Capacity building and data innovation in the ESS

International Conference on Big Data for Official Statistics
Data Science and Capacity Development in Official Statistics
Kigali (Rwanda), 2–3/5/2019

Contents

- *Study on training strategy for big data*
- *Big Data test infrastructure*
- *European Statistics Training Programme*
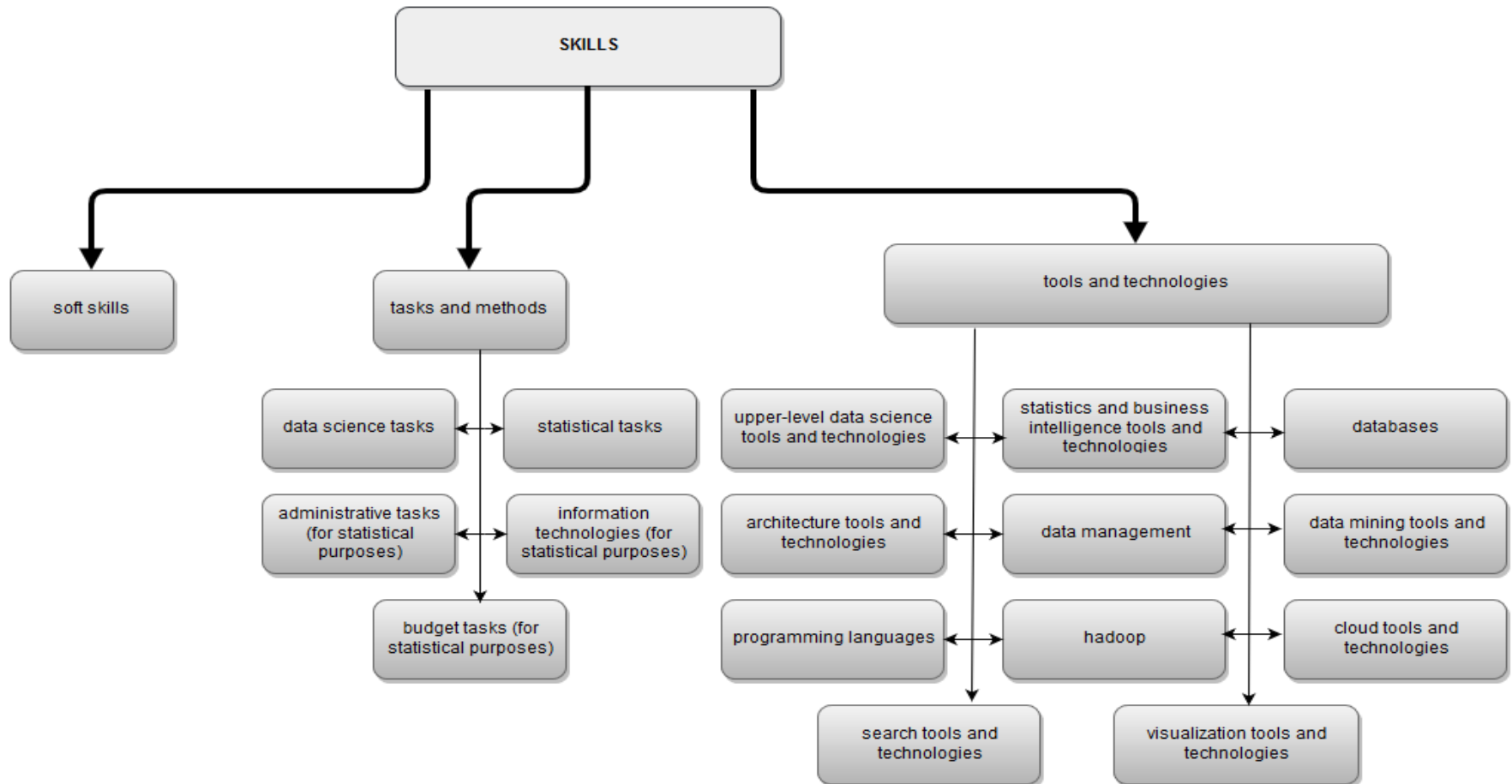- *European Master in Official Statistics*

