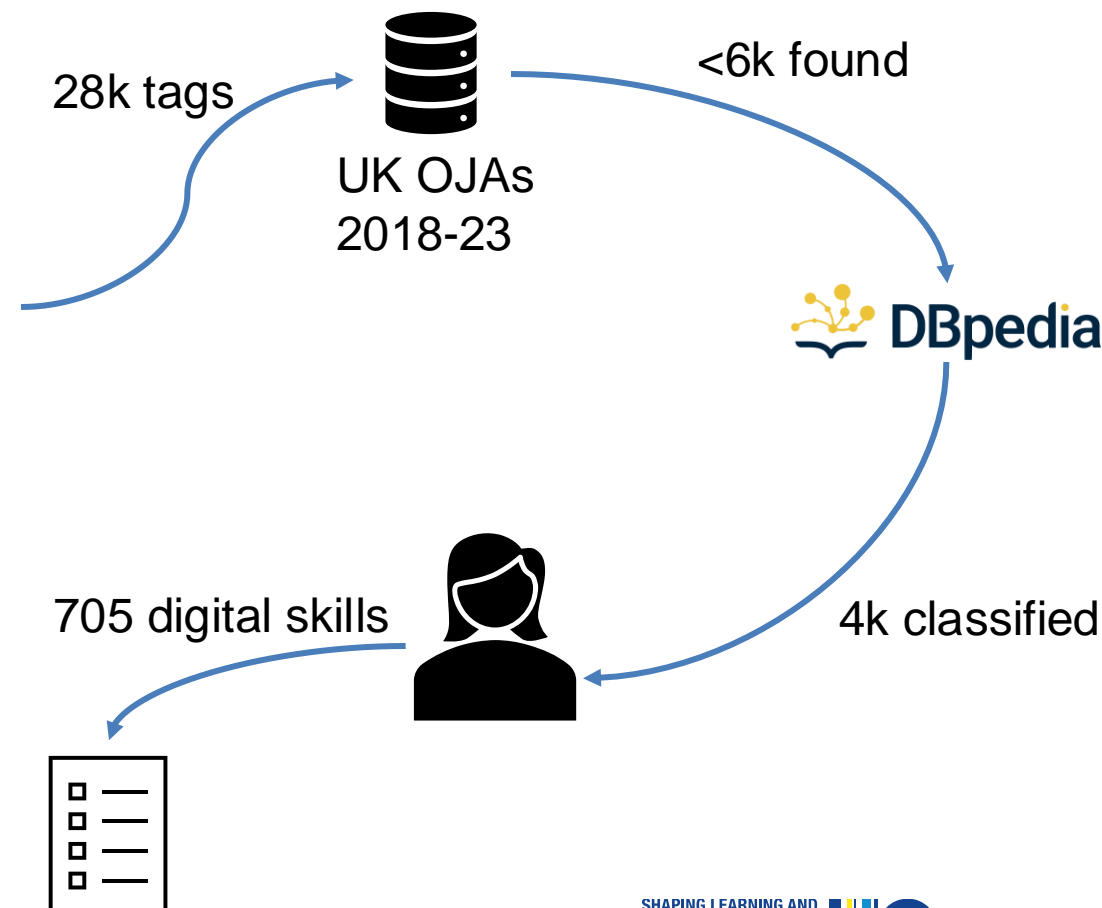
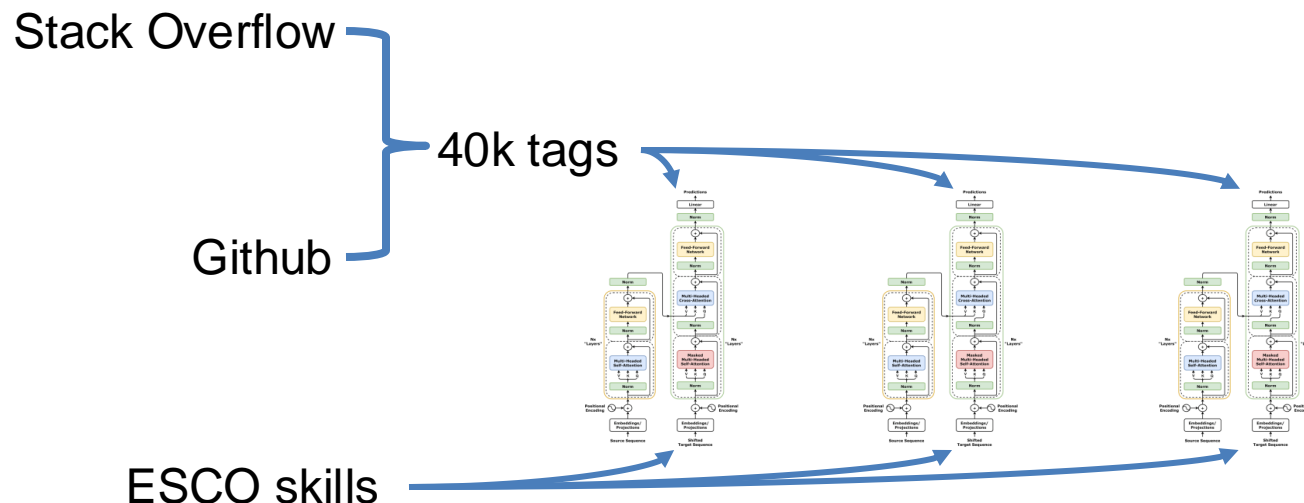
Twin Transition: Digital Skills

(4/5)



Twin Transition: Digital Skills

(4/5)



