# Big Data for Labour Market Intelligence

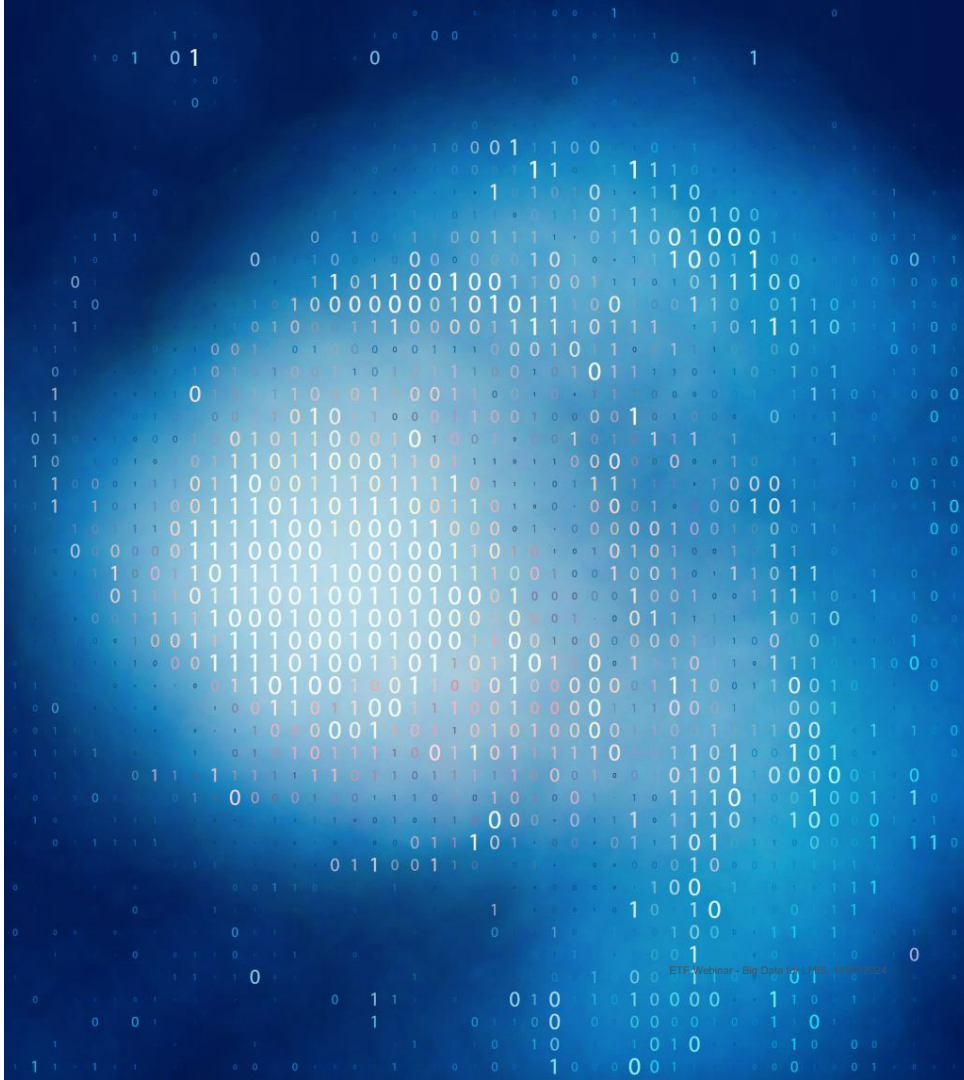
## Webinar 14 Feb 2024

ETF Working together Learning for life

European Training Foundation

Eduarda Castel-Branco

ecb@etf.europa.eu

# Welcome to the conversation

2 conversations: 14 and 16 February
09.00 – 13.00

## Main themes

- Big Data for LMIS: skills demand analysis
- Combination and comparison of OJV data with traditional data
- AI and education
- Occupational profiles: developing and testing a data driven approach based on Big Data

## Speakers

- Mauro Pelucchi
- Anna Gatti
- Emilio Colombo
- Fabio Mercorio
- Elena Magrini

ETF Working together Learning for life
European Training Foundation

# Getting acquainted

**Mentimeter 14/Feb – Webinar Big Data for LMIS**

**Menti.com – Code: 6130 8707**

https://www.menti.com/al77e8puecdj

ETF Working together Learning for life

European Training Foundation

# The journey of Big Data LMIS ETF project

- Handbook “Big Data for LMIS: an introductory Guide”
- 6 countries: Egypt, Morocco, Tunisia, Kenya; Ukraine, Georgia
- 3 languages of dashboards: EN-FR-AR
- 5 training programs
- 45 pedagogic videos on YouTube channel
- Over 12.3 million OJV collected
- New analyses: green skills, digital skills, remote work, supply side

ETF Working together Learning for life

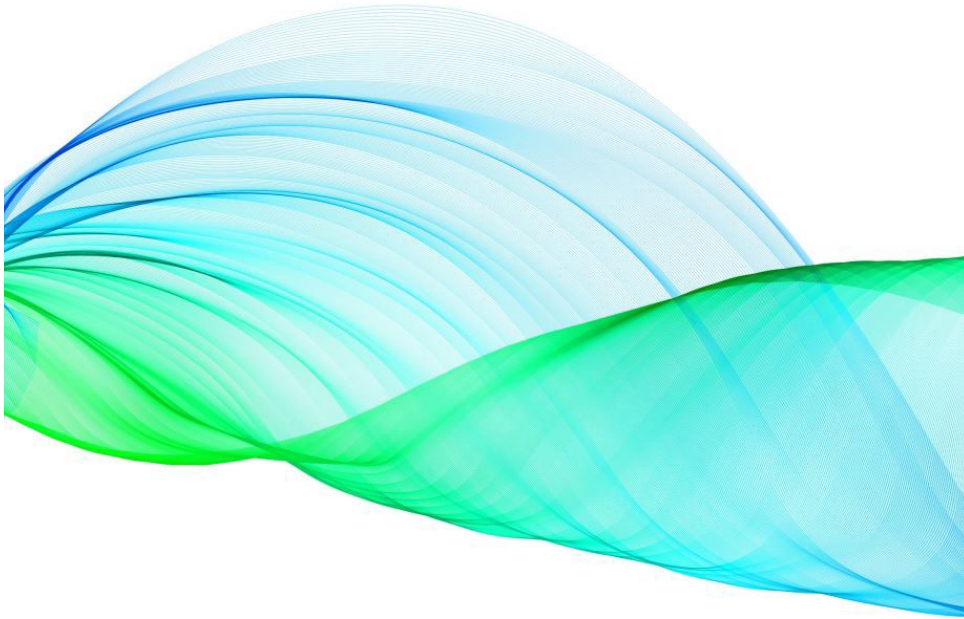
European Training Foundation

# Starting from the end…

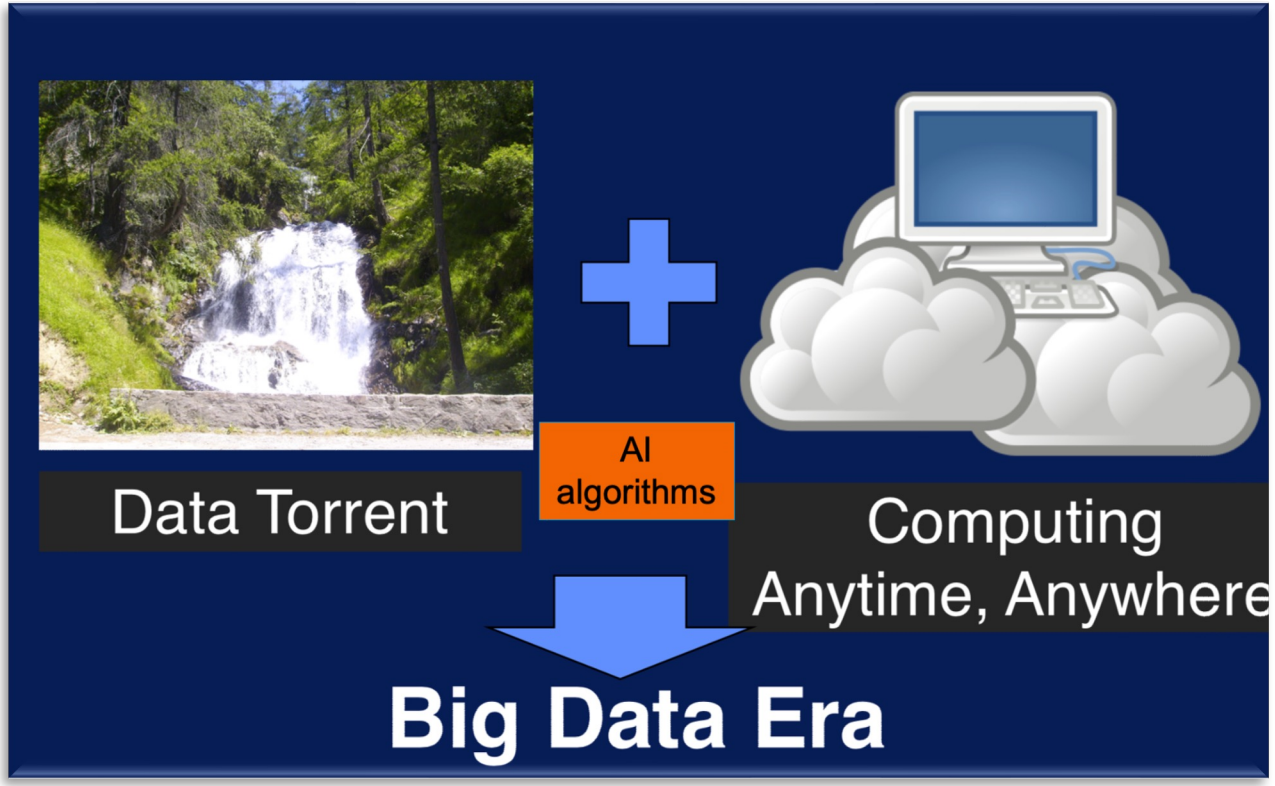
**The value added of Big Data for LMI**:

- Volume, velocity, variety
- Real-time
- Granularity
- Combinable
- International classifications
- Early spotting of new trends – for further deeper and wider analysis (combining other data sources)
- “Let the Data Speak”

**Challenges exist**



# Big Data for Labour Market and Skills Intelligence



# Large volumes of data from employers' job vacancies online: new source and value added for Labour Market and Skills Intelligence

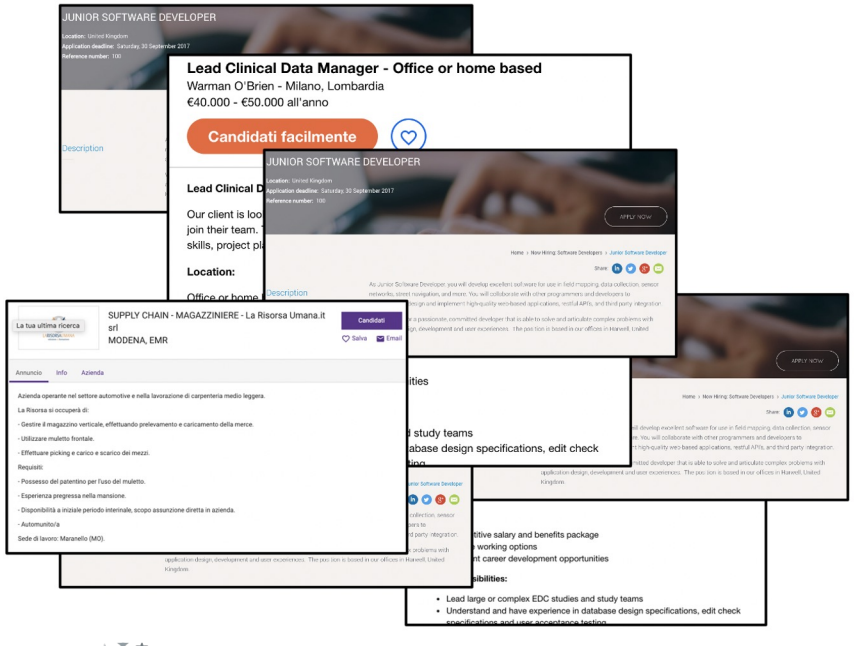
**Online Job advertisements / vacancies**

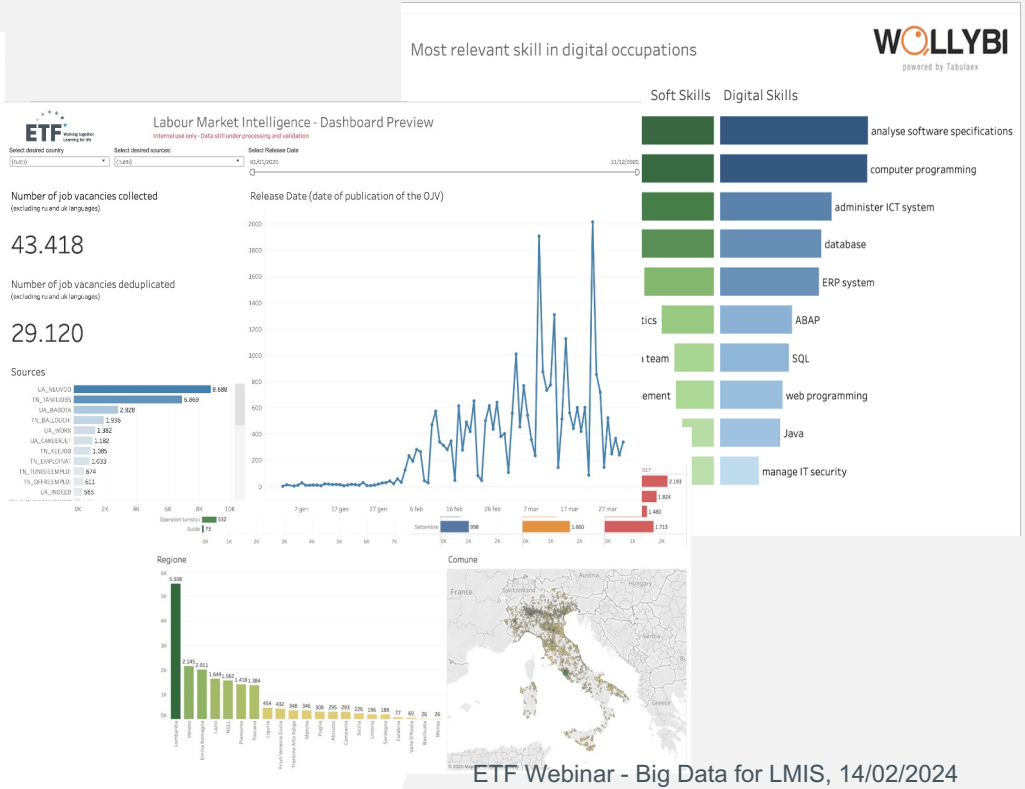
**(OJV)**

# THE GOAL OF Online Job Vacancy ANALYSIS IS…

## To transform this…

## …into value

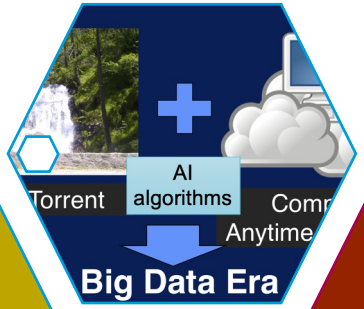



ETF Working together Learning for life

European Training Foundation

# OJV analysis system: building blocks




3. International Classifications ESCO, ISCO, NACE…

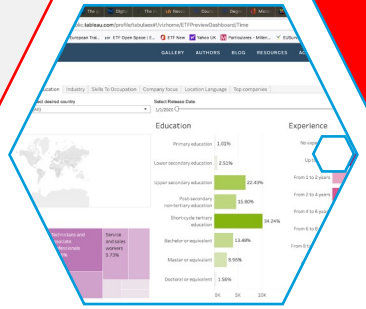
4. Visualisation Dashboard - variables

1. Data from OJV

Complements statistics

2. AI-aided data system

5. Let the data speak




6. Volume, Velocity, Variety, Veracity, Value

# Big Data for LMI - OJV DATA

**All country Dashboards at**: https://solutions.lightcast.io/?pc=x$fhADtD*cu$BjY9

**Methodological Guide**: https://www.etf.europa.eu/en/publications-and-resources/publications/big-data-labour-market-intelligence-introductory-guide

- Experimental project – data production system based on internet data
- Started: 2019 (Handbook)
- Data science expertise: Lightcast.
- Data: constant inflow; updates: monthly
- 12,5 million OJV collected in the Database

Countries in ETF Database & dashboards:

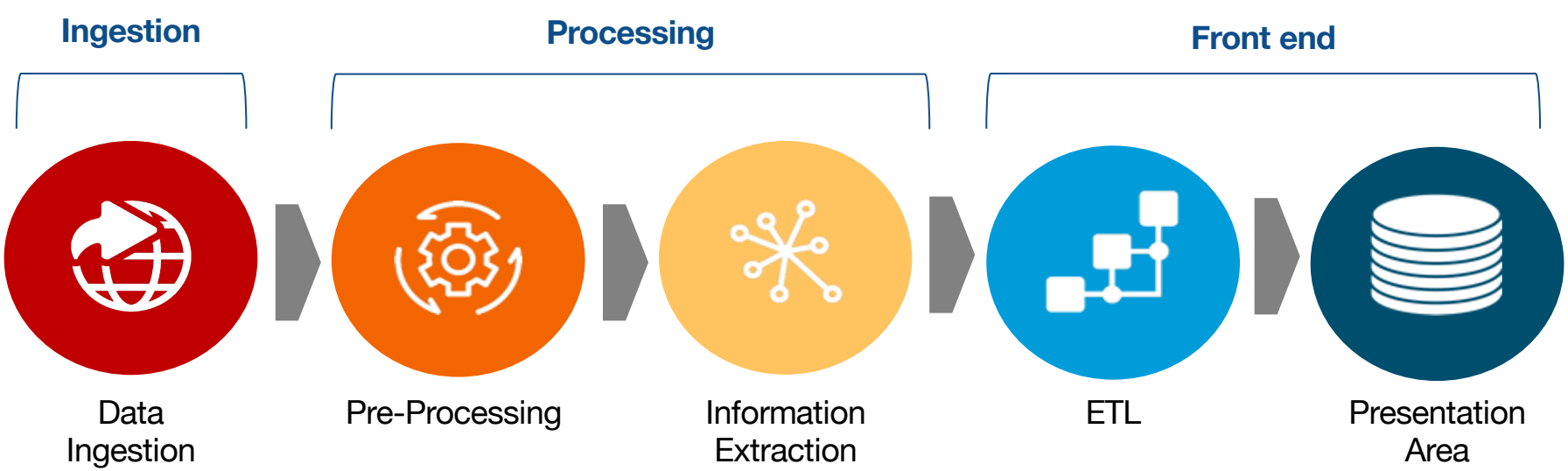
- Ukraine
- Tunisia
- Georgia
- Kenya
- Egypt
- Morocco