Never stop learning, because life never stops teaching

European Commission

Training Strategy for Big Data

To bridge the Big Data skills gap in European official statistics

- Identification of skills required for the use of big data sources
  - Skills framework

- Inventory of existing skills in Eurostat and in the NSIs in Europe
  - Questionnaire to NSIs and Eurostat

- Analysis of the big data training needs

- To define the training objectives and content
  - Competency-Based Education approach

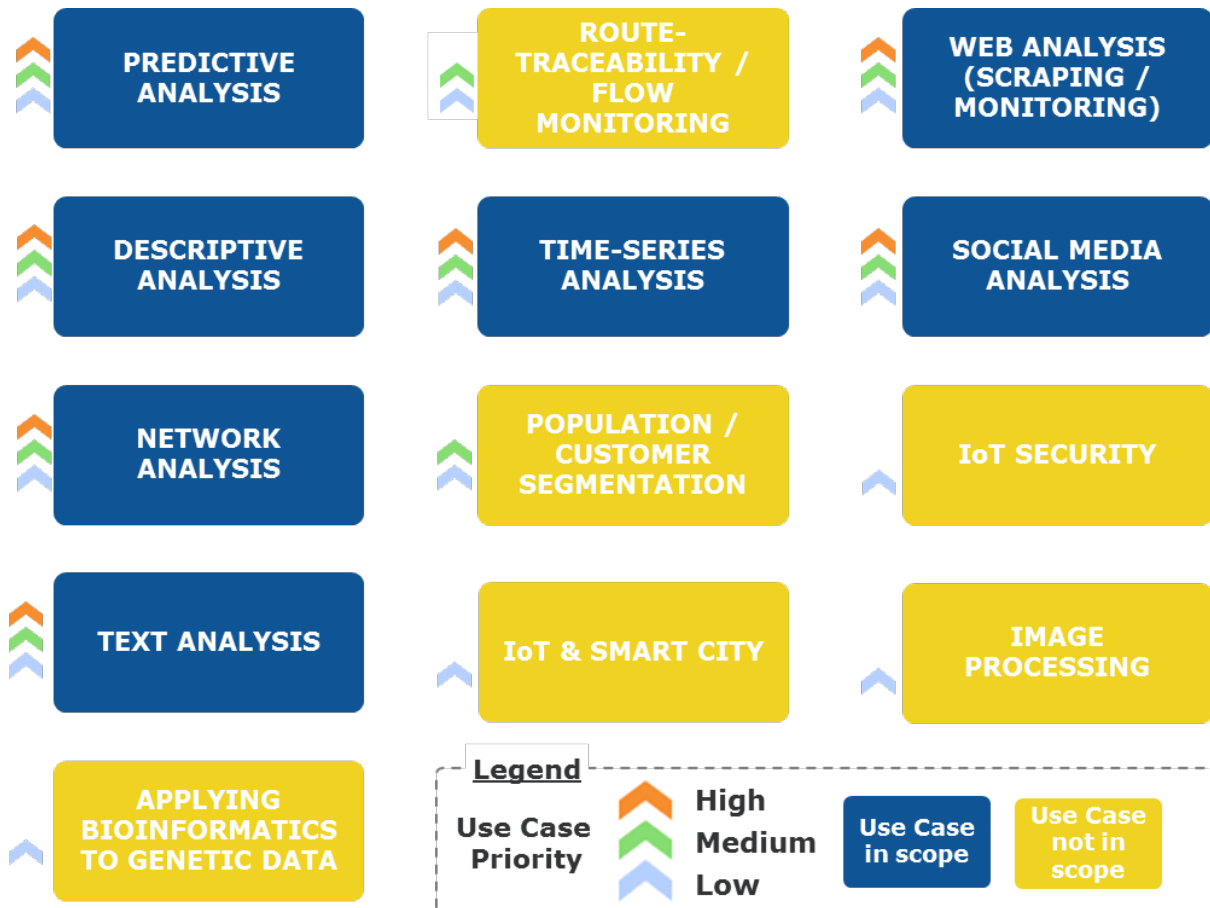- Develop a training provision strategy to bridge the skill gap using different instruments

