

Big Data for Labour Market Intelligence

Capacity development programme 2022

Module 2: Dissemination and analysis

Session 6


Role of ESCO and O*Net classifications in OJVs analysis.

Data Scientist

Data Science Analyst (Remote) ⋮

Yelp 3.4 ★

Remote

 **Full-time**

- A knack for communicating quantitative results to a broad audience through writing and compelling **data** visualization.

Posted 30+ days ago · More...

Machine Learning Research Engineer ⋮

Evolution Artificial Intelligence



Remote

- You'll be putting state of the art deep learning technology into production use.
- Designing, developing and rigorous testing of machine learning models.

Data Science Manager ⋮

Consortia

Remote


 **£120,000 a year**  **Permanent**

- As **Data Science Manager**, you will lead and grow a world-class

Applied Scientist (Remote)

Yelp 3.4 ★

Remote

 **Full-time**

- Experience with **data** analysis/statistical software and packages (pandas/statsmodels/sklearn within Python, R, etc.).

Posted 30+ days ago · More...

Biostatistician - Global Biostatistics ⋮

IQVIA 3.8 ★

Remote

 **Full-time**


- Perform protocol development, sample size calculation, protocol and CRF review, and **data** management on database design and data.

Posted 30+ days ago · More...

Senior Analyst, BI & Analytics ⋮

WorldRemit

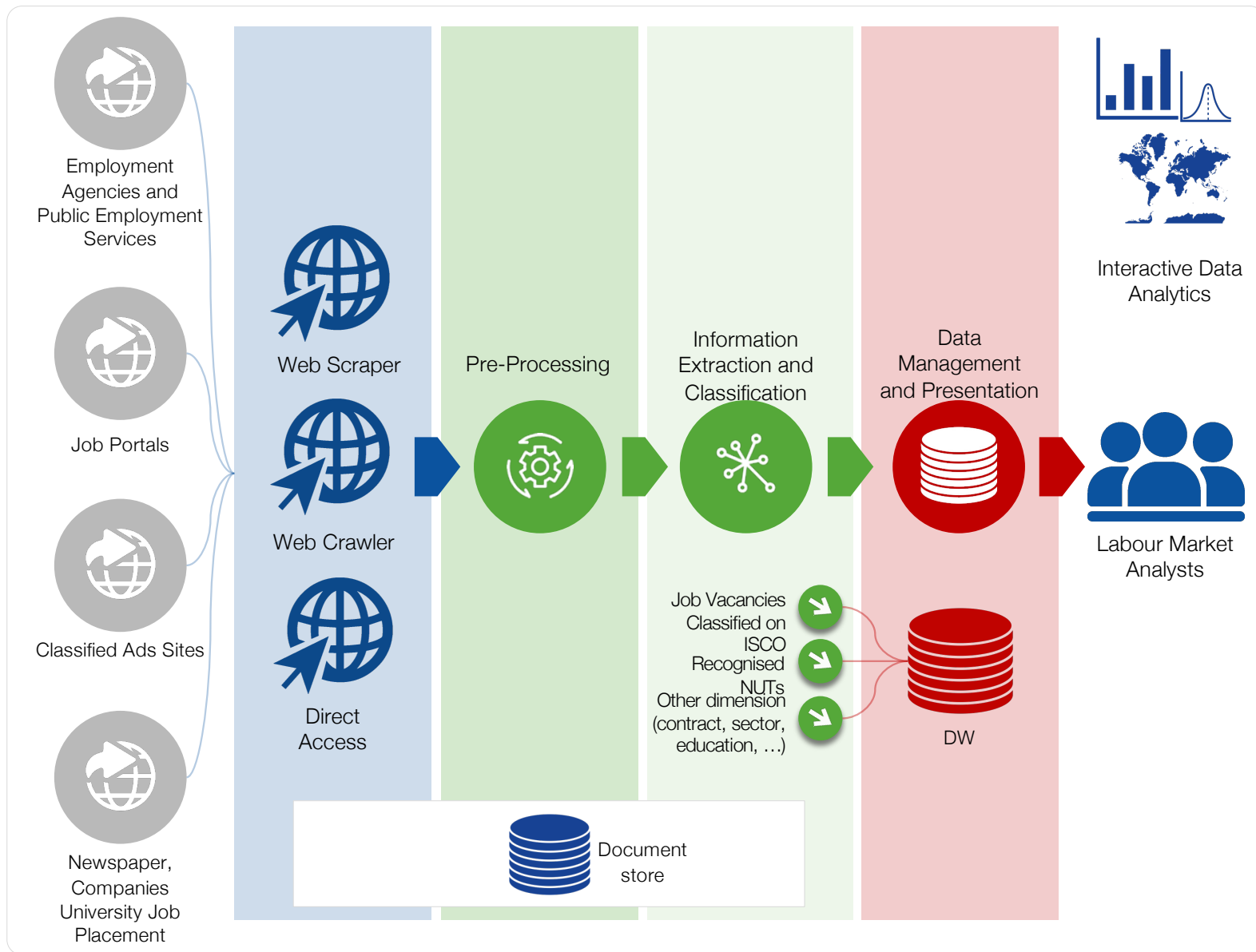
Remote

 **Full-time**

- Experience in an analytics or **data** science role.
- Strong SQL skills, able to handle large, complex **data** structures from multiple sources.

Posted 30+ days ago · More...

