Big Data Analytics for labour market intelligence

Fabio Mercorio
University of Milano-Bicocca

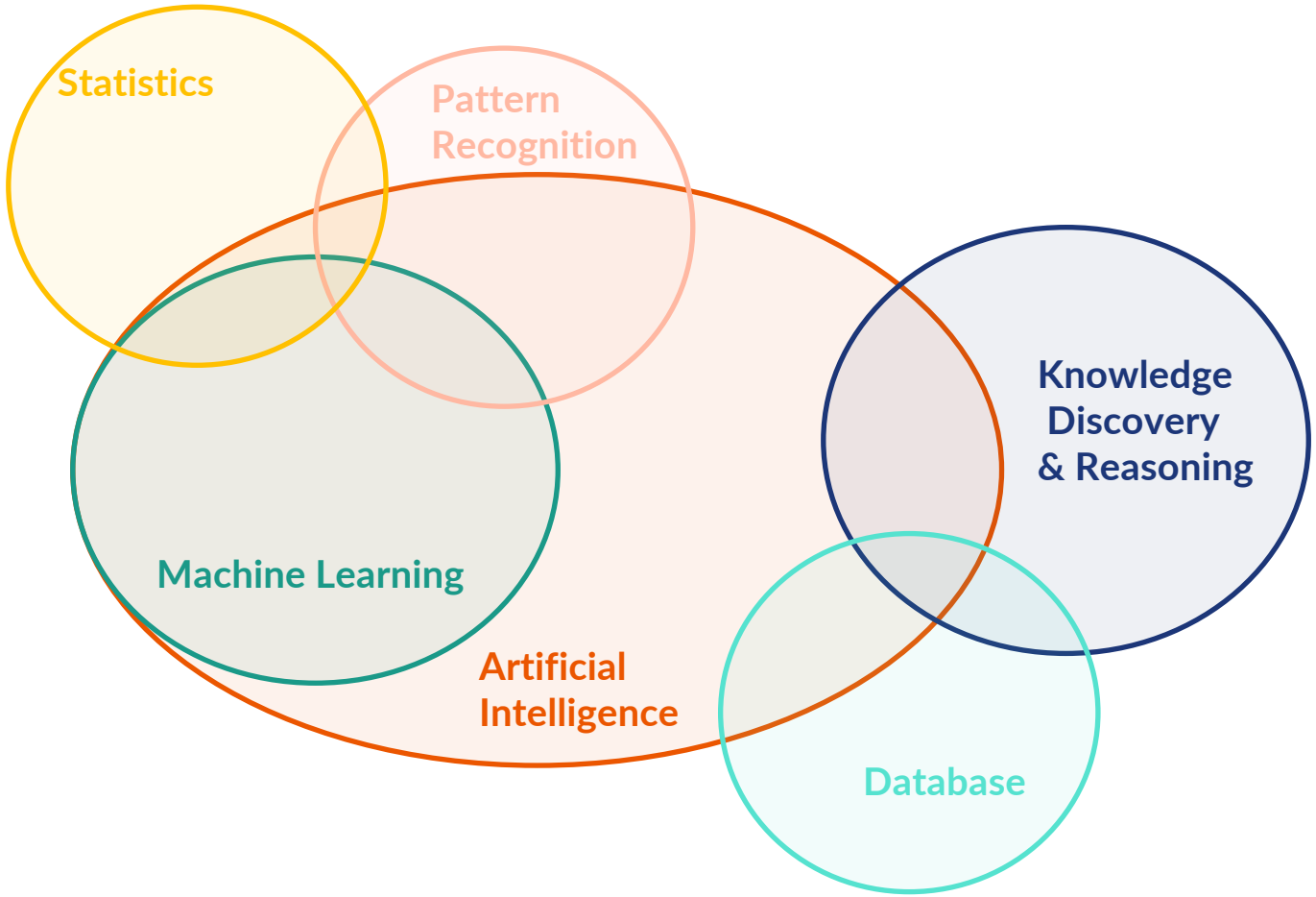
interuniversity research centre for public services

crispresearch.eu

Big Data is the fuel of AI

Artificial intelligence (AI) refers to **systems** that **show** intelligent behaviour: by analysing their **environment** they can perform various tasks with some degree of autonomy to achieve **specific** goals.