## Standard variables on ETF dashboards - Analysis by:

- Occupation
- Occupation– education
- Occupation – industry
- Industry
- Occupation to skill (ESCO)
- Occupation to skill (O*Net)
- Location – language
- Professional dashboard
- New: green skills, digital skills

# DATA FLOW

**Landscaping and assessing OJV sources in given labour market**



# Dashboards

## Interactive presentation of results – classified OJV data

All country Dashboards at:
https://solutions.lightcast.io/?pc=x$fhADtD*cu$BjY9

## Examples

# Professional dashboard (all major variables) – Kenya

Home | Time | Professional Dashboard | Green skill | Digital skill | Occupation | Occupation to skill | Occupation to skill - O*NET | Occupation - Education | Occupation - Industry | Industry | Location Language | Fo

ETF Working together Learning for life

If you want, you can filter by occupation family:

ESCO (Level 1): (All)

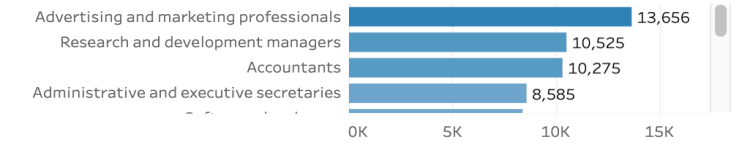
ESCO (Level 2): (All)

ESCO (Level 3): (All)

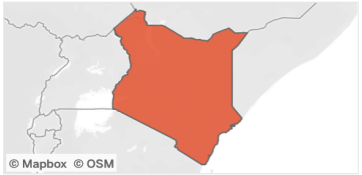
Select Release Date

October 2022 — December 2023

## Select the Occupation that you want to Analyze (then select it againg to reset your choice)

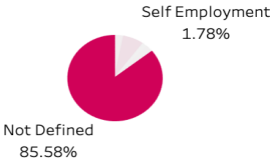



## Location

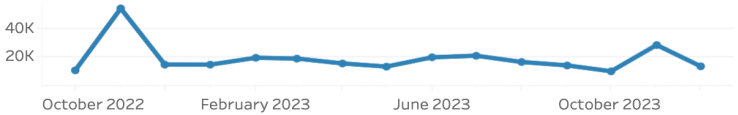



## Industry (level 1 NACE)

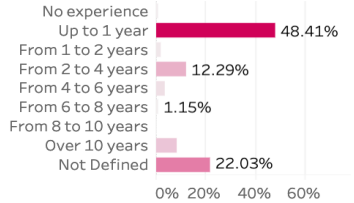
## Contract



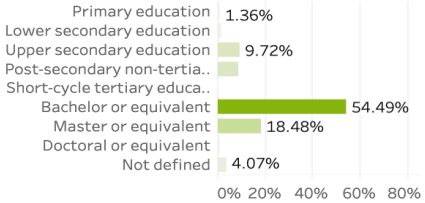

## Distribution by Release Date (date of publication of the OJV)



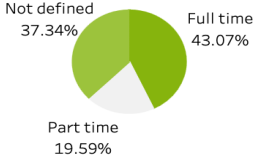

## Experience



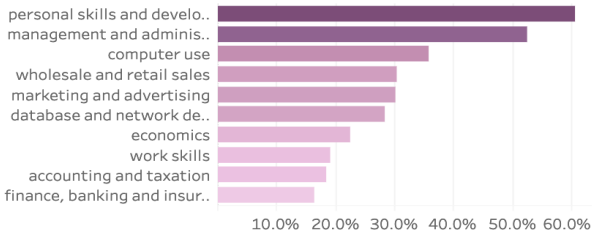

## Education



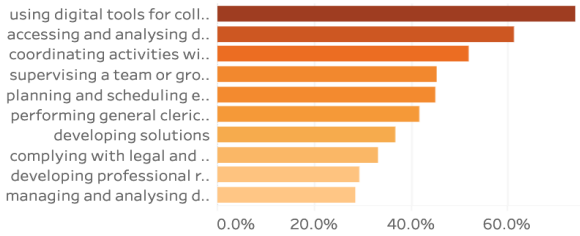

## Working hours



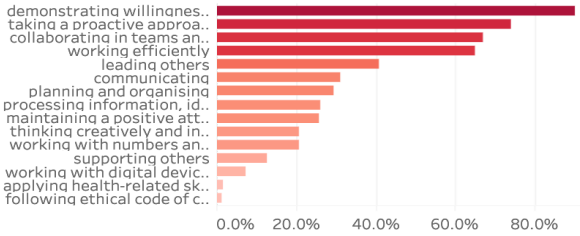

## Knowledge




## Skills

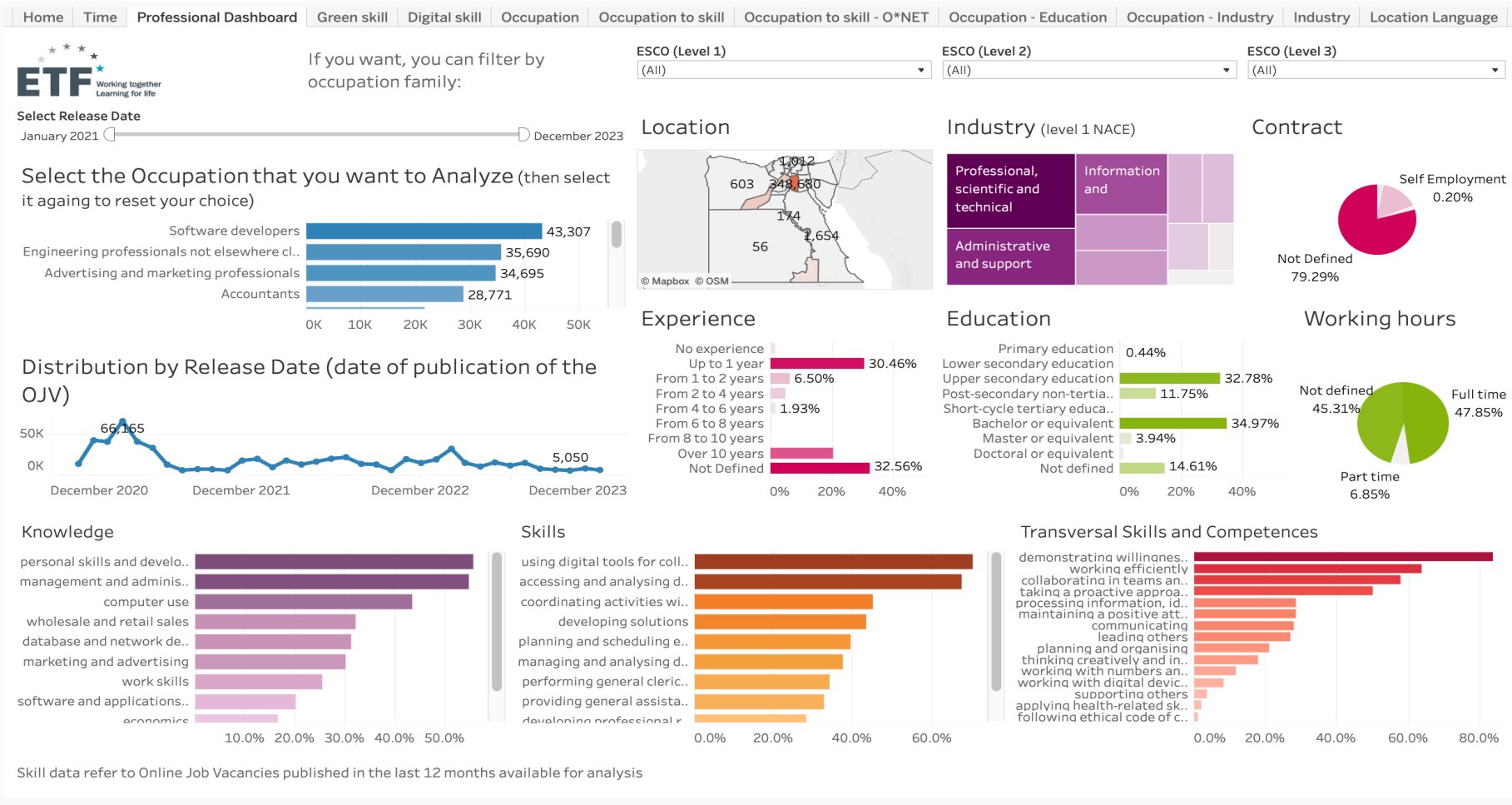



## Transversal Skills and Competences




Skill data refer to Online Job Vacancies published in the last 12 months available for analysis