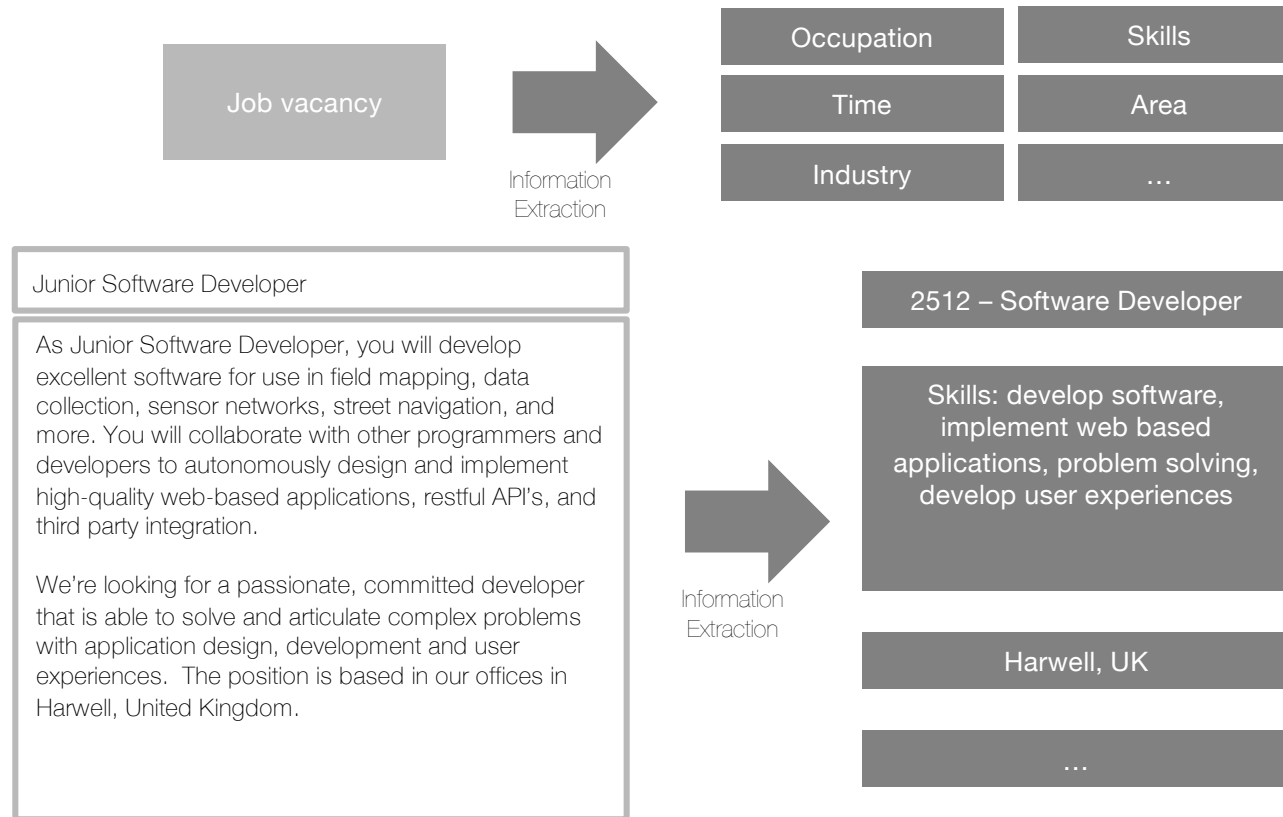
data center engineer artificial intelligence specialist
artificial intelligence developer
big data architect database administrator deep learning engineer dba sql server
project manager artificial intelligence artificial intelligence engineer
data protection officer analista energetico
privacy protection consultant machine learning engineer
business intelligence & dwh specialist
data scientist statistico business intelligence developer analista di mercato
data quality specialist associate scientist nlp expert data analyst
computer vision developer dba oracle data compliance consultant
big data developer computer vision engineer



Data classification


- **Goal:**
 - Extract and structure information from data, to be provided to the presentation layer
- **Challenges:**
 - Handle massive amount of heterogeneous data written in different languages
- **Approach:**
 - Develop an adaptable framework, language dependent, tailored on different information features. Some relevant challenges:
 - **Occupation** feature classification: combined methods such as Machine Learning, Topic Modeling and Unsupervised Learning
 - **Skill** feature classification: another different combined methods, such as Text Analysis with corpus based or Knowledge based similarity
- **Features:**
 - Guarantee Explainable information extraction, logging classification methods and relevant features.

Data Classification - An example



Do we really need a
classification standard?



An aerial photograph of a dense urban landscape, likely a major city center. The image shows a variety of skyscrapers and high-rise buildings, some with glass facades reflecting the sky. A wide, multi-lane road or highway runs through the center of the image, flanked by greenery and smaller buildings. The overall scene is hazy, suggesting a clear but slightly misty day. The text is overlaid in the center of the image.

The quality of big data analysis depends on the quality of the classification system. It's like the road to follow in a big maze.

Can you imagine some big data analysis collected so far?