European Commission. AI for Europe (2018)

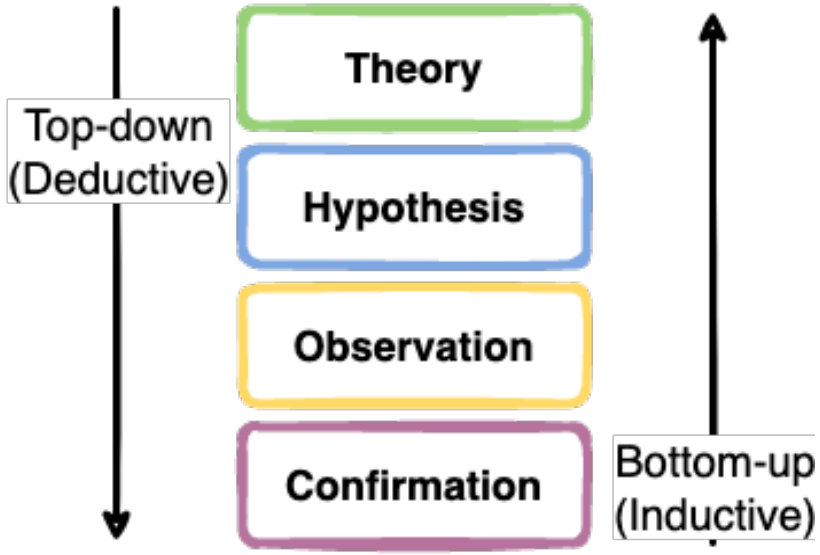


A paradigm shift

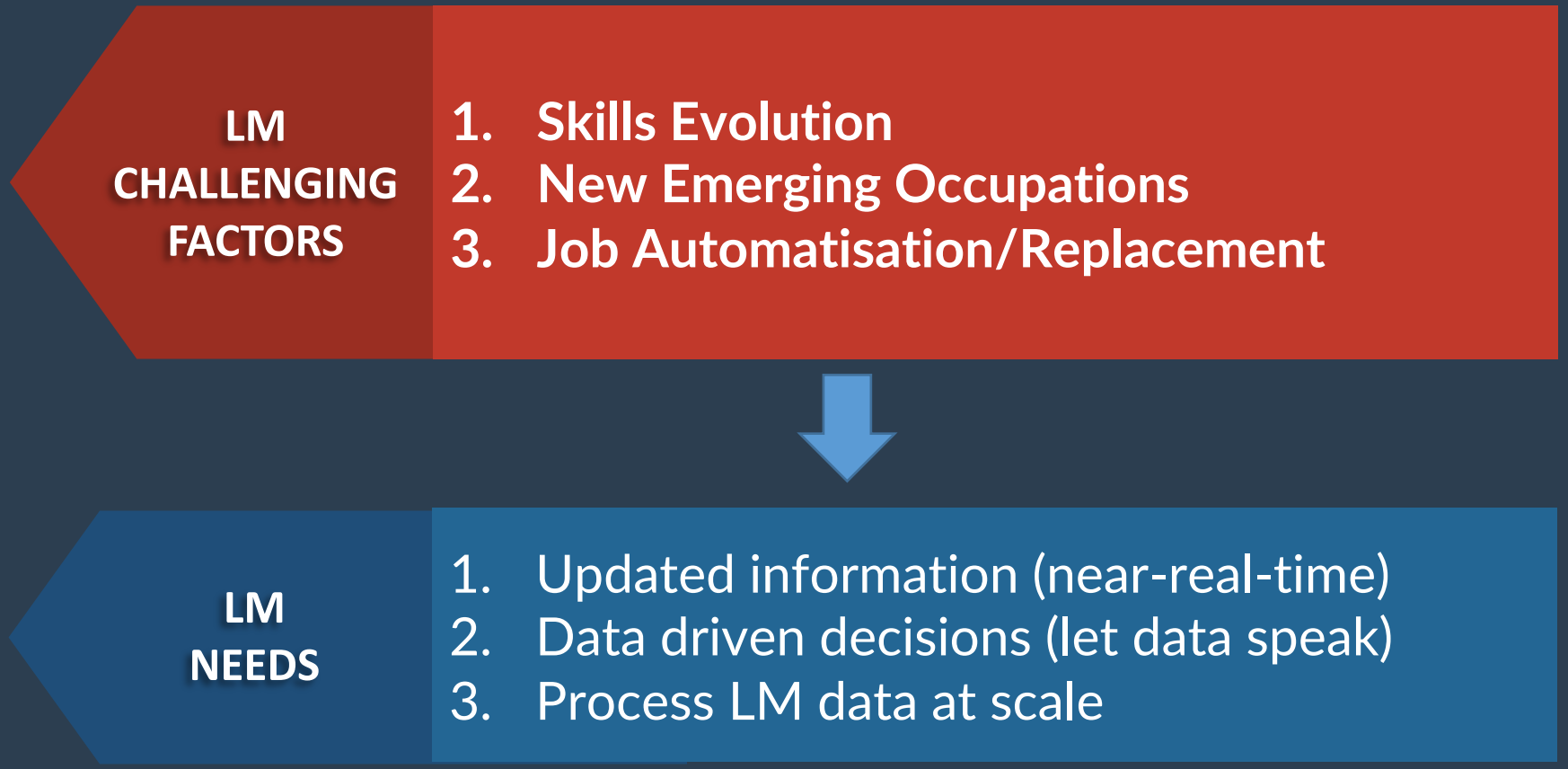
Big Data: large-scale data, usually having a varied and complex structure. They
Big Data Analytics (BDA): processing big data and looking for valuable information, correlations and patterns within them.