European Commission

Skills classification by groups of skills

Intended outcomes of strategic training plan

Time	Change in	Effects
Short	• Knowledge • Skills • Attitude • Motivation • Awareness	• Awareness raising • Change of attitude • Skills acquisition
Medium	• Behaviors • Practices • Policies • Procedure	• Constant support of level of expertise • Change of behavior and practices using of big data and other emerging technologies • Adoption of new standards, policies and procedures
Long	• Situation – environment, social and economic conditions	• Knowledge and new technologies will drive the statistical production

The Big Data Test Infrastructure

PREDICTIVE ANALYSIS

ROUTE-TRACEABILITY / FLOW MONITORING

WEB ANALYSIS (SCRAPING / MONITORING)

DESCRIPTIVE ANALYSIS

TIME-SERIES ANALYSIS

SOCIAL MEDIA ANALYSIS

NETWORK ANALYSIS

POPULATION / CUSTOMER SEGMENTATION

IoT SECURITY

TEXT ANALYSIS

IoT & SMART CITY

IMAGE PROCESSING

APPLYING BIOINFORMATICS TO GENETIC DATA

Big Data use cases in scope

Legend

Use Case Priority

High
Medium
Low

Use Case in scope

Use Case not in scope

European Commission

The Big Data Test Infrastructure

PaaS for implementing Big Data Use Cases

Community Building and Innovation Portal

Big Data and Analytics software catalogue

Advisory

Data Catalogue and Data Exchange APIs

Support for Analytics implementation

Analytics as a Service

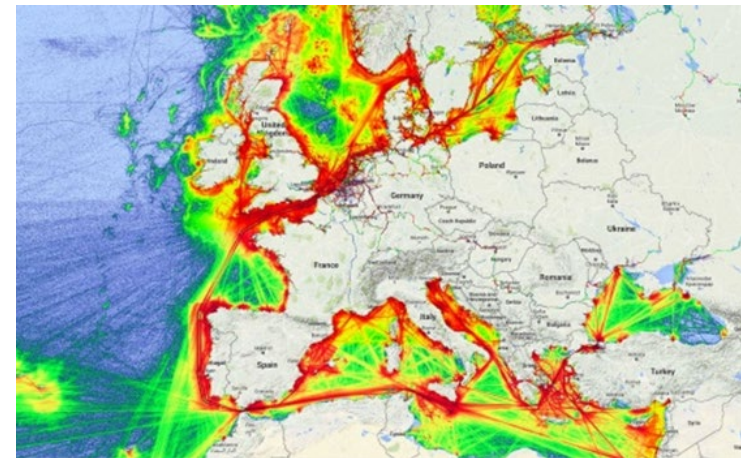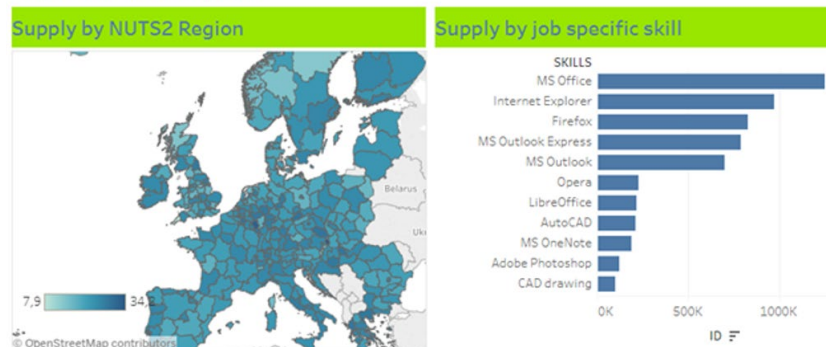**Big Data Test Infrastructure service offering**

European Commission

BDTI and the ESS

BDTI sandbox test with Wikipedia page views
First pilot: European Big Data Hackathon 2019
ESSnet big data pilots

- **Online Job Vacancies**
- **AIS**
- **Mobile phone**

Matching Skills Demand (CEDEFOP JVs) and Supply (EURES CVs):
*Absolute and relative frequencies*