# Professional dashboard - Egypt



# Remote work - Ukraine

Green skill | Digital skill | Occupation | Occupation to skill | Occupation to skill - O*NET | Occupation - Education | Occupation - Industry | Industry | Remote work | Location Language | Company focus

ETF Working together Learning for life

Select Release Date

April 2020 — December 2023

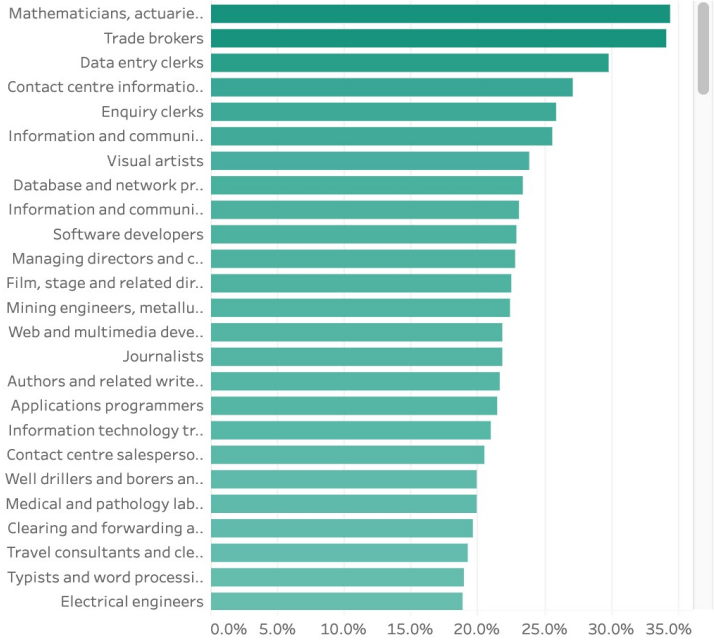
## Remote Work Share

5.41%

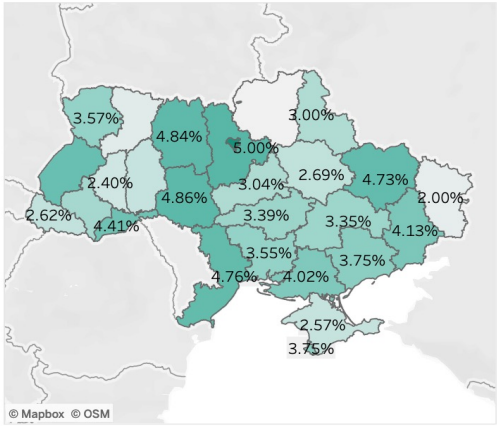
## Number of remote work OJV

161,210

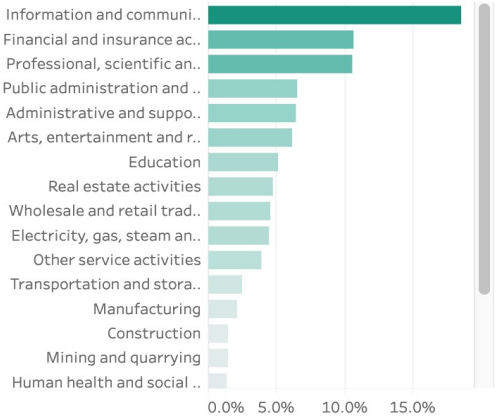
## Select the Occupation that you want to Analyze (then select it againg to reset your choose)



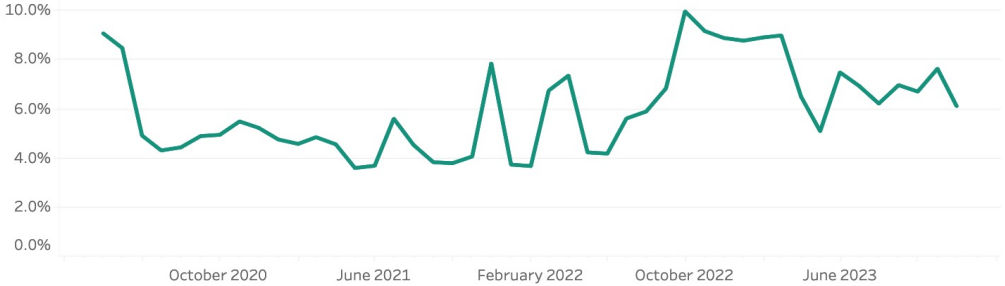

## Share of remote work by region




## Share of remote work by industry




## Share of remote work by release date

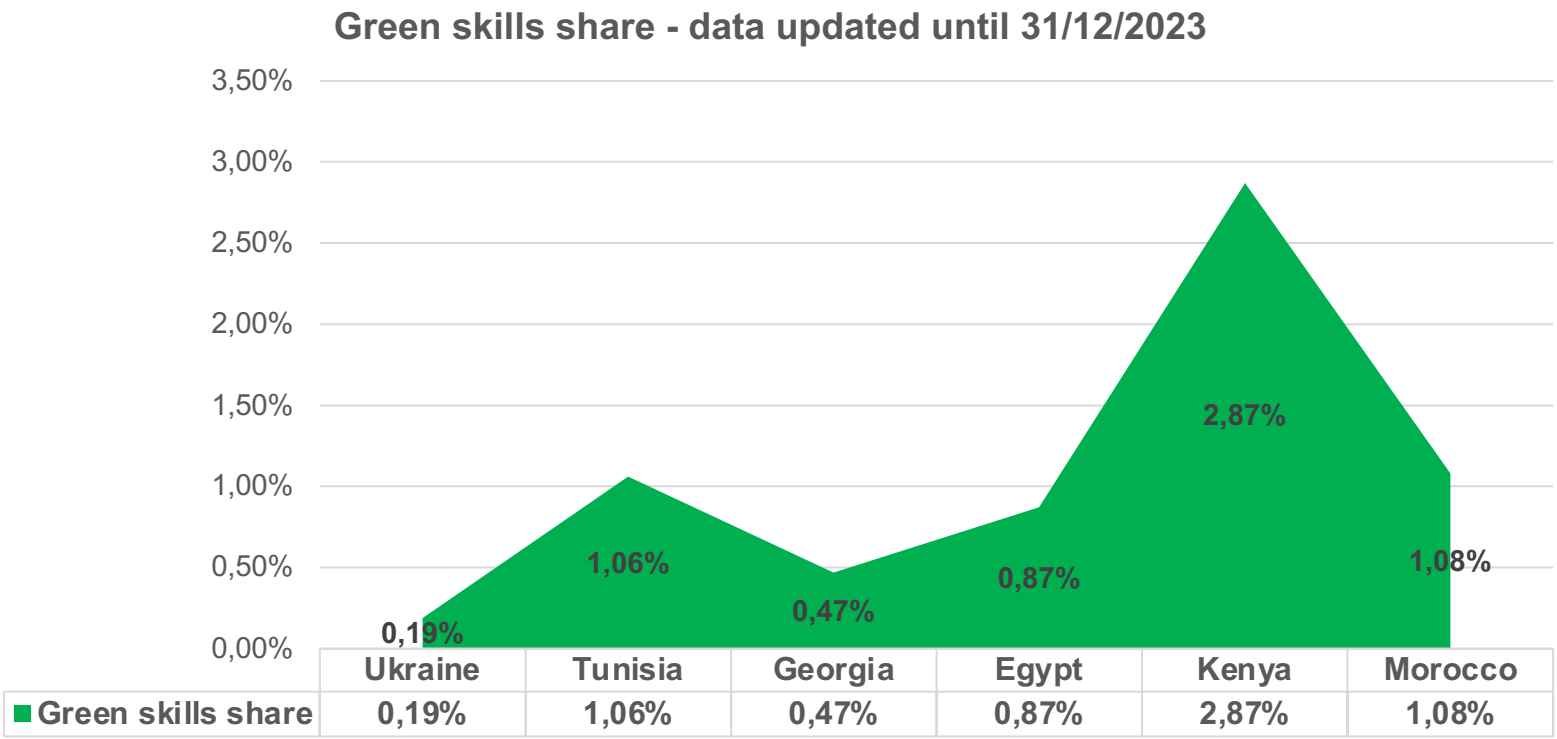



tableau

# GREEN SKILLS

## IN ETF OJV DATA

# ETF project: Online job vacancy – Big data for LMI




Green skills share = share of online job vacancies that required at least 1 green skill

Kenya: highest green skills share, Ukraine – the lowest

Used green skills taxonomy: ETF (225 terms)