in EU
by Broad occupations



SKILLS PANORAMA
EUROPEAN COMMISSION

in EU
in Professionals

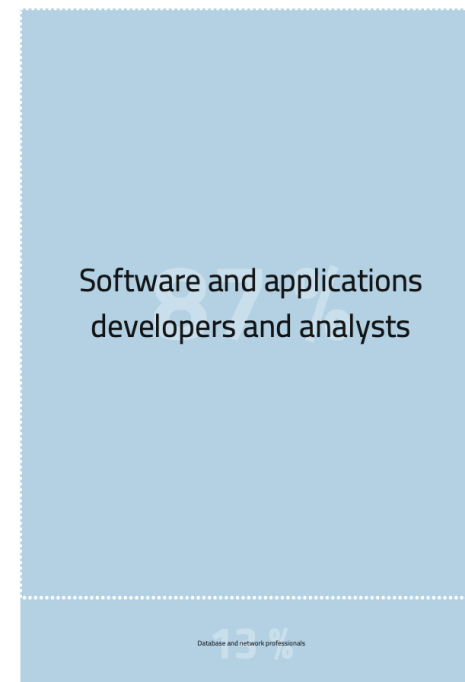
Legal & social professionals : 7.0%



SKILLS PANORAMA
EUROPEAN COMMISSION

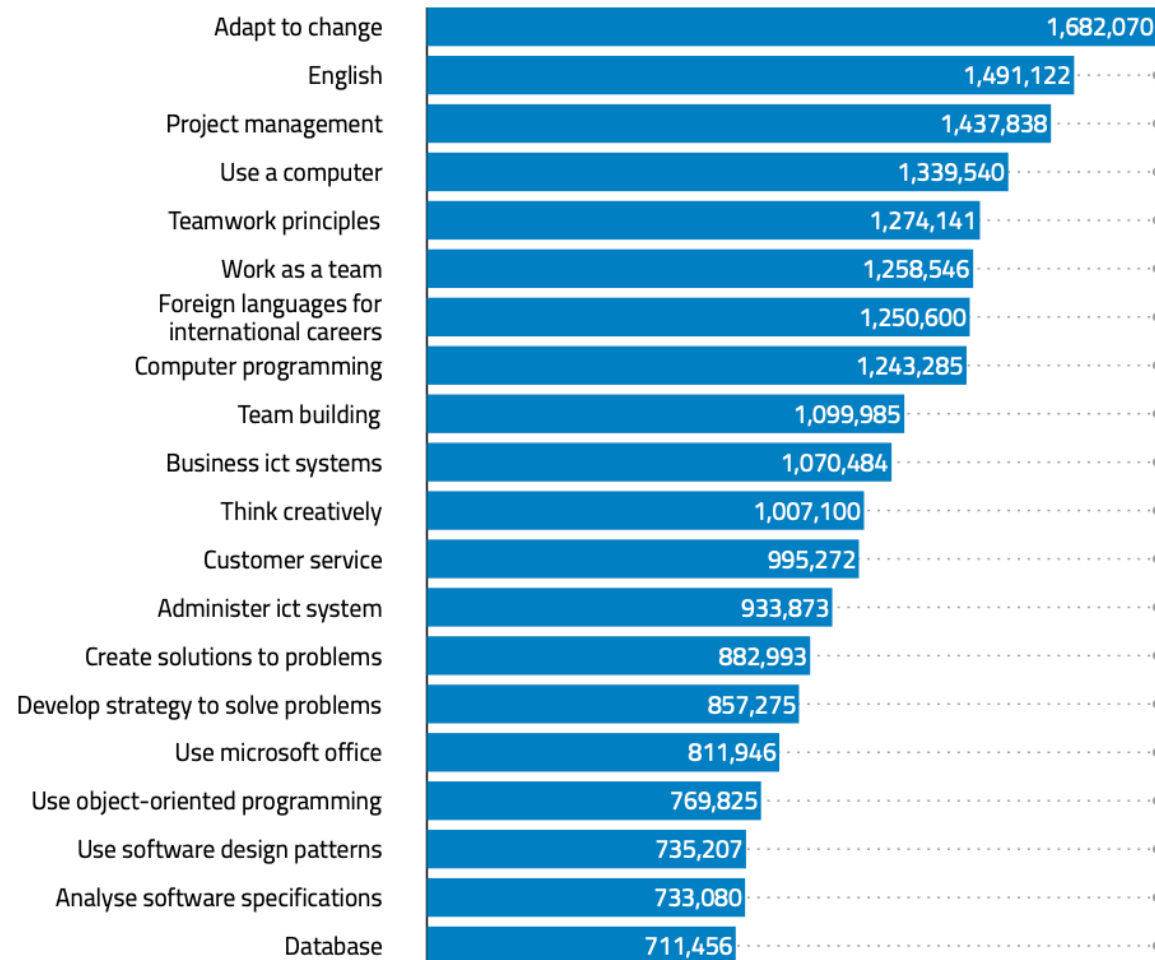
in EU
in ICT professionals

Database and network professionals : 13.3%

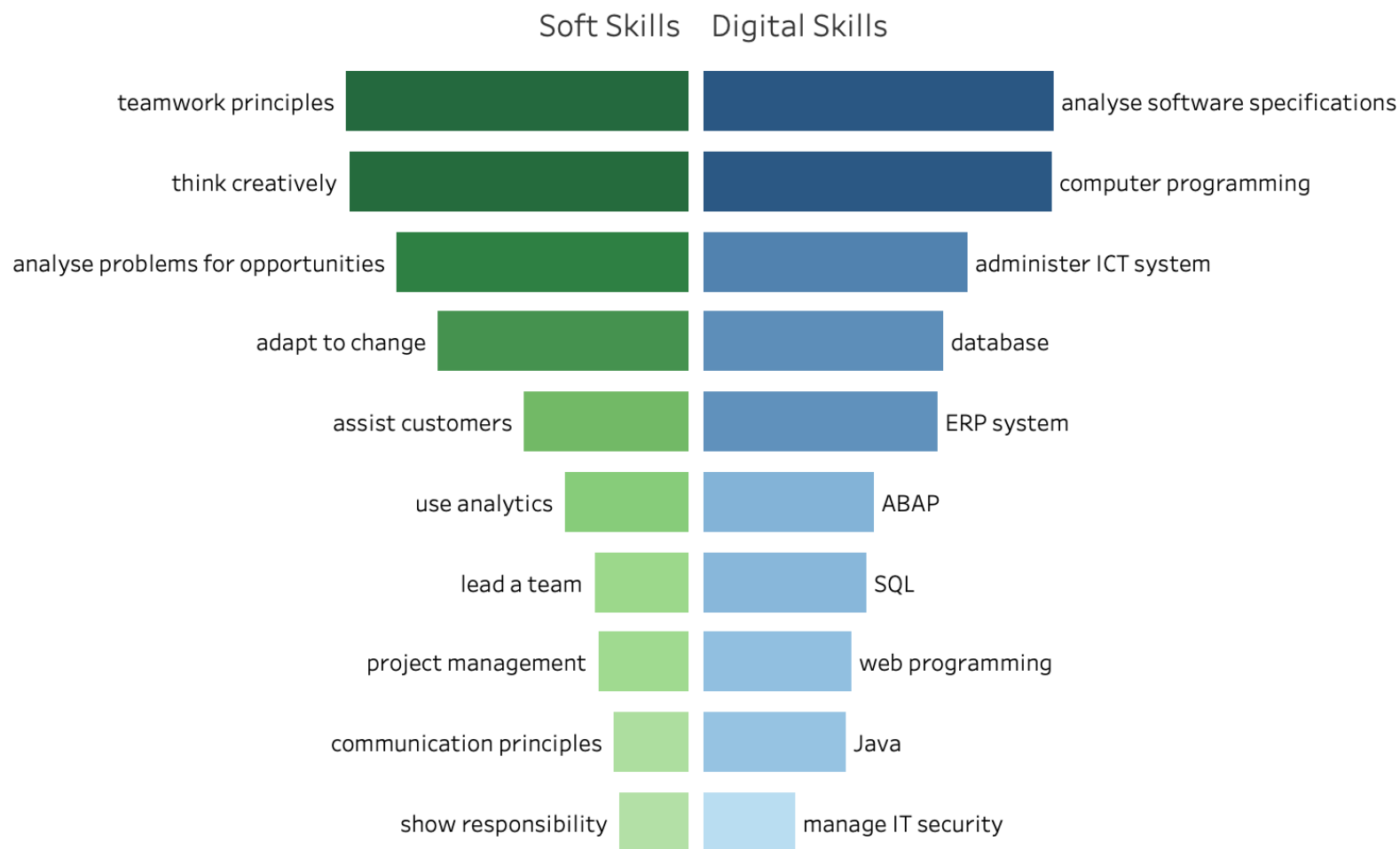


SKILLS PANORAMA
EUROPEAN COMMISSION

Most requested Skills in EU in ICT professionals



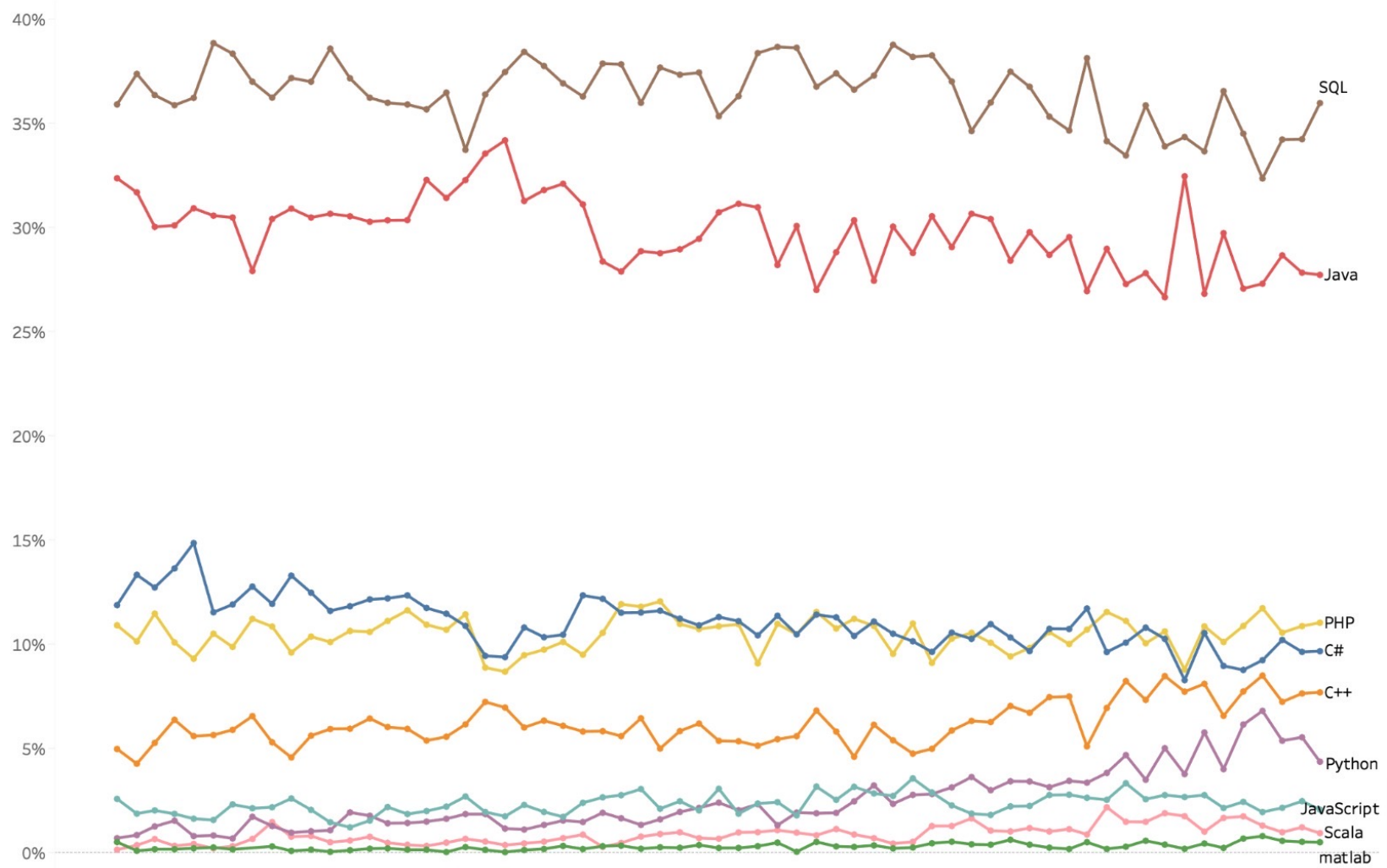
Most relevant skill in digital occupations



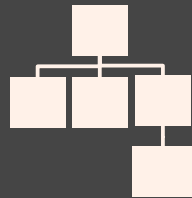
Source: WollyBI Data - Italy 2018

Estimated on 146.567 digital OJVs over 1.330.623 OJVs

Number of Web Job Vacancy by Programming Language

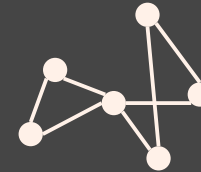


Structured data vs Big data



Structured data

Purposefully collected and collated data which comes in neat, tidy structure. This is typically data from government statistical surveys, designed to ask explicit questions of targeted samples of specific audiences.