Recent progress – Overview

Large Language Models

ICT classification

New emerging skills

Region extraction

ESCO-V digit
classification

Data Science for the
Labour Market

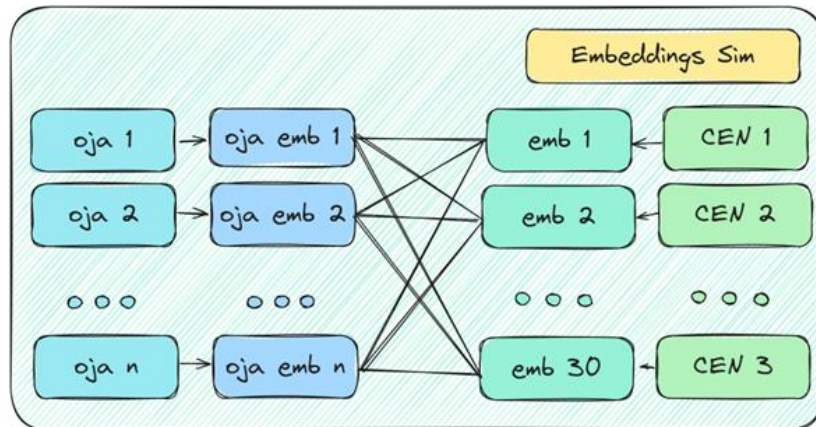
Entry level jobs

Wage variable

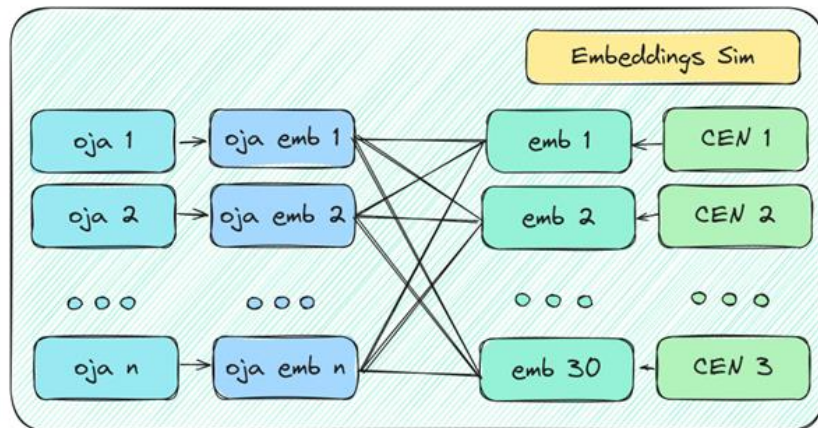
Time series indices

Skills Novelty Index

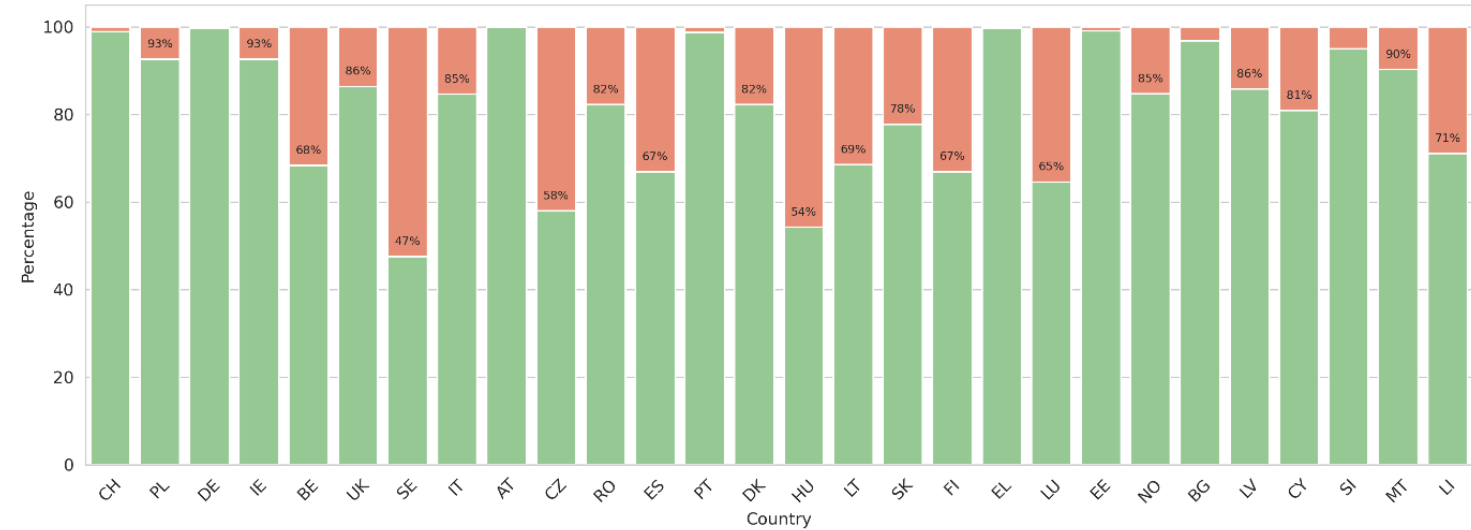
Recent progress – LLMs for ICT classification



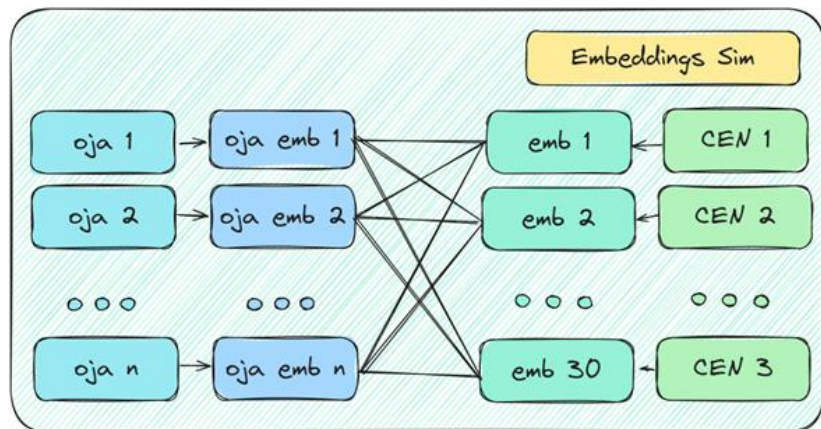
Recent progress – LLMs for ICT classification