ETF Working together Learning for life
European Training Foundation

# ETF - BigData LMI - Green Dashboard by Lightcast

ETF Working together Learning for life

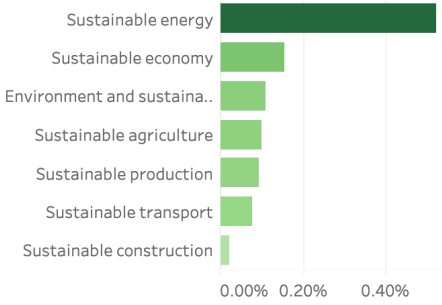
**Select Release Date**

October 2022 — December 2023

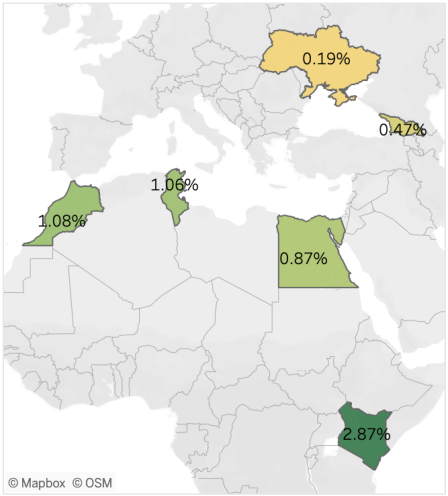
## Green Share

**1.02%**

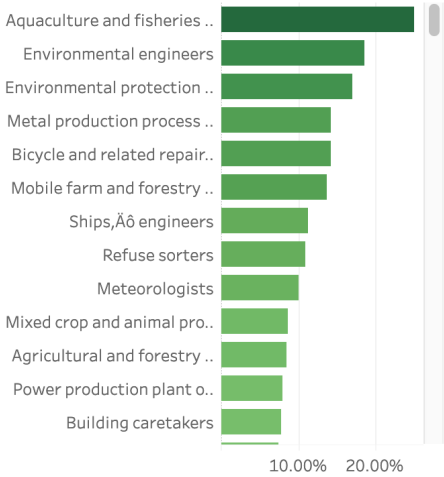
Select the Green skill family that you want to Analyze (then select it againg to reset your choose)



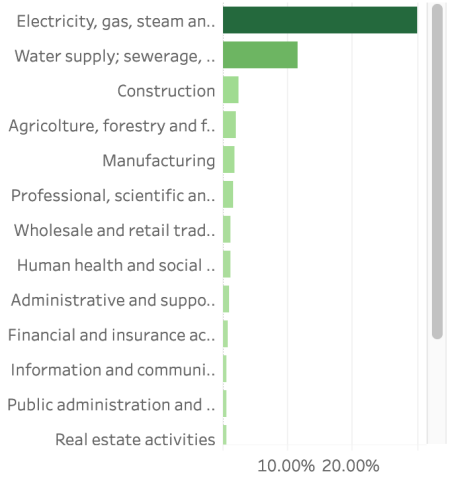

## Share of green skill by region




## Share of green skill by occupation

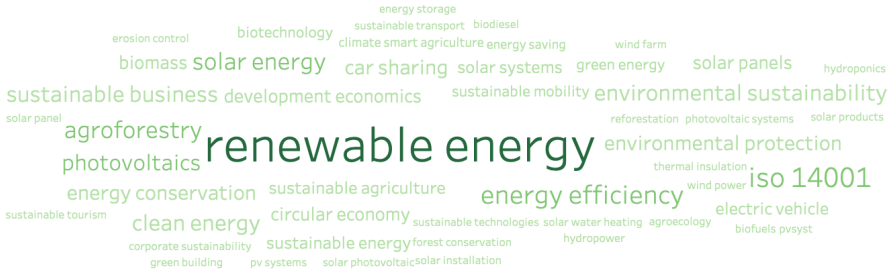



## Share of green skill by industry




## Share of green skill by release date

## Most requested Green Skill



# KENYA

Home | Time | Professional Dashboard | Green skill | Digital skill | Occupation | Occupation to skill | Occupation to skill - O*NET | Occupation - Education | Occupation - Industry | Industry | Location Language

ETF Working together Learning for life

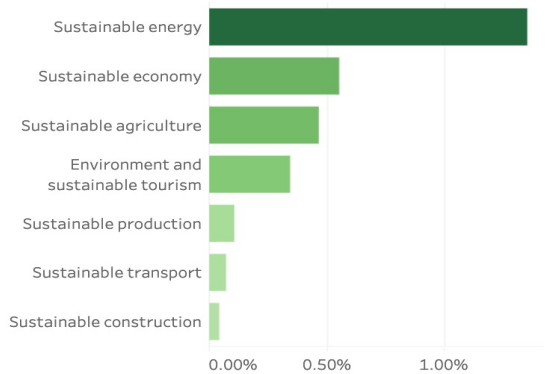
Select Release Date

October 2022 — August 2023

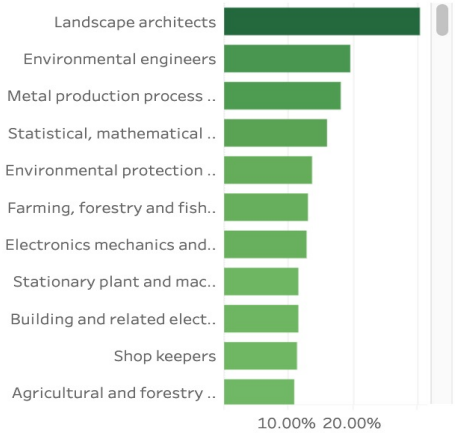
## Green Share

2.64%

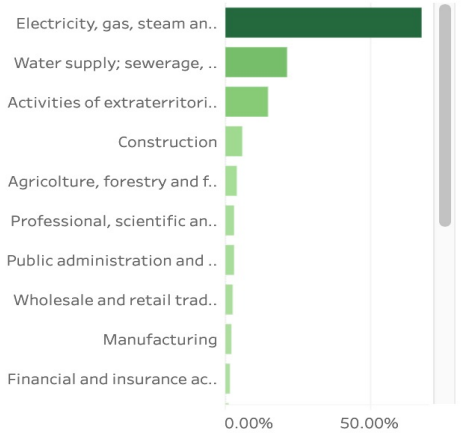
## Select the Green skill family that you want to Analyze (then select it againg to reset your choose)



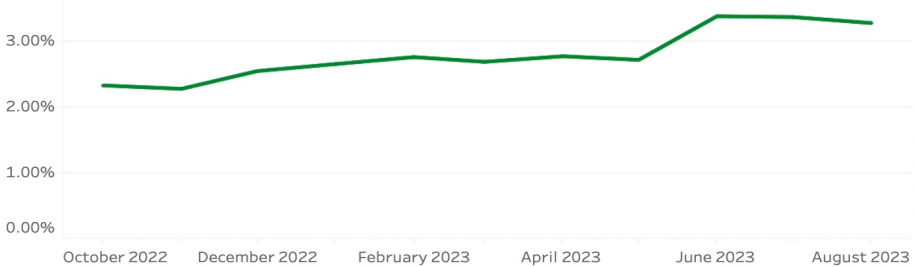

## Share of green skill by occupation




## Share of green skill by industry




## Share of green skill by release date




## Most requested green skill

# Top 20 Green Skills Kenya

| Skills / skill set | % (from 20) | Nº unique job postings |
|---|---|---|
| renewable energy | 22,95% | 2.598 |
| agroforestry | 9,08% | 1.028 |
| clean energy | 6,54% | 740 |
| sustainable business | 6,47% | 732 |
| solar energy | 6,27% | 710 |
| circular economy | 5,64% | 638 |
| environmental sustainability | 5,48% | 620 |
| development economics | 5,05% | 572 |
| biomass | 4,22% | 478 |
| solar systems | 3,45% | 390 |
| environmental protection | 3,41% | 386 |
| energy efficiency | 3,41% | 386 |
| sustainable energy | 3,38% | 382 |
| sustainable agriculture | 2,99% | 338 |
| climate smart agriculture | 2,54% | 288 |
| green energy | 2,31% | 262 |
| iso 14001 | 2,19% | 248 |
| agroecology | 1,80% | 204 |
| solar products | 1,63% | 184 |
| electric vehicle | 1,18% | 134 |

# Kenya: Green skills required for selected occupations

| Occupation | Green skill | Unique Job postings (Oct 22 - Jul 23) |
|---|---|---|
| Electronics mechanics and servicers | solar panels | 7 |
| | solar energy | 2 |
| | environmental protection | 2 |
| Environmental engineers | environmental sustainability | 32 |
| | environmental protection | 19 |
| | sustainable procurement | 9 |
| | renewable energy | 9 |
| | sustainable business | 7 |
| | energy efficiency | 7 |
| | sustainable energy | 5 |
| | circular economy | 4 |
| | development economics | 3 |
| | climate smart agriculture | 3 |
| | green energy | 2 |
| | ecotourism | 2 |
| | clean energy | 2 |
| | carbon footprint reduction | 2 |
| | sustainable agriculture | 1 |
| | sustainability performance | 1 |
| | iso 14001 | 1 |
| | hydroponics | 1 |
| | clean technology | 1 |
| | agroforestry | 1 |
| Metal production process controllers | iso 14001 | 2 |
| Statistical, mathematical and related associate professionals | environmental protection | 18 |
| | clean energy | 18 |
| | energy efficiency | 11 |
| | green building | 9 |
| | sustainable materials | 6 |
| | sustainable building | 6 |
| | sustainable agriculture | 6 |
| | renewable energy | 5 |
| | circular economy | 5 |
| | biomass | 4 |
| | agroforestry | 4 |
| | sustainable energy | 3 |
| | hydropower | 1 |
| | biofuels | 1 |