The use of BDA to support decision making requires a paradigm shift (not a replacement!) in the way decisions are made, moving from deductive reasoning to inductive reasoning.

Both deductive and inductive approach are needed!



Labour Market Challenging Factors



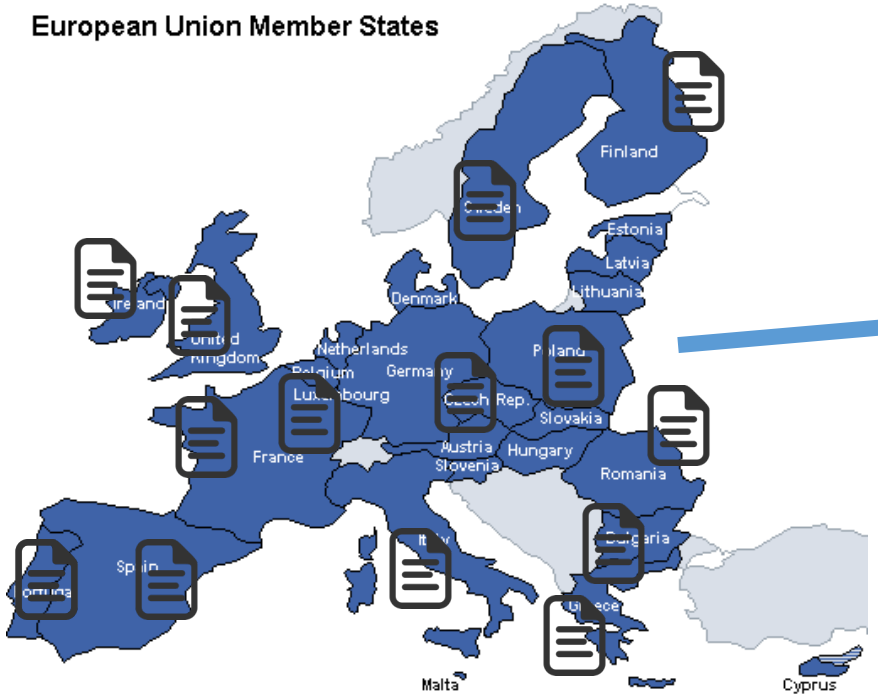
Labour Market Intelligence (LMI): Design, define and implement machine-based frameworks and algorithms to extreact knowledge from labour market information