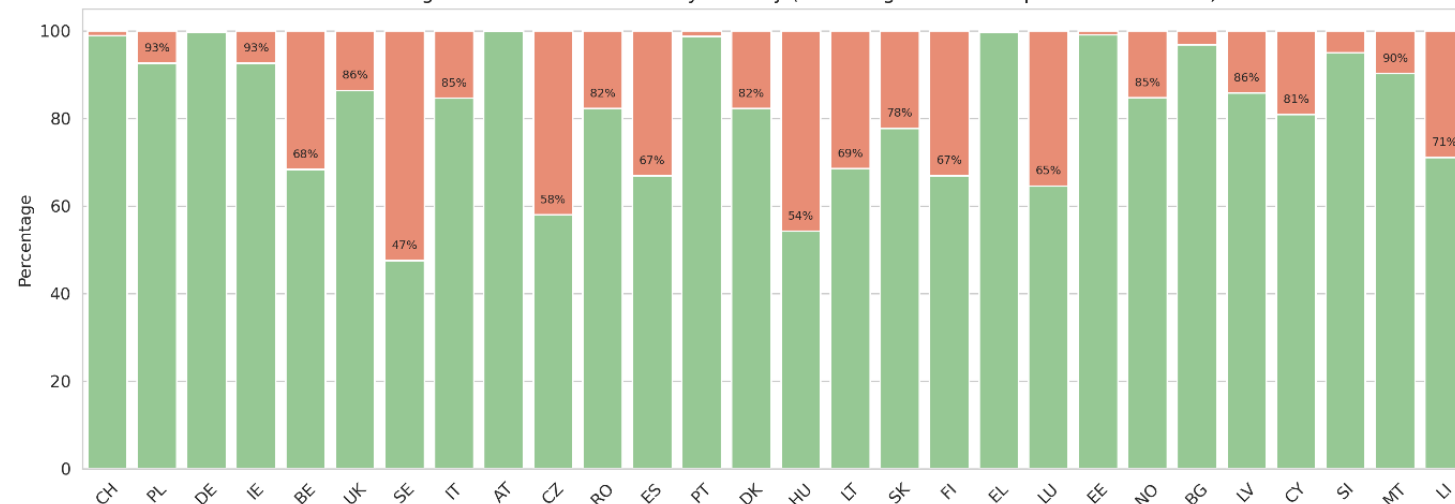
Percentage of Correct CEN Labels by Country (Excluding Not an Occupation and Not ICT)



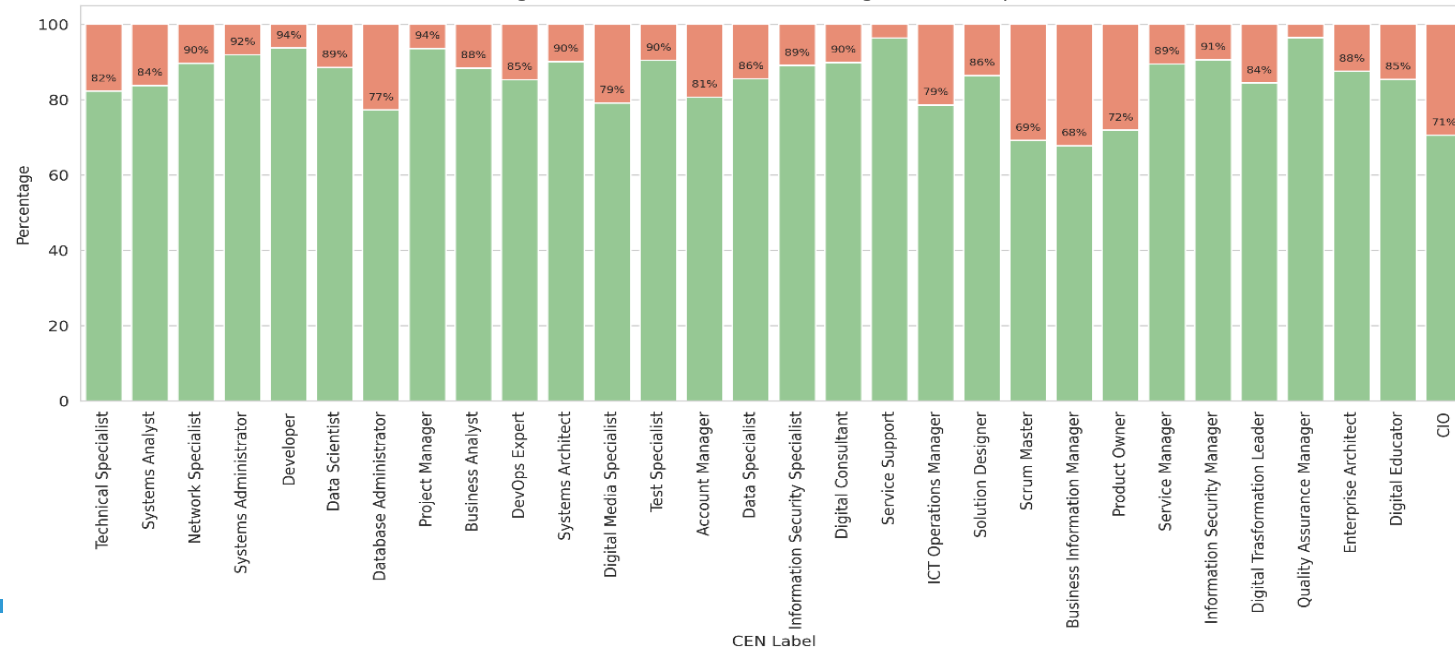
Recent progress – LLMs for ICT classification



Percentage of Correct CEN Labels by Country (Excluding Not an Occupation and Not ICT)



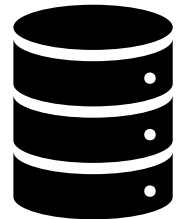
Overall Percentage of Correct CEN Labels (Excluding Not an Occupation and Not ICT)