# Digital skills – overview of some features of demand

Analysis based on ETF data OJV - Egypt

## Taxonomy

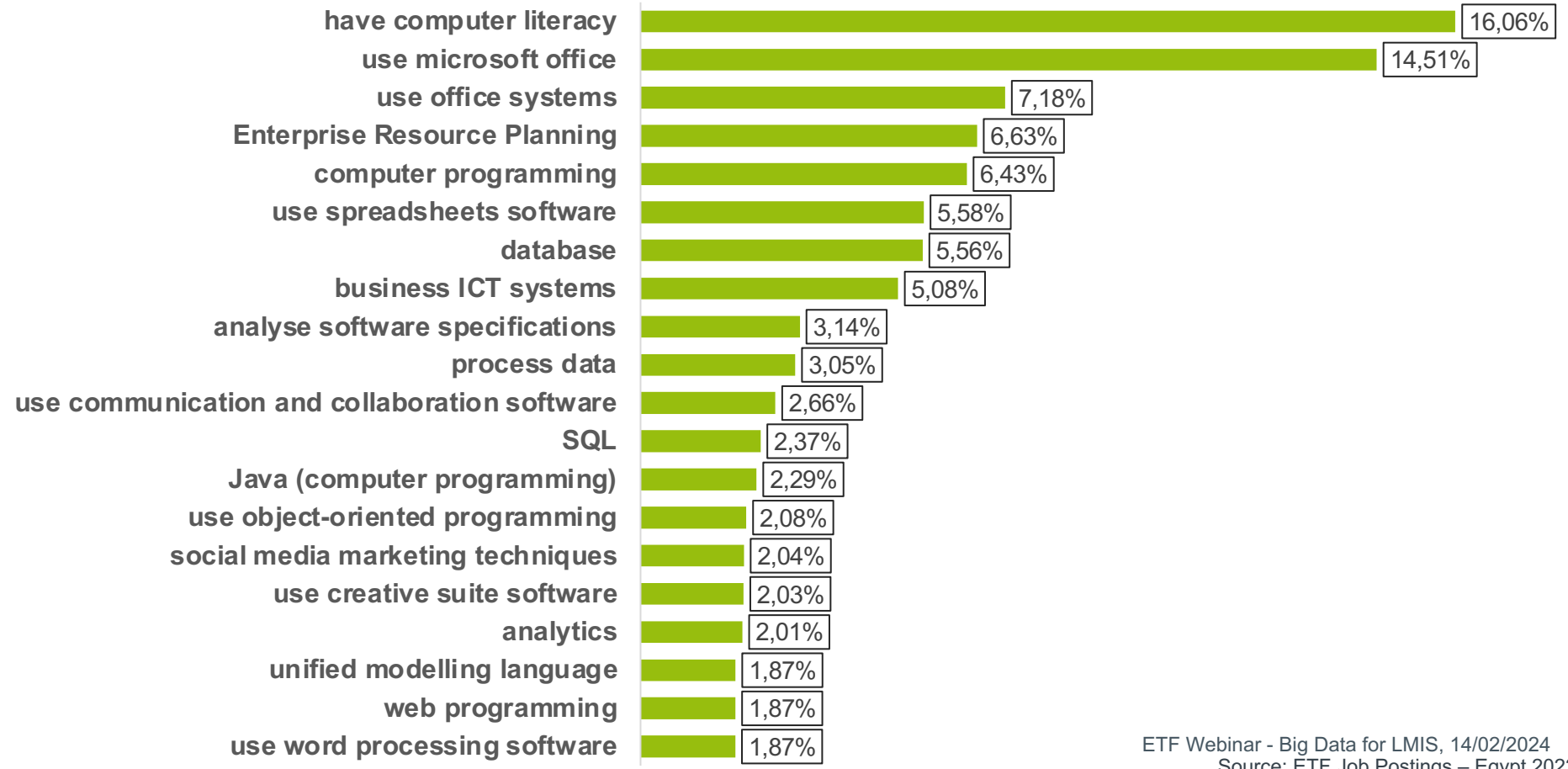
- **ESCO: 1,100 digital skills terms**

- Other sources of digital skills terms: Stackoverflow and Git-Hub – to ease classification

# Kenya: Top 20 digital skills in Job Postings (online)

| Digital skill | | Unique Job postings (Oct 22-Jul 23) |
|---|---|---|
| use microsoft office | 21,2% | 133.629 |
| have computer literacy | 19,2% | 121.093 |
| office software | 7,6% | 47.684 |
| database | 5,7% | 35.886 |
| use spreadsheets software | 5,4% | 33.935 |
| computer programming | 5,4% | 33.775 |
| business ICT systems | 4,0% | 25.000 |
| perform data analysis | 3,8% | 24.040 |
| online analytical processing | 3,5% | 21.879 |
| process data | 3,2% | 19.941 |
| social media marketing techniques | 3,0% | 19.083 |
| use communication and collaboration software | 2,9% | 18.554 |
| use word processing software | 2,6% | 16.306 |
| analyse software specifications | 2,3% | 14.409 |
| digital marketing techniques | 2,0% | 12.549 |
| use creative suite software | 1,9% | 11.825 |
| use software design patterns | 1,7% | 10.686 |
| computer science | 1,7% | 10.611 |
| use object-oriented programming | 1,5% | 9.463 |
| administer ICT system | 1,5% | 9.337 |