Ongoing Research Projects on AI& Big Data for LMI

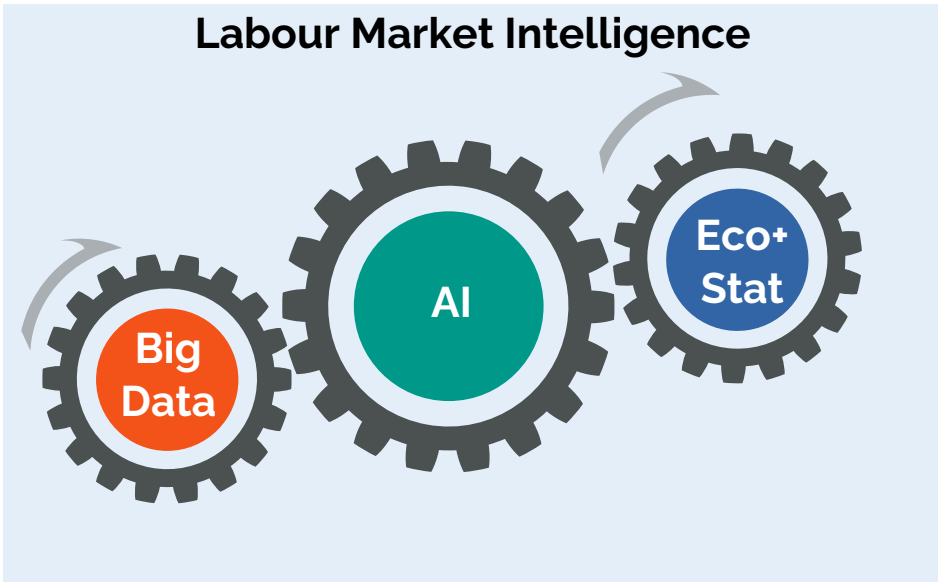
- (2021-2024) Eurostat+Cedefop **[Putting OJA into Statistics & novel research studies]**. 27+1 EU Countries, 32 languages supported.
- (2021-2023) H2020 **[PILLARS: emerging tech and LM transformation]**
- (2022-2026) European Training Foundation ETF **[Human Capital Developement Expertise Services]**
- (2021-2023) Cedefop **[Short-term anticipation of skills trends and VET demand]**
- (2022-2024) LMI-EUniv **[Innovating the use of Labour Market Intelligence within European Universities]**. Erasmus+
- (2021-2022) ACSOL **[Acquiring crisis-proof skills through online learning].** Erasmus+
- (2021-2022) E-MLSR **[European Mid Life Skills Review]**. Erasmus+

How can AI monitor the labour market demand from the web?

EUROPEAN Real-Time Labour Market Monitor



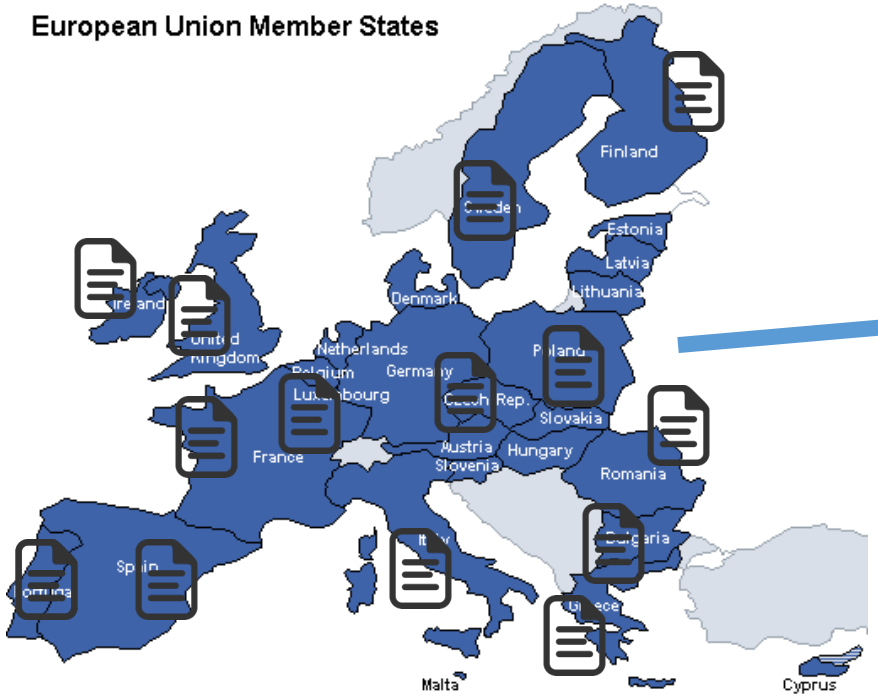
(2016-2020)