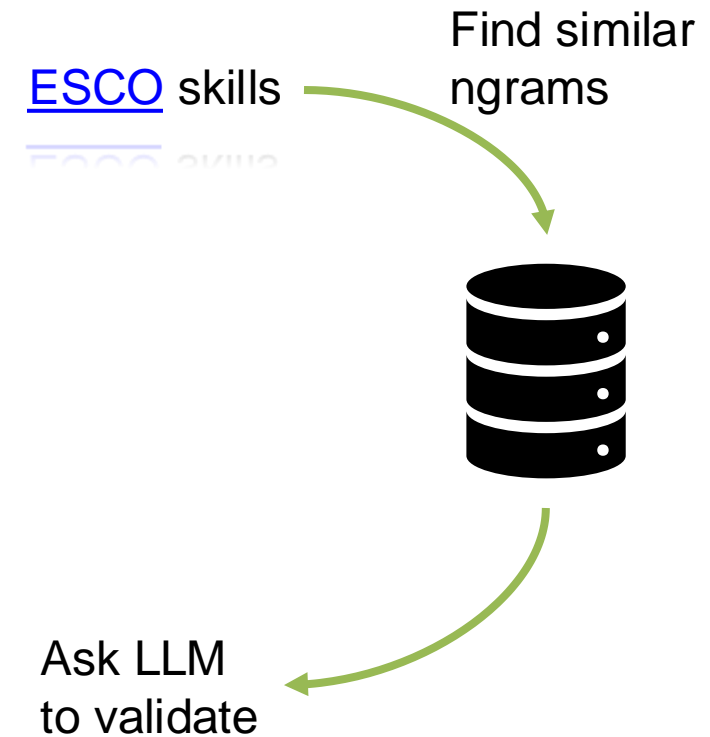
Recent progress – LLMs for new emerging skills