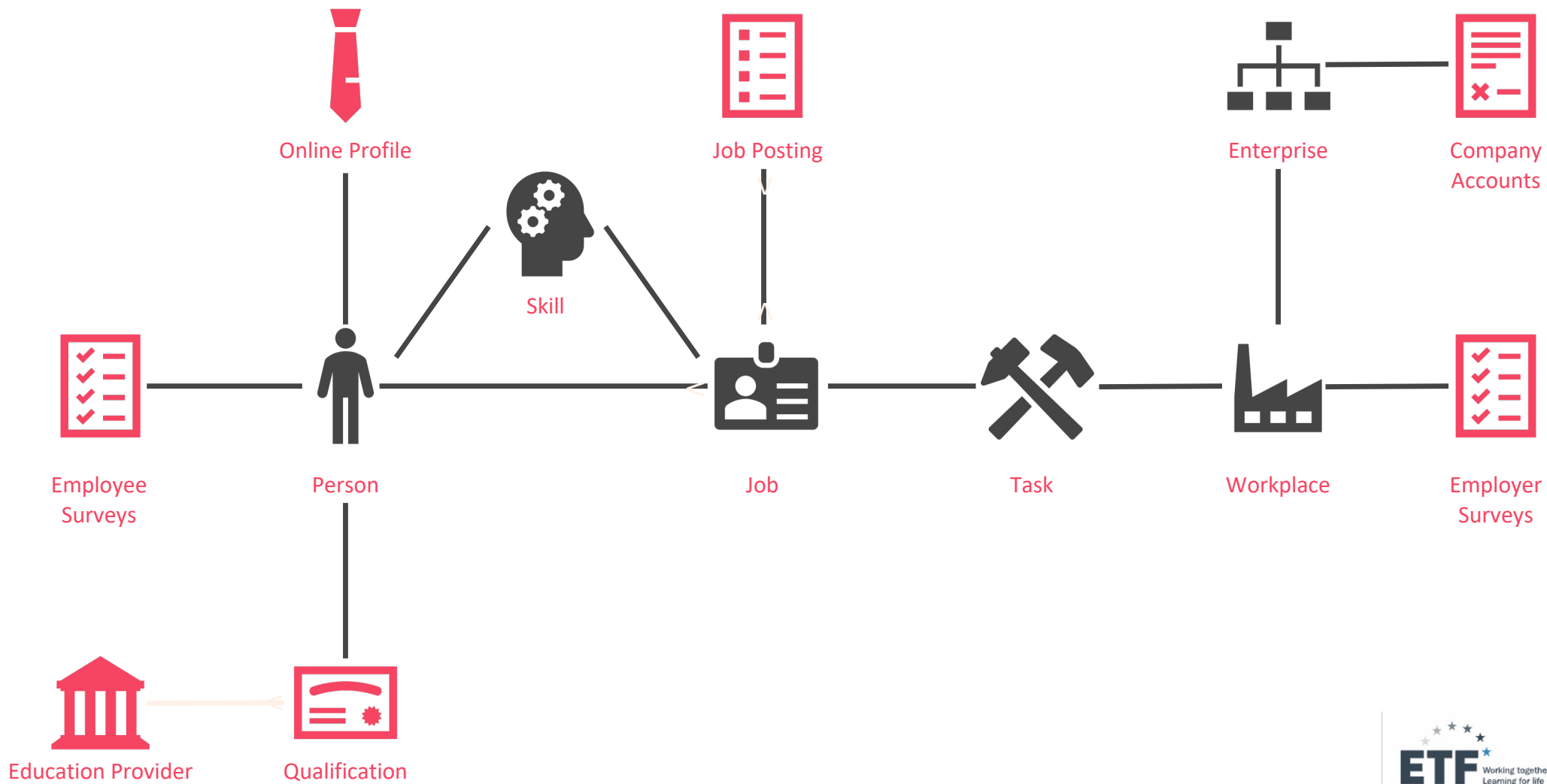
Big data

Extremely large scale data captured from some transactions rather than as a specific data collection exercise. This means harvesting job postings and worker profiles from different web-based sources.

Why use Big data?

We make data available through three primary channels

- **Granular:** big data provides coverage down to 4-digit industry and occupation, down to Local Area Units, with more detail possible.
- **Comprehensive:** big data is made available across all industry/occupation/area cells, covering all cases in the country.
- **Fresh:** job postings data is updated monthly
- **Integrated:** big data is presented as a single set without exceptions, and is made available through tools which allow access for decision-makers rather than only expert analysts



Taxonomies

The three main taxonomies

ISCO08

The International Standard Classification of Occupations is an International Labour Organization classification structure for organizing information on labour and jobs. It is part of the international family of economic and social classifications of the United Nations

NACE

The Statistical Classification of Economic Activities in the European Community, commonly referred to as NACE, is the industry standard classification system used in the European Union. The current version is revision 2 and was established by Regulation No 1893/2006.

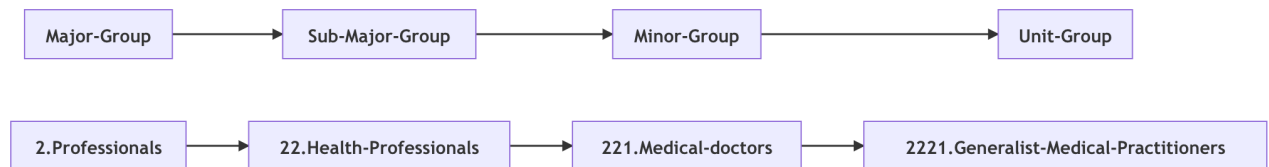
ESCO

ESCO (European Skills, Competences, Qualifications and Occupations) is the European multilingual classification of Skills, Competences and Occupations.

Taxonomies are fully hierarchical

ISCO08

The International Standard Classification of Occupations is an International Labour Organization classification structure for organizing information on labour and jobs. It is part of the international family of economic and social classifications of the United Nations