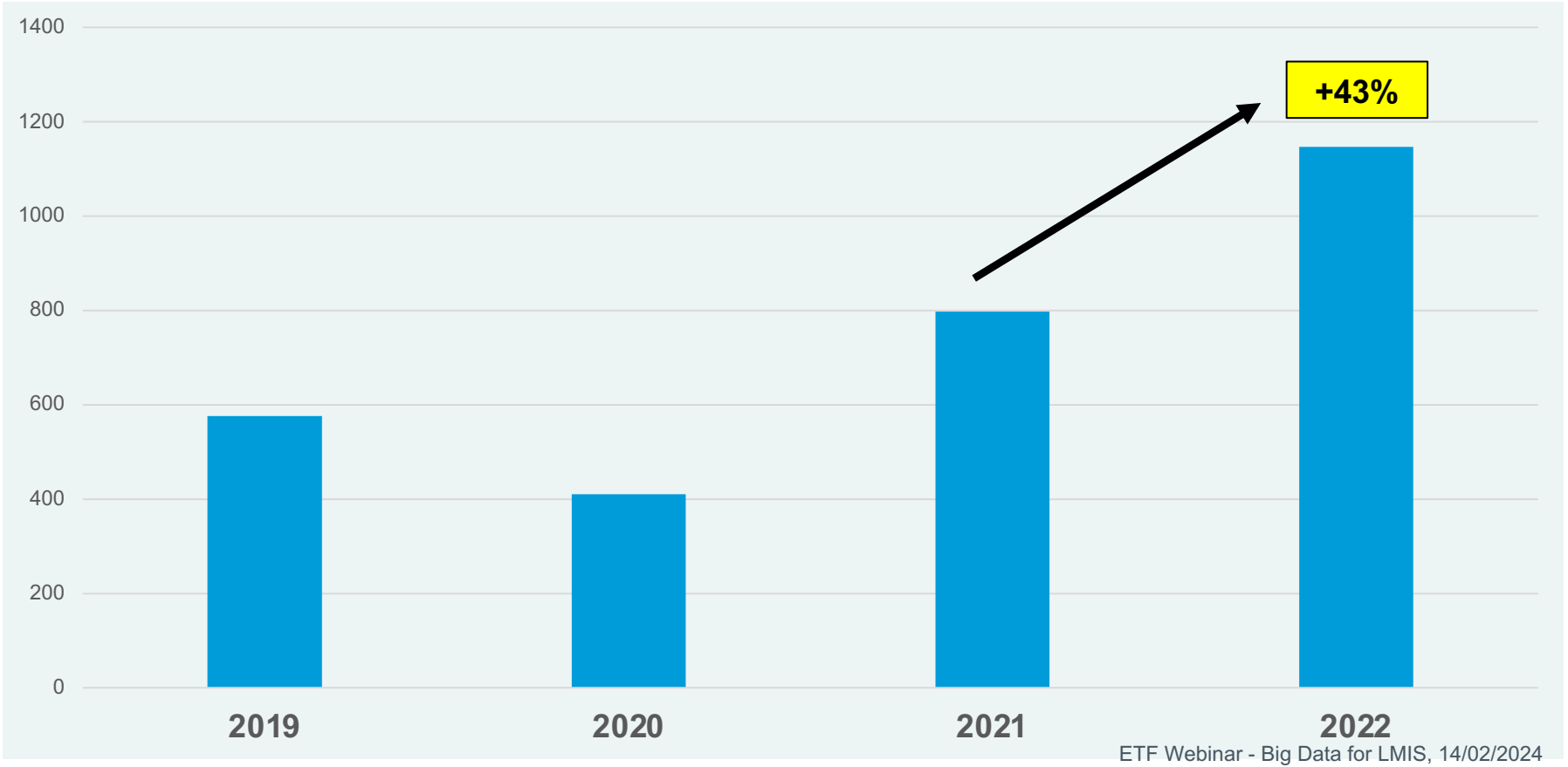
# Egypt: DIGITAL SKILLS: TOP 20





Source: ETF Job Postings – Egypt 2022

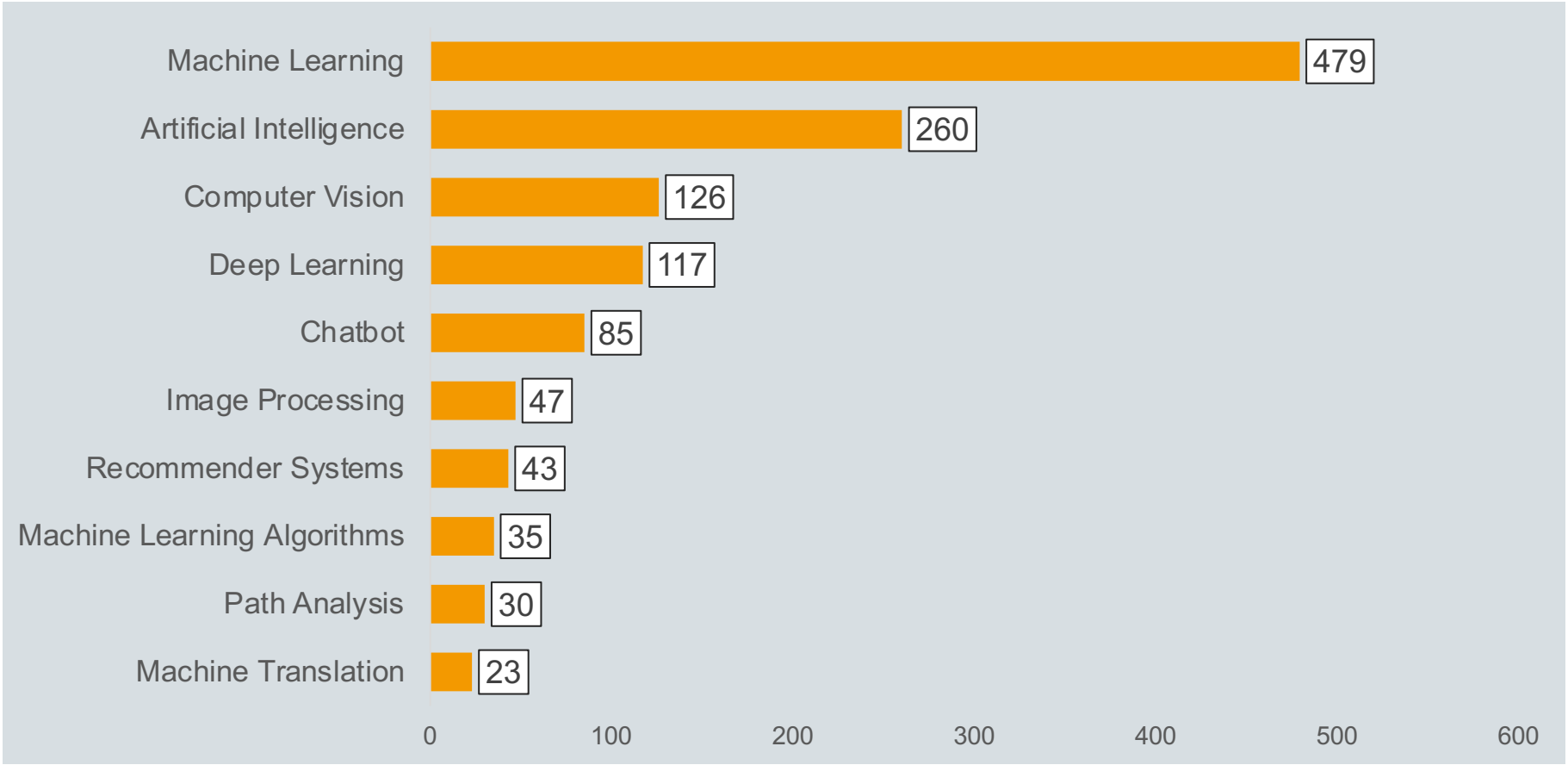
# AI SKILLS: UNIQUE JOB POSTINGS






Source: ETF Job Postings – Egypt – Lightcast AI Skills list

# AI SKILLS: TOP SKILLS IN OJV




Source: ETF Job Postings – Egypt – Lightcast AI Skills list

# DIGITAL SKILLS: IN DIGITAL AND NON-DIGITAL OCCUPATIONS

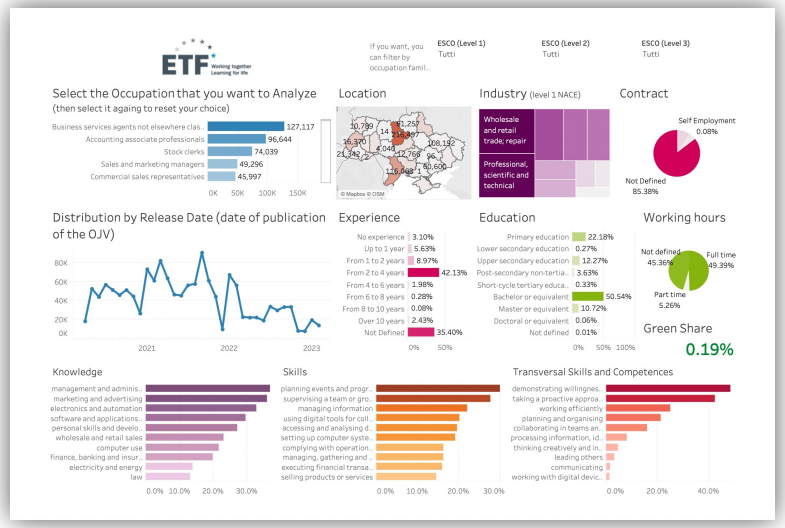
| Occupation (ISCO 08) | Unique OJV 2022 | Digital skills rate | Soft skills rate | Occupational-specific non digital skills rate |
|---|---|---|---|---|
| Database designers and administrators | 158 | 54,00% | 20,00% | 26,00% |
| Web and multimedia developers | 5418 | 48,00% | 23,00% | 29,00% |
| Systems administrators | 1366 | 47,00% | 25,00% | 28,00% |
| Applications programmers | 1343 | 44,00% | 24,00% | 32,00% |
| Software developers | 14752 | 40,95% | 26,32% | 32,73% |
| Data entry clerks | 351 | 40,13% | 30,64% | 29,23% |
| Mathematicians, actuaries and statisticians | 174 | 39,08% | 28,53% | 32,38% |
| Database and network professionals | 1780 | 38,77% | 26,57% | 34,66% |
| Computer network professionals | 397 | 38,29% | 29,08% | 32,63% |
| Systems analysts | 6593 | 28,64% | 33,62% | 37,74% |
| Graphic and multimedia designers | 3686 | 25,39% | 25,93% | 48,69% |
| General office clerks | 2022 | 21,92% | 44,19% | 33,89% |
| Environmental engineers | 58 | 19,53% | 36,15% | 44,31% |
| Business services and administration managers | 1785 | 15,78% | 40,38% | 43,84% |
| Personnel and careers professionals | 4122 | 15,77% | 39,97% | 44,26% |
| Accounting and bookkeeping clerks | 240 | 11,88% | 33,99% | 54,13% |
| Handicraft workers in textile, leather and related materials | 2022 | 19,75% | 31,85% | 48,41% |



Source: ETF Job Postings – Egypt 2022

# Big Data LMIS ETF project

## New initiatives using our database, other data sources and ESCO - 2024




- Occupational profiles – data driven, ESCO-based (ACQF-II – for common profiles of qualifications)
- Demand for green skills – data driven, ESCO-based - report
- Demand for digital skills - report
- Supply side – social profiles (ESCO Skills)
- Combination of data (OJV, conventional statistics, administrative)

# FINAL NOTES

## 1. Value of OJV data

- ***Near real-time data***
- ***Volume:*** allows different angles of analysis on skills and occupations; granularity
- ***Finder*****!** Early spotting of new skills / new patterns of skills mix (digital, soft, technical, green…). For further analysis in combination with conventional statistical and other data sources
- ***Green and digital transition***: Identification of demanded skills; quantitative trends over time – by occupations and sectors; green and digital skills profiles of occupations; AI skills and occupations; and setting up international benchmarking / comparisons.
  - Identification of **green jobs** in high demand; lead industries and regions in green job creation; qualifications for green jobs
  - The evolution of the **digital skills** and new emerging occupations; growth in jobs requiring new knowledge & skills.
- ***Value-added to LMI and skills development policies and practices***: emerging trends in LM and skills; inform policy responses;

## 2. Limitations

- Over-representation of certain occupational groups (professionals – 55%); underrepresentation (in general occupations requiring lower level of skills and qualifications).
- Classification of OJV data: robust techniques, constant evolution, taxonomies in transformation.