ESCO skills

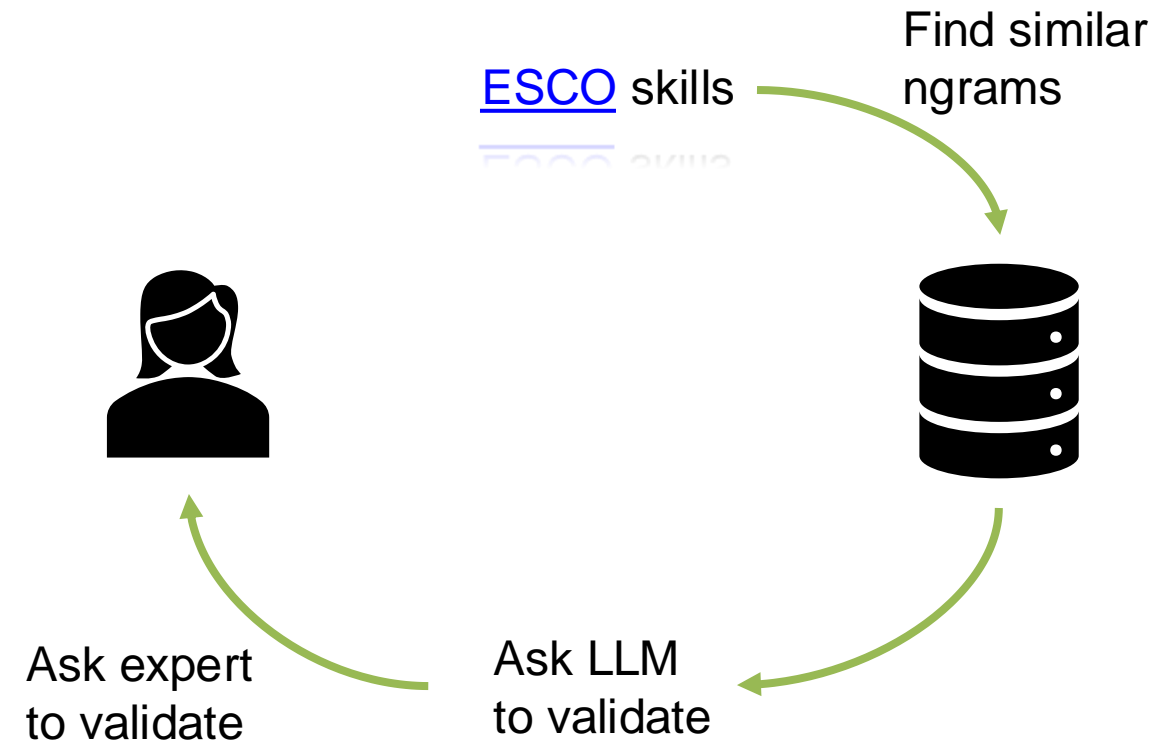
Find similar ngrams



Recent progress – LLMs for new emerging skills

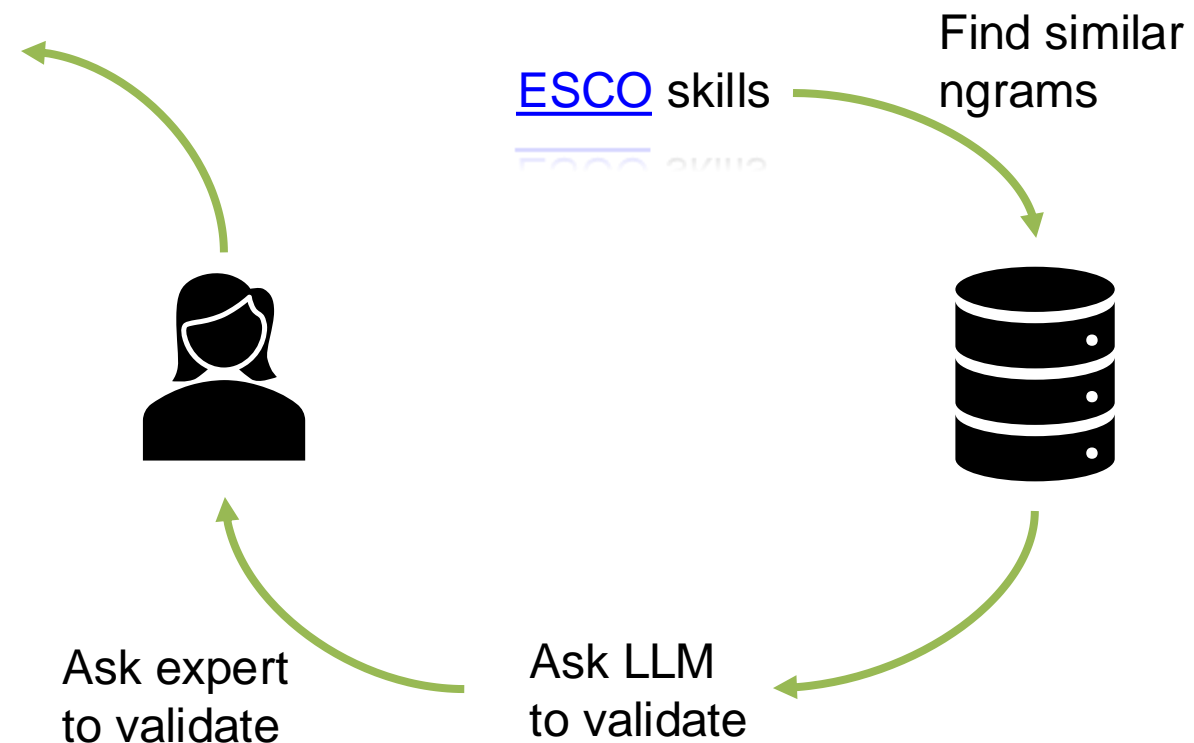
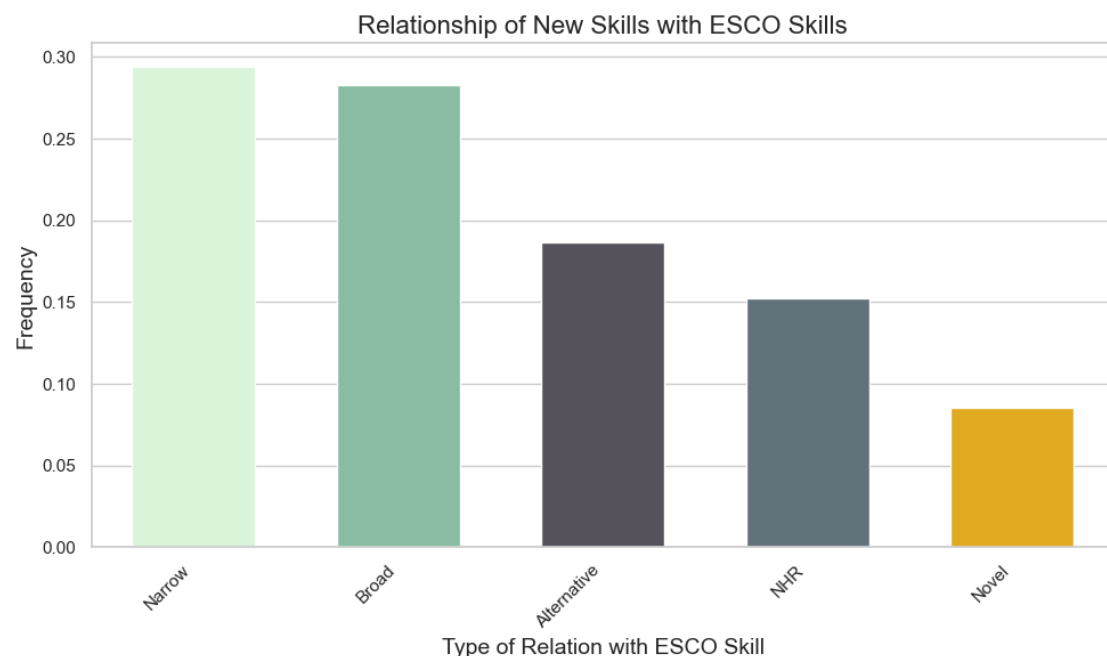