Major Groups

1 **Managers**

2 **Professionals**

3 **Technicians and Associate Professionals**

4 **Clerical Support Workers**

5 **Service and Sales Workers**

6 **Skilled Agricultural, Forestry and Fishery Workers**

7 **Craft and Related Trades Workers**

8 **Plant and Machine Operators, and Assemblers**

9 **Elementary Occupations**

0 **Armed Forces Occupations**

Total number of groups

Sub Major Groups Minor Groups Unit Groups Skill level

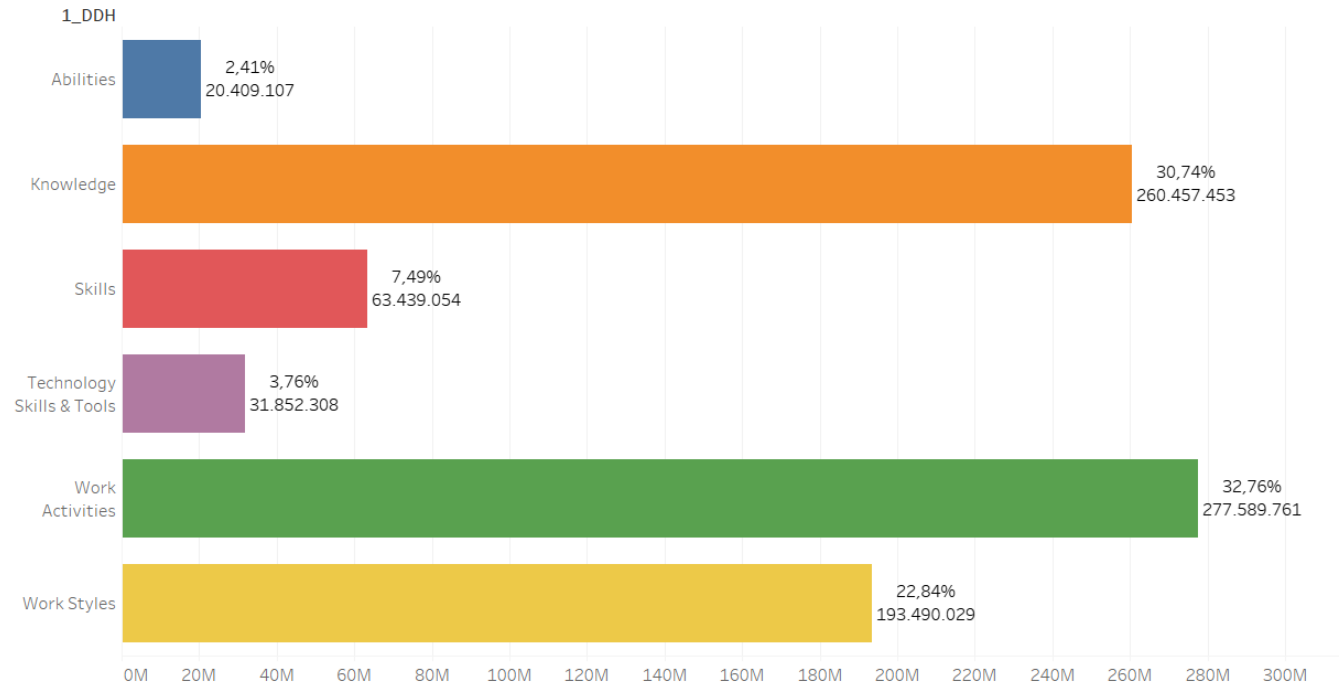
4	11	31	3 and 4
6	27	92	4
5	20	84	3
4	8	29	2
4	13	40	2
3	9	18	2
5	14	66	2
3	14	40	2
6	11	33	2
3	3	3	1, 2 and 4
43	130	436	

Taxonomies are fully hierarchical

ESCO

ESCO (European Skills, Competences, Qualifications and Occupations) is the European multilingual classification of Skills, Competences and Occupations.

DDH1



CEDEFOP

European Centre
for the Development
of Vocational Training



ESCO Skills main groups

- **Abilities** - Enduring attributes of the individual that influence performance
- **Knowledge** - Organized sets of principles and facts applying in general domains
- **Skills** - Developed capacities that facilitate learning or the more rapid acquisition of knowledge
- **Work Activities** - General types of job behaviors occurring on multiple jobs
- **Work Styles** - Personal characteristics that can affect how well someone performs a job.
- **Technology Skills & Tools** - Information technology and software skills or machines, equipment, and tools essential to the performance of an occupational role

Occupation taxonomy

- Group similar roles together to enable easier analysis.
- A role is defined by a distinct mix of knowledge, skills, abilities, activities, and tasks.
- They may include many variations of similar job titles.

Occupation taxonomy

Clustering: Occupations represent clusters of **Job Titles** and **Skills**

Android Developer
ADB Shell Engineer
Linux Mobile
Senior Engineer – Android Mobile Dev
JavaScript, XML, Struts = 2512 Software Developer

Occupation taxonomy **are building blocks**

5821 Publishing computer games

5829 Other software publishing

6110 Wired telecommunications activities

6120 Wireless telecommunications activities

6130 Satellite telecommunications activities

6190 Other telecommunications activities

6201 Computer programming activities

6202 Computer consultancy activities

6203 Computer facilities management activities

6209 Other information technology and computer services

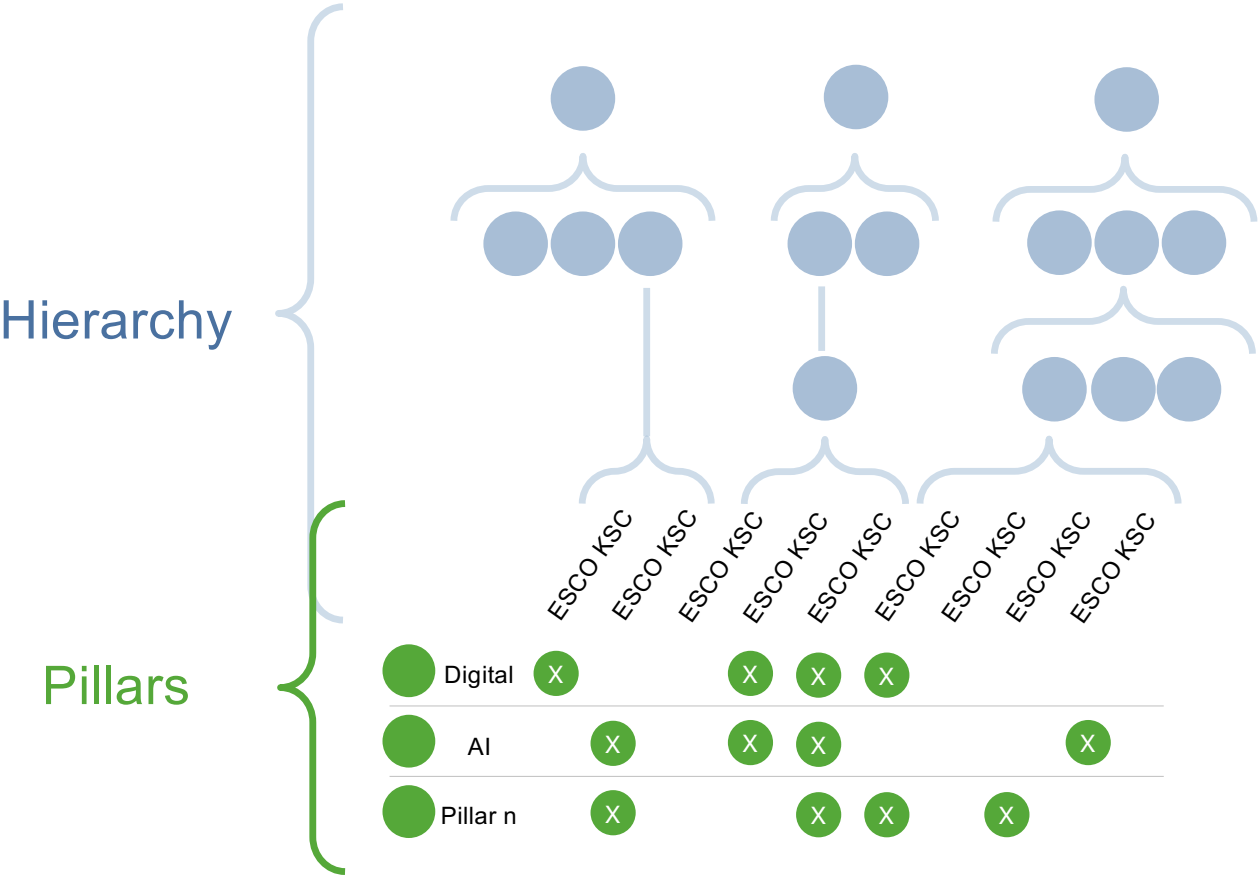
6311 Data processing, hosting and related activities