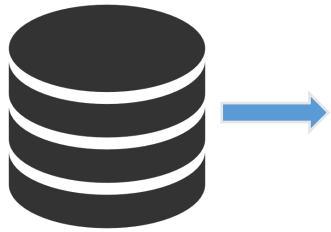


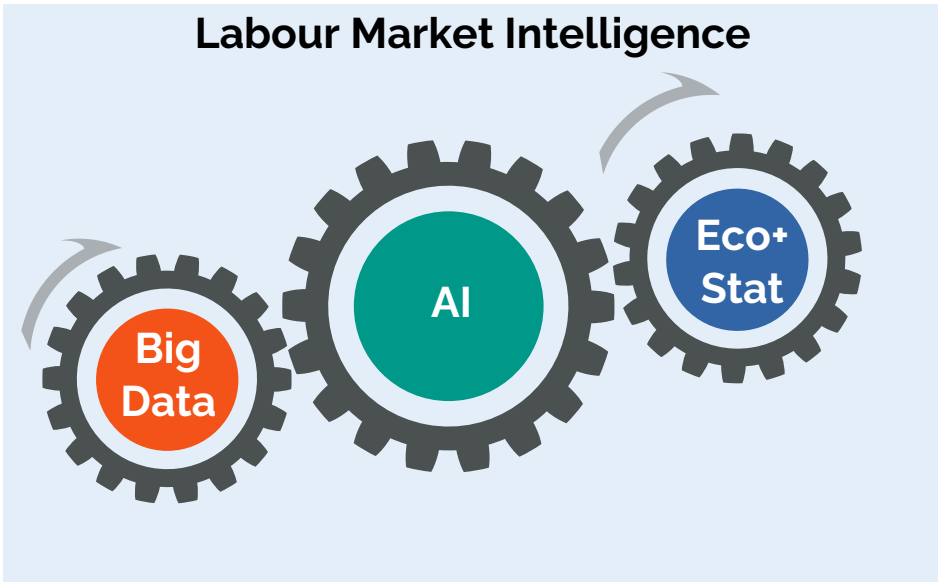
CEDEFOP
European Centre
for the Development
of Vocational Training

27+1 EU Countries – 28 Languages – more than 4M unique vacancies per month

EUROPEAN Real-Time Labour Market Monitor







(2021-2024)

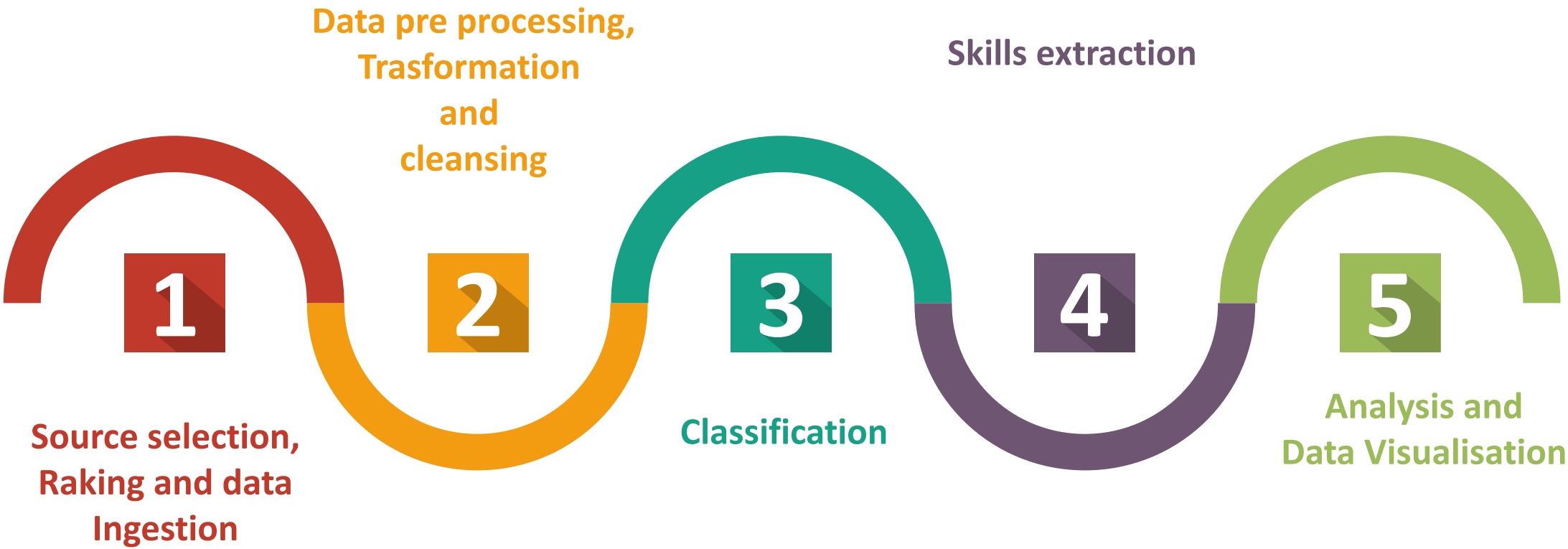
32 Countries 28 Languages

CEDEFOP
European Centre for the Development of Vocational Training

eurostat

Putting job ads into official statistics for the whole Europe

The Process


1
Source selection, Raking and data Ingestion
2
Data pre processing, Trasformation and cleansing
3
Classification
4
Skills extraction
5
Analysis and Data Visualisation