Recent progress – LLMs for new emerging skills

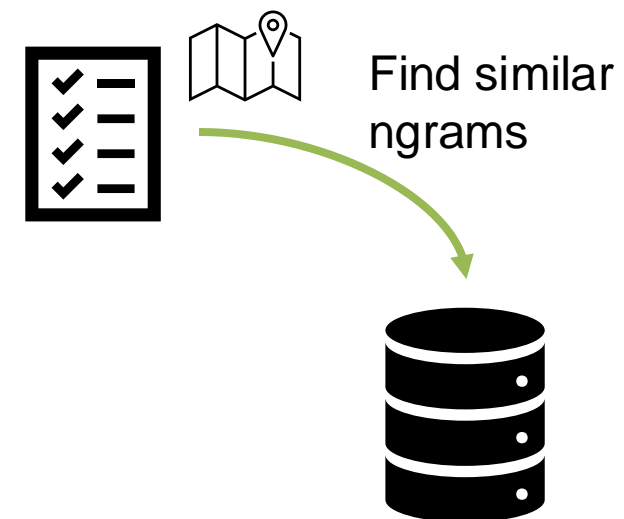


Recent progress – LLMs for new emerging skills

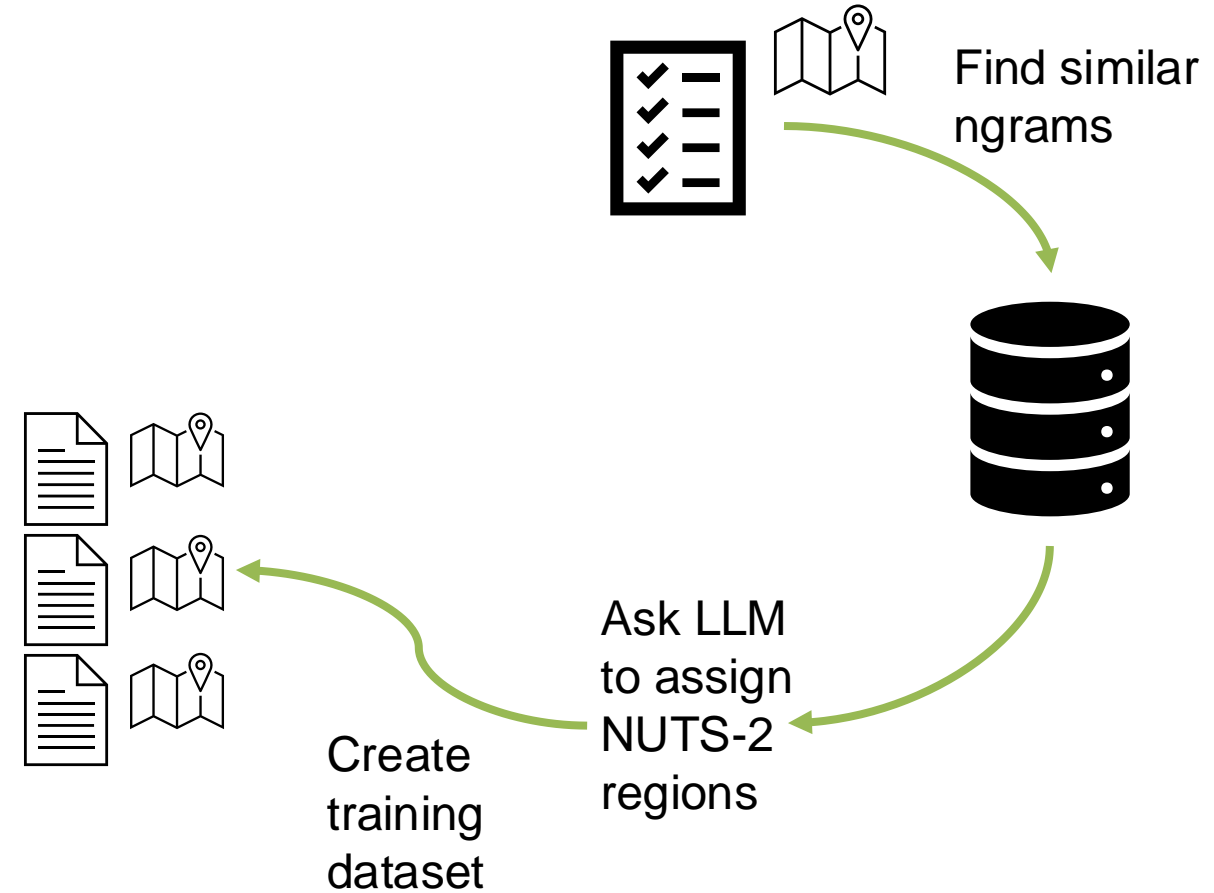
- ICEs from EU27 member states + UK
- 3,552 (from 4,941) candidates recognised as 'new skills'



Recent progress – LLMs for region classification



Recent progress – LLMs for region classification



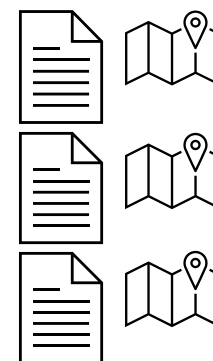
Recent progress – LLMs for region classification

- 93% accuracy

Koutsikakis et al. GREEK-BERT: The Greeks Visiting Sesame Street. In Hellenic Conference on Artificial Intelligence. 2020.



Fine-tuning



Create training dataset

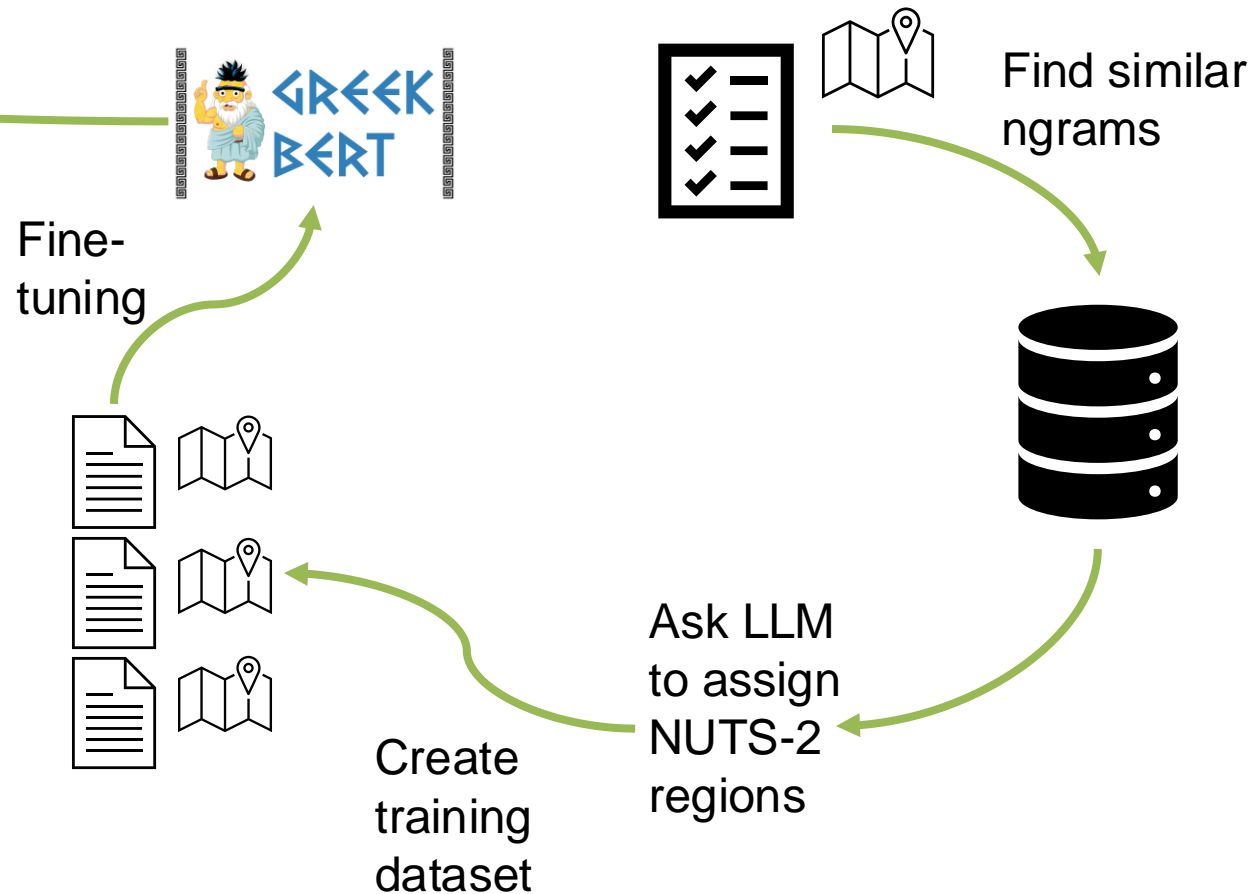
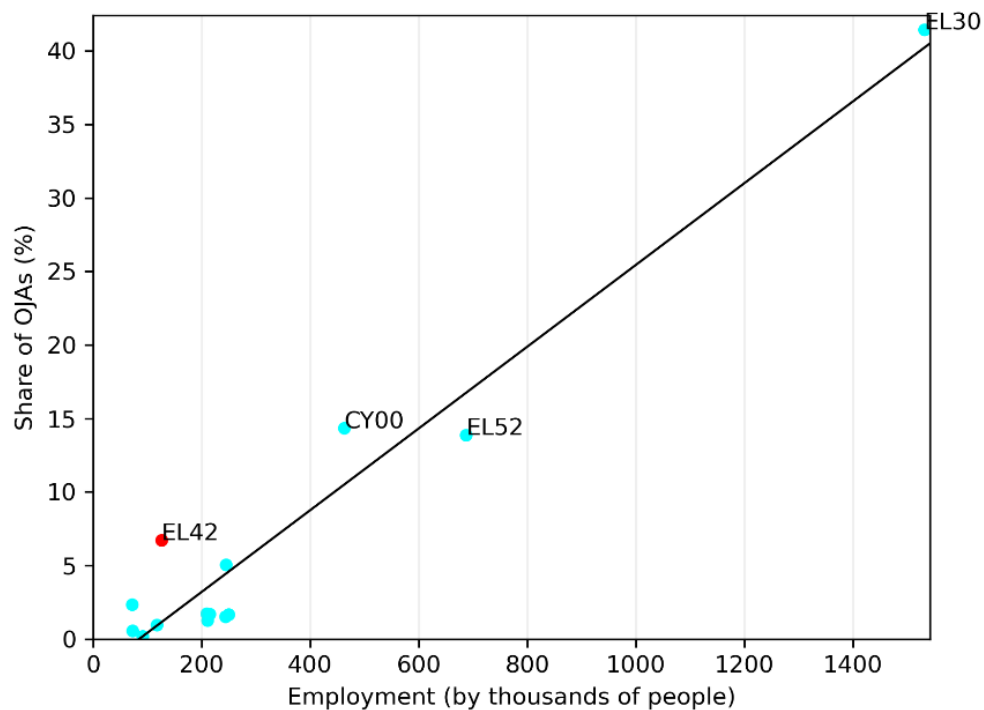
Ask LLM to assign NUTS-2 regions