6312 Web portals



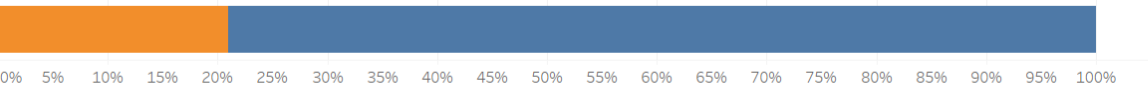
Digital
cluster

Pillars

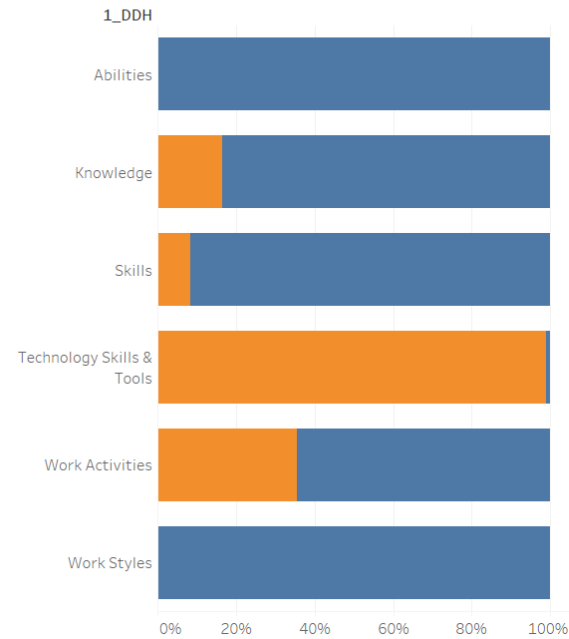


Results on Pillars Digital skills

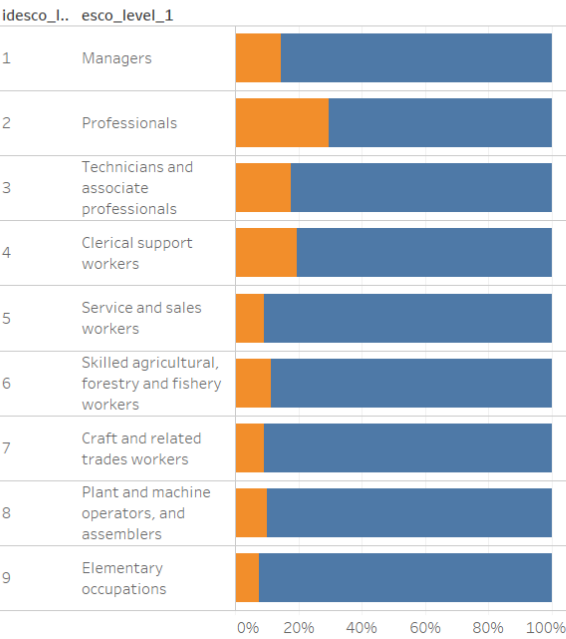
Digital



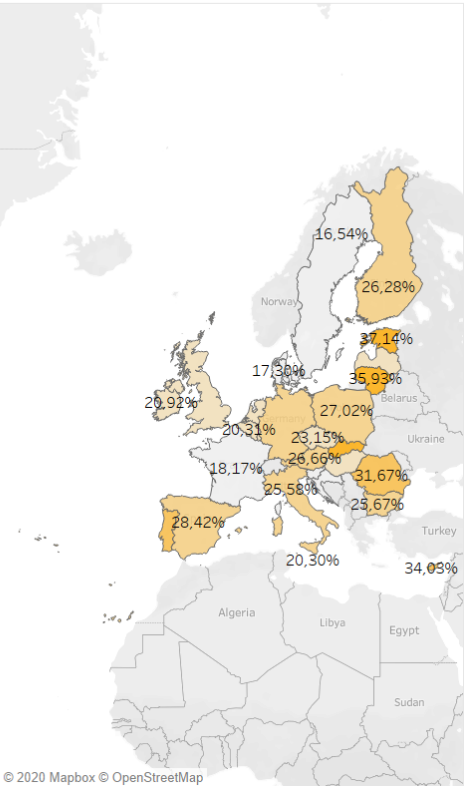
DDH1Digital



Esco1Digital



ContryDigital



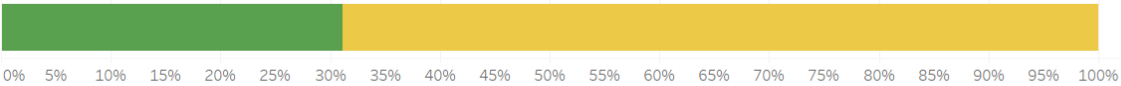
CEDEFOP

European Centre
for the Development
of Vocational Training

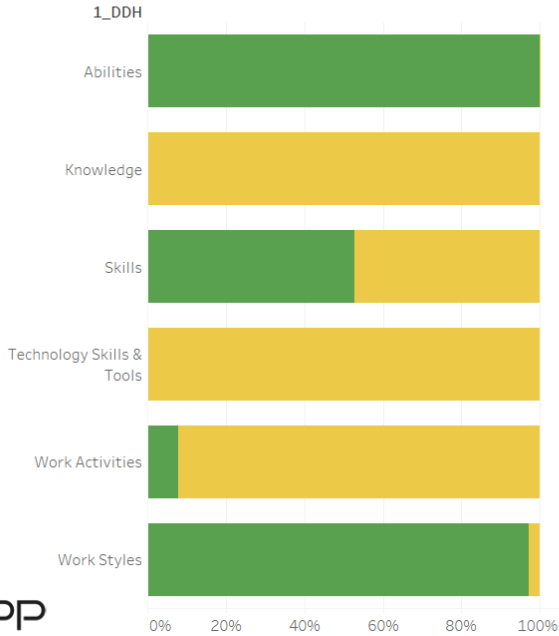


Results on Pillars Soft skills

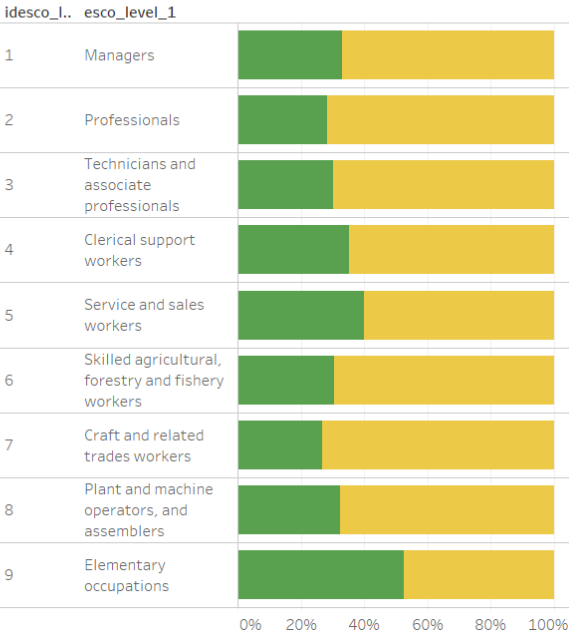
Soft



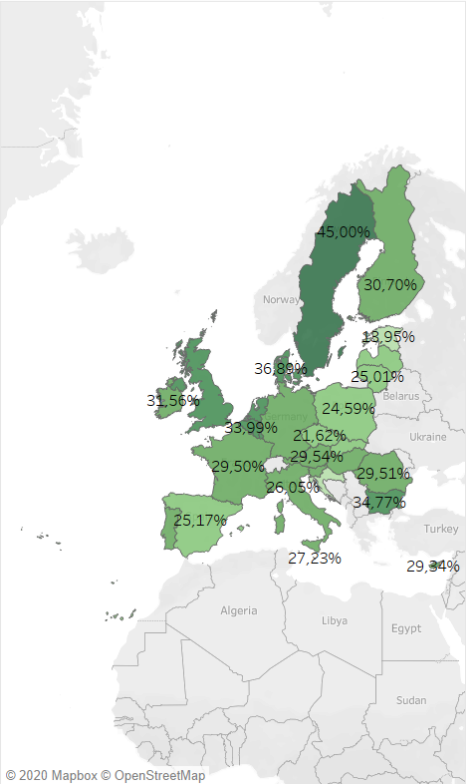
DDH1Soft



Esco1Soft



ContrySoft



Using ONET data

- ONET data is produced by the US Government Department of Labor (v25.2). A range of occupations are measured against the ONET Content Model using an ongoing survey process. The Content Model is described here: <https://www.onetcenter.org/content.html>
 - Knowledge, Skills, Tasks and Work Conditions are measured by an Importance scale (IM) from 1 (Not important) to 5 (Extremely important)
 - Knowledge and Skills are also measured by a Level scale (LV) from 0 to 7.
 - Tasks are measured by a Relevance scale (RT) which is a percentage scale.
 - Work Context is measured by Context scale (CX) from 1 to 5.

O*Net-SOC Example

O*NET SOC takes things one step farther

Major Group ♦	Minor Group ♦	Broad Occupation ♦	Detailed Occupation ♦	Detailed O*NET-SOC ♦	SOC or O*NET-SOC 2019 Title
11-0000					Management Occupations
	11-1000				Top Executives
		11-1010			Chief Executives
			11-1011		Chief Executives
				11-1011.03	Chief Sustainability Officers

Hierarchical Comparison

O*NET SOC takes things one step farther

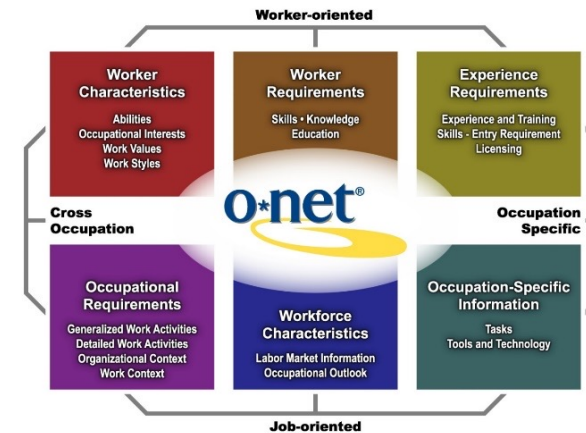
Major Group	Minor Group	Broad Occupation	Detailed Occupation	Detailed O*NET-SOC	SOC or O*NET-SOC 2019 Title
11-0000					Management Occupations
	11-1000				Top Executives
		11-1010			Chief Executives
			11-1011		Chief Executives
				11-1011.03	Chief Sustainability Officers

About O*NET -1