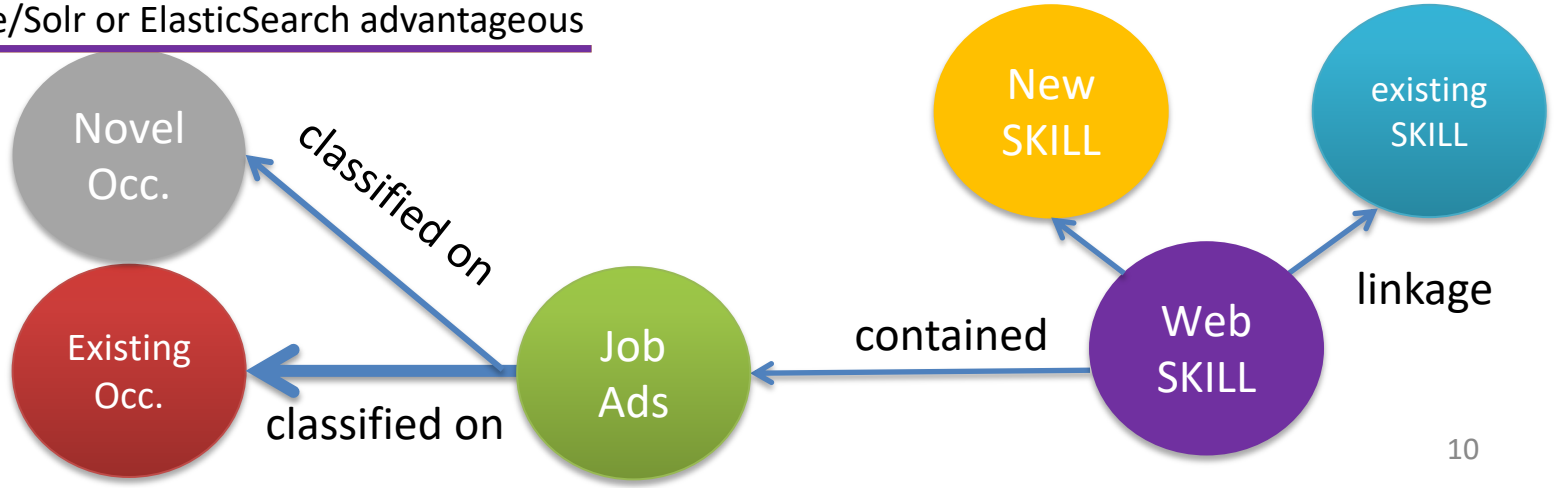
Online Job Ads example

Job Title: Data Scientist.

Description: We're looking for a talented Computer Scientist to join our growing development team. Your expertise in data will help us take this to the next level. You will be responsible for identifying opportunities to further improve how we connect recruiters with jobseekers, and designing and implementing solutions. […] Required skills and experience:

- SQL and relational databases;
- Data analysis with R (or Matlab);
- Processing large data sets with MapReduce and Hadoop);
- Real time analytics with Spark, Storm or similar;
- Machine Learning;
- Natural Language Processing (NLP) and text mining;
- Development in C++, Python, Perl;
- Experience with search engines e.g. Lucene/Solr or ElasticSearch advantageous

1. **Data driven approach**
2. **High granularity**
3. **High frequency (updated)**
4. **Focus on important skills**



Eurostat – Cedefop research activities (16-ongoing)

Goal:

1. Set-up real-time LMI system from 5 EU countries up to 27+1 (and 4 EU border countries)
2. Extract value from OJAs to support decision makers in understanding labour market dynamics, trends, and to put big data into official statistics

Activities (selection)

- Identify novel occupations and skills on the basis of skill variation;
- Estimate the representativeness, stability, and coverage of online job ads;
- Study indicator to estimate the impact of green economy within occupations
- Estimate the relevance of digital/non-digital/soft skills within jobs and among sectors/territory/etc…
- …