Find similar ngrams



Recent progress – LLMs for region classification

- 93% accuracy



Recent progress – Data Science for the Labour Market

Recent progress – Data Science for the Labour Market

- **Entry level jobs:** new labour market entrants
 - **No experience** (114 terms, 11 identified)
 - **Apprenticeship** (90 terms, 36 identified)
 - **Traineeship** (129 terms, 52 identified)

Recent progress – Data Science for the Labour Market

- **Entry level jobs:** new labour market entrants
 - **No experience** (114 terms, 11 identified)
 - **Apprenticeship** (90 terms, 36 identified)
 - **Traineeship** (129 terms, 52 identified)

- **Wage variable**
 - **OJA vs SES:** correlation (ISCO08-1d)
 - **Education, experience, worktime**
 - **Imputing missing wages** in OJAs

Recent progress – Data Science for the Labour Market

- **Entry level jobs:** new labour market entrants
 - **No experience** (114 terms, 11 identified)
 - **Apprenticeship** (90 terms, 36 identified)
 - **Traineeship** (129 terms, 52 identified)

- **Wage variable**
 - **OJA vs SES:** correlation (ISCO08-1d)
 - **Education, experience, worktime**
 - **Imputing missing wages** in OJAs

- **OJA time-series indices**
 - **Based on growth rates** (per country)
 - **Correlation vs official statistics** (IP, JVS)
 - **Source quality assessment**

Recent progress – Data Science for the Labour Market

- **Entry level jobs:** new labour market entrants
 - **No experience** (114 terms, 11 identified)
 - **Apprenticeship** (90 terms, 36 identified)
 - **Traineeship** (129 terms, 52 identified)

- **OJA time-series indices**
 - **Based on growth rates** (per country)
 - **Correlation vs official statistics** (IP, JVS)
 - **Source quality assessment**

- **Wage variable**
 - **OJA vs SES:** correlation (ISCO08-1d)
 - **Education, experience, worktime**
 - **Imputing missing wages** in OJAs

- **Skills Novelty index**
 - **Further noise removal**
 - **Skill ‘importance’ index** (by occupation)
 - **Variation of skills** over time

Recent progress – Data Science for the Labour Market

- **Entry level jobs:** new labour market entrants
 - **No experience** (114 terms, 11 identified)
 - **Apprenticeship** (90 terms, 36 identified)
 - **Traineeship** (129 terms, 52 identified)

- **Wage variable**
 - **OJA vs SES:** correlation (ISCO08-1d)
 - **Education, experience, worktime**
 - **Imputing missing wages** in OJAs

- **OJA time-series indices**
 - **Based on growth rates** (per country)
 - **Correlation vs official statistics** (IP, JVS)
 - **Source quality assessment**

- **Skills Novelty index**
 - **Further noise removal**
 - **Skill ‘importance’ index** (by occupation)
 - **Variation of skills** over time

In parallel: crawling/scraping/APIs, preprocessing, deduplication, classification, database, dashboards, support, analysis...