- The 'Occupational Information Network' ([O*NET](#)) is a free online database that contains hundreds of occupational definitions developed during the 1990s under the sponsorship of the US Department of Labor/Employment and Training Administration (USDOL/ETA).
- USDOL/ETA describes the O*NET as: a database of occupational requirements and worker attributes. It describes occupations in terms of the **skills** and **knowledge** required, how the work is performed, and typical work settings.

About O*NET-2

- The [Content Model](#) is the
- conceptual foundation of O*NET.
- The Content Model is organized
- into six major domains and provides a framework that identifies the most important types of information about work and integrates them into a theoretically and empirically sound system.
- [O*NET OnLine](#) is a tool for career exploration and job analysis, in this tool you can browse O*NET data according to the O*NET Content Model.



O*NET OnLine categories - 1

At its first level O*NET includes the following Data Descriptors:

- **Abilities** - Enduring attributes of the individual that influence performance.
- **Interests** - Preferences for work environments and outcomes.
- **Knowledge** - Organized sets of principles and facts applying in general domains.
- **Skills** - Developed capacities that facilitate learning or the more rapid acquisition of knowledge.
- **Work Activities** - General types of job behaviors occurring on multiple jobs.

O*NET OnLine categories - 2

- **Work Context** - Physical and social factors that influence the nature of work.
- **Work Styles** - Personal characteristics that can affect how well someone performs a job.
- **Work Values** - Global aspects of work that are important to a person's satisfaction.
- **Technology Skills & Tools** - Information technology and software skills or machines, equipment, and tools essential to the performance of an occupational role

Real-Time Labour Market Information System on Skill Requirements

Continuously evolving Labour Market