Come cambia la domanda di skill nelle professionalità

Area professionale

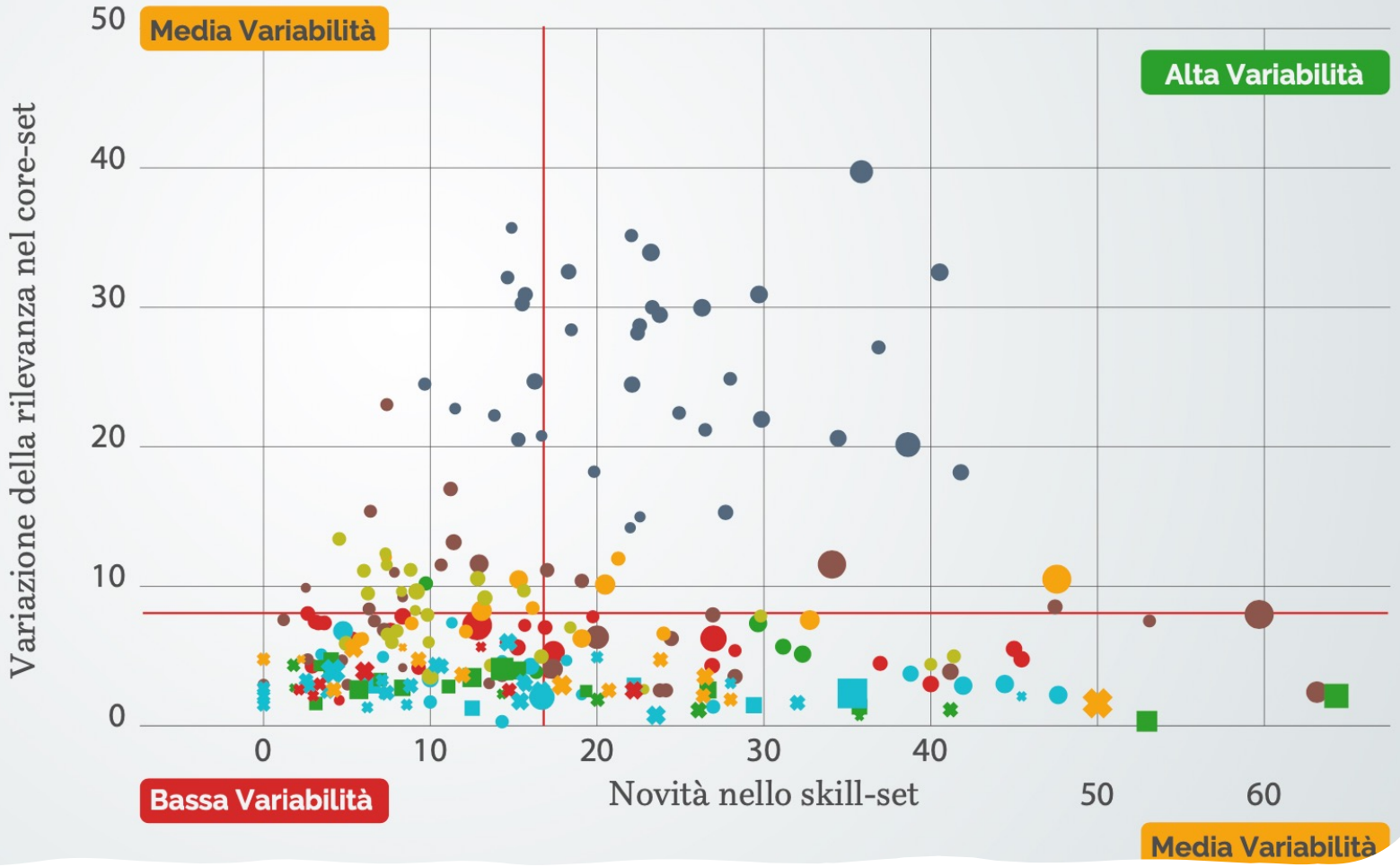
- Ciclo produttivo beni
- Amministrazione, contabilità segreteria e gestione del personale
- Progettazione, ricerca e sviluppo
- Marketing, assistenza clienti e vendita
- Ciclo erogazione servizi
- Sistemi informativi
- Direzione generale