The way forward – Ideas

- Additional information sources



The way forward – Ideas

- Additional information sources
- Gold standard datasets:
 - State-of-the-art LLMs for labelling or creating artificial datasets

The way forward – Ideas

- Additional information sources
- Gold standard datasets:
 - State-of-the-art LLMs for labelling or creating artificial datasets
- Time series & NLP-inspired longitudinal tasks, adjusted to the labour market:
 - Change point detection
 - Anomaly detection
 - First story detection
 - Semantic change detection



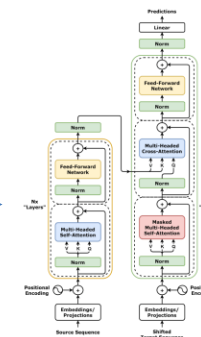
The way forward

- Regular DPS quarterly updates



The way forward

- Regular DPS quarterly updates
- Quality improvements



By [dvgodoy](#), licensed under [CC BY 4.0](#)

The way forward

- Regular DPS quarterly updates
- Quality improvements
- New skill taxonomies in line with EC priorities
 - Examples: healthcare, STEM, AI

The future of European competitiveness: Report by Mario Draghi



Mario Draghi – former European Central Bank President and one of Europe's great economic minds – was tasked by the European Commission to prepare a report of his personal vision on the future of European competitiveness.

The report looks at the challenges faced by the industry and companies in the Single Market. It outlines how Europe will no longer be able to rely on many of the factors that have supported growth in the past and lays out a clear diagnosis and provides concrete recommendations to put Europe onto a different trajectory.

[Download the report >](#)

https://commission.europa.eu/topics/eu-competitiveness/draghi-report_en

Union of Skills

Investing in people for a competitive European Union

The European Union needs skilled people to respond to new challenges and stay competitive. But as job requirements change, many workers struggle to keep up and businesses cannot find the right talent. These skills and labour gaps are hindering European competitiveness. That is why the European Commission is introducing the Union of Skills, a plan to improve high quality education, training, and lifelong learning.

It aims to

- deliver higher levels of basic and advanced skills
- provide opportunities for people to regularly update and learn new skills
- facilitate recruitment by businesses across the EU
- attract, develop and retain top talent in Europe

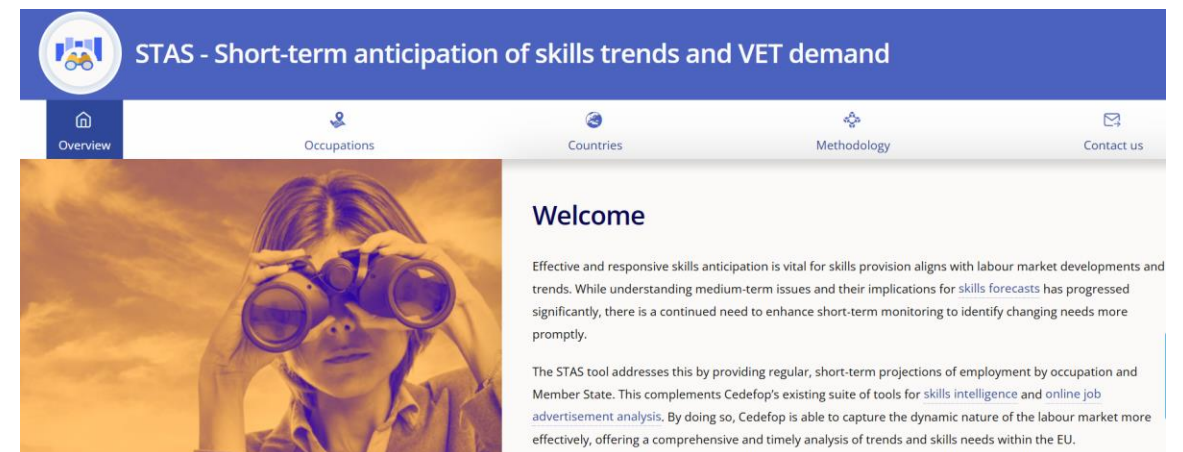


https://commission.europa.eu/topics/eu-competitiveness/union-skills_en

The way forward

- Regular DPS quarterly updates
- Quality improvements
- New skill taxonomies in line with EC priorities
 - Examples: healthcare, STEM, AI
- Integration to other tools

<https://www.cedefop.europa.eu/en/tools/stas>



The screenshot shows the STAS website interface. At the top, there is a blue header with the text "STAS - Short-term anticipation of skills trends and VET demand" and a logo. Below the header is a navigation bar with five tabs: "Overview" (selected), "Occupations", "Countries", "Methodology", and "Contact us". The main content area features a large image of a person looking through binoculars against a sunset background. To the right of the image, there is a "Welcome" section with a paragraph of text: "Effective and responsive skills anticipation is vital for skills provision aligns with labour market developments and trends. While understanding medium-term issues and their implications for skills forecasts has progressed significantly, there is a continued need to enhance short-term monitoring to identify changing needs more promptly." Below this, another paragraph states: "The STAS tool addresses this by providing regular, short-term projections of employment by occupation and Member State. This complements Cedefop's existing suite of tools for skills intelligence and online job advertisement analysis. By doing so, Cedefop is able to capture the dynamic nature of the labour market more effectively, offering a comprehensive and timely analysis of trends and skills needs within the EU."

The way forward

- Regular DPS quarterly updates
- Quality improvements
- New skill taxonomies in line with EC priorities
 - Examples: healthcare, STEM, AI
- Integration to other tools
- Resources & collaboration

Conclusion

- Delivering Skills Intelligence based on OJAs: a decade of rapid progress

Conclusion

- Delivering Skills Intelligence based on OJAs: a decade of rapid progress
- From scraping/crawling to data insights
- Cautionary note on bias

Conclusion

- Delivering Skills Intelligence based on OJAs: a decade of rapid progress
- From scraping/crawling to data insights
- Cautionary note on bias
- Moving forward: (a) production; (b) quality; (c) tasks; (d) collaboration...

Conclusion

- Delivering Skills Intelligence based on OJAs: a decade of rapid progress
- From scraping/crawling to data insights
- Cautionary note on bias
- Moving forward: (a) production; (b) quality; (c) tasks; (d) collaboration...

...starting from now!

Using data from the web to shape next-generation labour market and skills intelligence

Thank you

www.cedefop.europa.eu

Follow us on social media



@adtsakal

Additional Links:

[Skills OVATE](#)

[Cedefop Skills Intelligence](#)



CEDEFOP

European Centre
for the Development
of Vocational Training