# Thank you

**European Training Foundation**

Eduarda Castel-Branco

ecb@etf.europa.eu

ETF Working together Learning for life

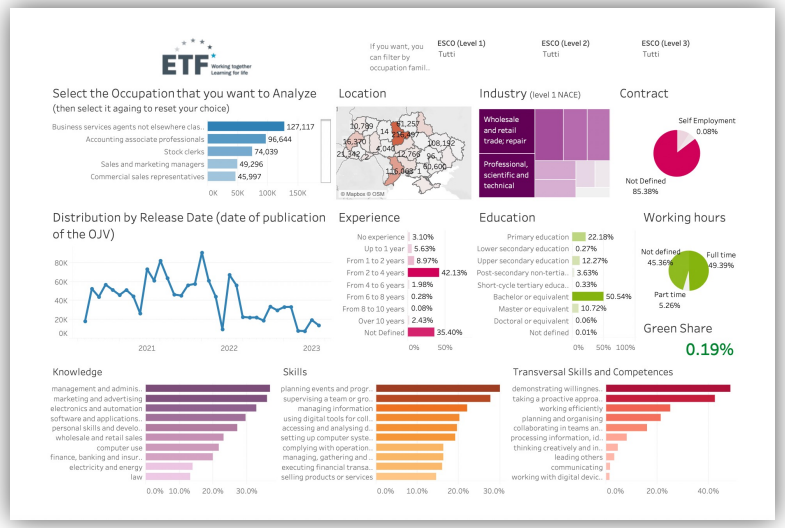
European Training Foundation

02

# ANNEX: for reference and further reading (not presentation)

# ETF project

# Big Data for LMI 2018-2024



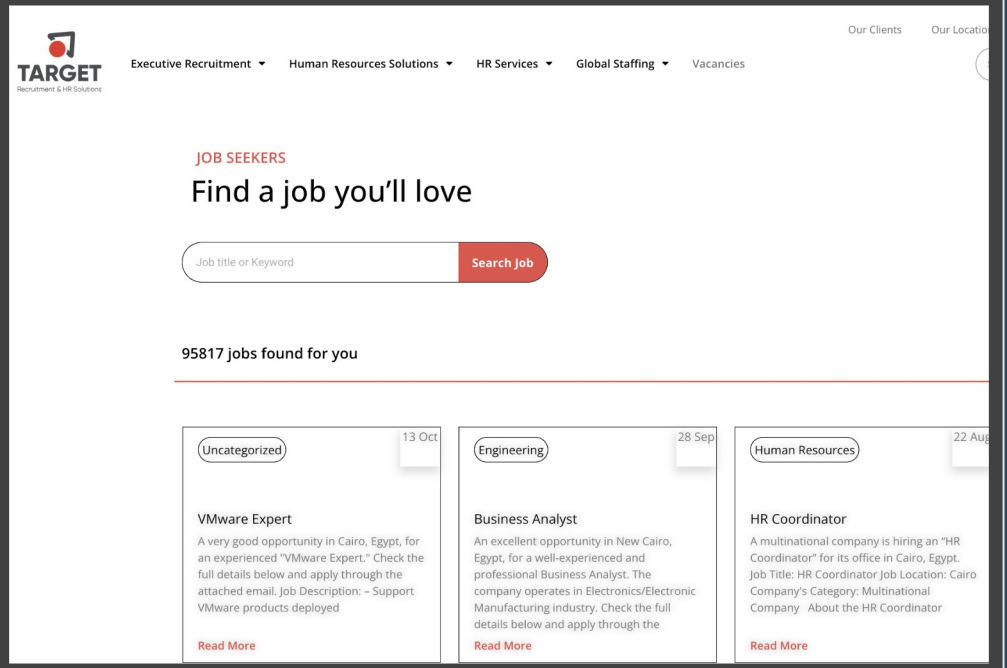

- **2018-2019: Methodology**: first step - brief methodological handbook "Big Data for labour market intelligence: an introductory guide" (published in **2019**).
- **2019: First application**: Feasibility analysis – Landscaping of Web Labour Markets Tunisia and Morocco
- **2019-2021:** 3 main training programmes for experts of the partner countries and other regions(Asia, Africa)
- **2020:** Creation of the complete OJV analysis system and dashboards: Tunisia and Ukraine
    - Analytical reports: LM and skills Ukraine and Tunisia
- **2021**:
    - New country – Georgia;
    - **Green dashboard 3 countries**
- The data system is based exclusively on **demand** – based on **job vacancies (OJV)** posted on web portals
- Full comparability with the Real-Time data system of the EU-27 (same methodology)
- ETF works with the data analytics specialists of **University Milano-Bicocca and LightCast**

**2022:** expansion new countries, new themes analysis: Egypt, Kenya; **2023**: Morocco. Selection – based on results of landscaping study.

New themes: Supply side analysis.

- Training and capacity development programme: materials, PPTs, videos – all accessible online for free use and sharing

# ONLINE JOB ADVERTISEMENTS



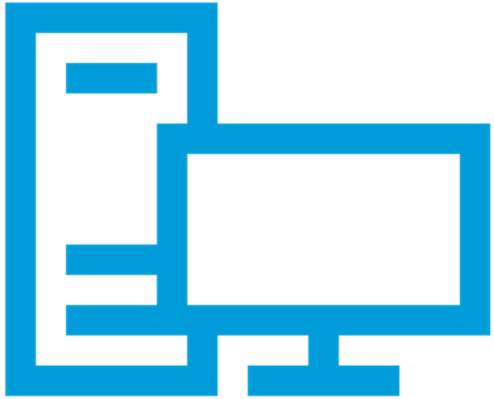

- Online job advertisements (OJAs) refer to advertisements published on the internet.
- Volume of OJAs is growing
- OJAs usually include data on the characteristics of the job (e.g. occupation and location), characteristics of the employer (e.g. economic activity) and requirements (e.g. education/skills).
- Part of this information is available only as natural language textual data.
- This type of big data requires specific methodologies for processing and analysis but also provides much more detailed information (compared to alternative data sources) and avoids pre-conceived classifications (important to identify emerging skills).

# BIG DATA LMI: FOCUS ON OJV KEY FEATURES

- Data collected from the web, based on feasibility analysis (landscaping)
- 5 Vs of Big Data
- Data management flow: landscaping, ingestion, processing, extraction, ETL, presentation
- Several data quality steps and tools
- Data classification: a) AI-aided – with human in the loop; b) International classifications & taxonomies
- ISCED, ISCO, ESCO, NACE; CEN-CENELEC, Stackoverflow, GitHub; ETF green skills
- Big Data: complements conventional statistics ("*Trusted Smart Statistics*") – points topics and issues for further / wider analysis

ETF Working together Learning for life
European Training Foundation

ACQF



# Digital skills

Ability to use digital technologies confidently, critically and responsibly and engage with them for learning, at work and for participation in society.

It includes information and data literacy, communication and collaboration, media literacy, digital content creation (including coding), security (including digital wellbeing and cybersecurity-related skills), intellectual property issues, problem-solving and critical thinking.

# Definitions

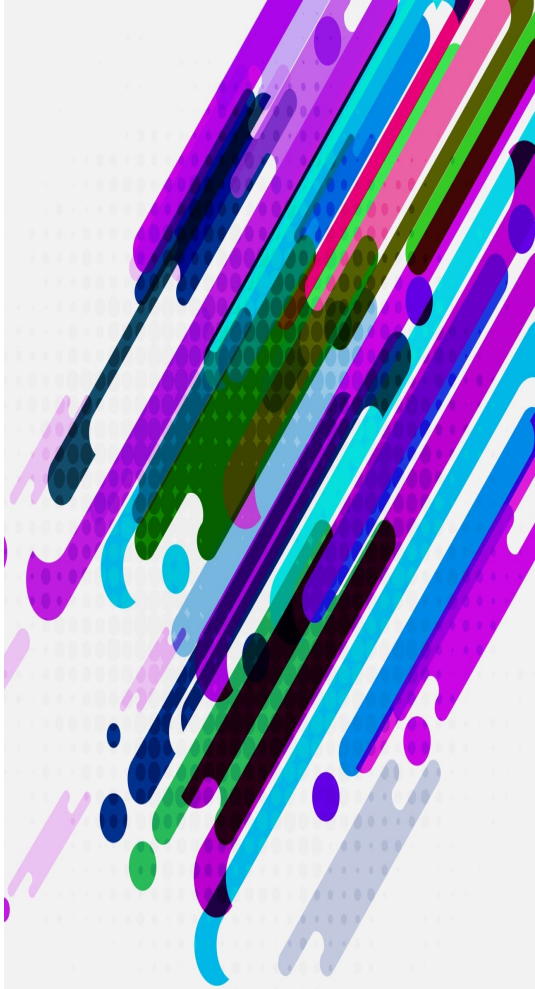
**Green skills – an important area of debate and policy action and which has become a priority in research (quantitative and qualitative) and social communication. Several international organisations are working on the analysis and taxonomies related to green competences.**

**Cedefop defines** green skills as "the knowledge, skills, values and attitudes needed to live, work and act in economies and societies that seek to reduce the impact of human activity on the environment".

**Skills for the green economy consist of:**

- **transversal skills, linked to sustainable thinking and acting, relevant to all economic sectors and professions;**
- **specific skills needed to adapt or implement standards, processes and services to protect ecosystems and biodiversity and reduce energy, materials and water consumption;**
- **highly specialised skills needed to develop and implement green technologies such as renewable energy, wastewater treatment or recycling;**

**Skills for the green economy are also referred to as skills for green jobs, skills for the green transition or green skills.**



# Taxonomy

- ETF green skills list
- 225 terms (data-driven identification) – now integrated in ESCO
- 7 clusters of sustainable activity (inspired by the European Green Deal)

# THANK YOU

**European Training Foundation**

Eduarda Castel-Branco

ecb@etf.europa.eu