Variazione tendenziale 15 vs 19

- -200
- 0
- 200
- 400
- 600

Livello della professione

- High level
- Medium level
- Low level



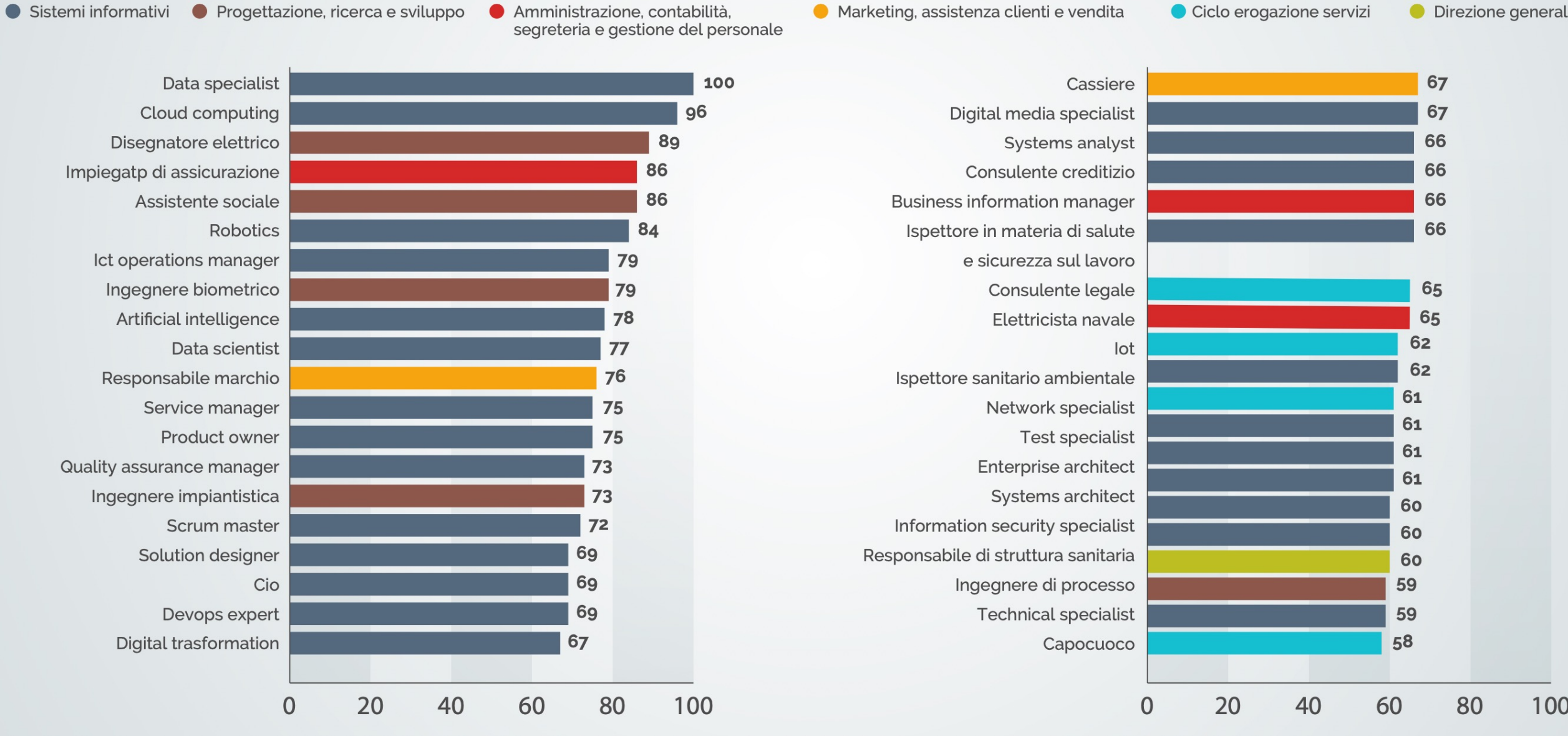
Come cambia la domanda di skill nelle professioni ICT – CEN*

Periodo di analisi 2015-2020

* The European Committee for Standardization

Indice di cambiamento

Graduatoria delle top-40 professioni per indice di cambiamento



Concluding Remarks

1. Data representativeness as a challenging issue to be addressed
2. Towards Trusted Smart Statistics:

   To build trust into smart statistics the data life-cycle needs to be auditable, transparent, with guarantees of accuracy and privacy by design
3. It enables third-party research activities
   - Identify novel occupations and skills on the basis of skill variation;
   - Estimate the representativeness, stability, and coverage of online job ads;
   - Study indicator to estimate the impact of green economy within occupations
   - Estimate the relevance of digital/non-digital/soft skills within jobs and among sectors/territory/etc
   - …

Thank you