Supply by NUTS2 Region

Supply by job specific skill

SKILLS

MS Office
Internet Explorer
Firefox
MS Outlook Express
MS Outlook
Opera
LibreOffice
AutoCAD
MS OneNote
Adobe Photoshop
CAD drawing

0K   500K   1000K

ID

7,9   34,

© OpenStreetMap contributors

European Statistics Training Program (ESTP)

- *The **purpose** of the ESTP is to provide European statisticians with continuous training in new methods, techniques and best practices and integrate the application of European concepts and definitions*

- *The programme is tailored to **meet the specific needs** of the European Statistical System (**ESS**)*

- *The programme is **managed by Eurostat**, the Statistical Office of the European Union*

- *The ESTP **complements national training schemes** and meets the challenges of comparable statistics at European and international level*

- *Training **delivered** by external contractors or staff members*

- ***Cost** for delivering training ca. 500.000 € over four years*
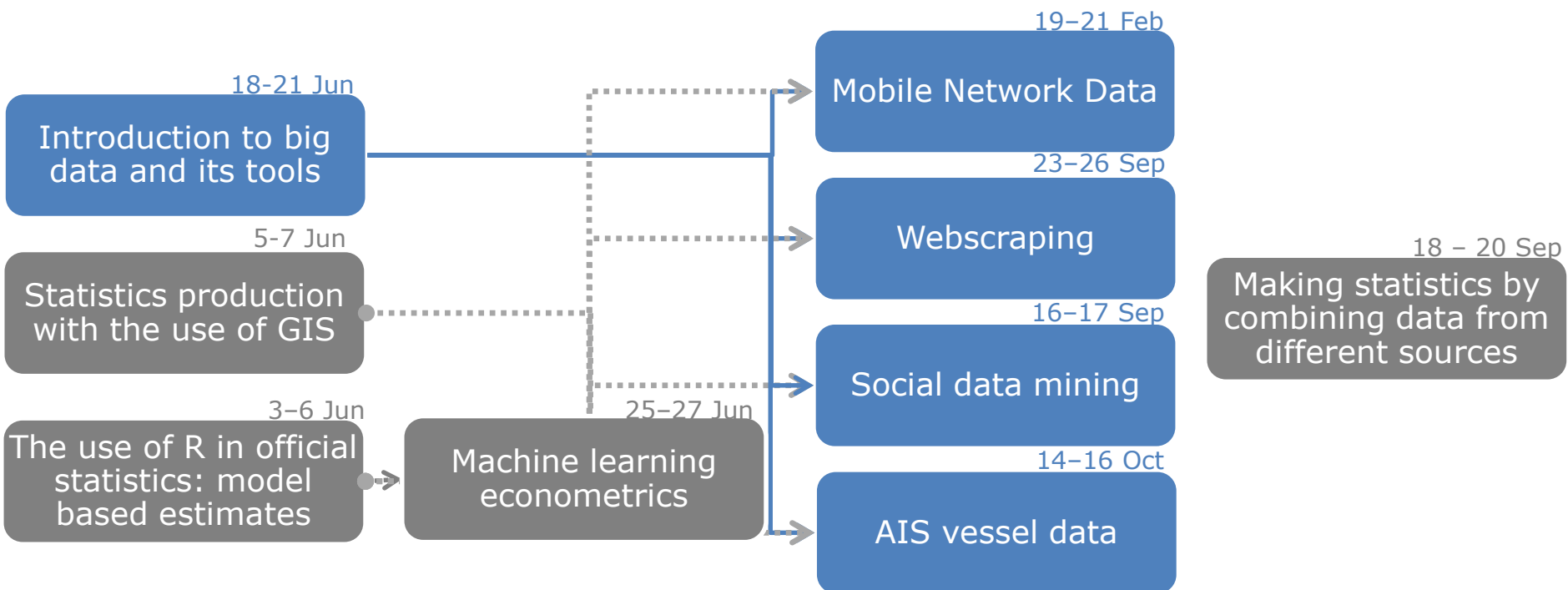
ESTP - Target audience

Staff of NSI/ONAs in EU Member States, EFTA countries, candidate countries and Eurostat

Applicants from non-European countries may be admitted on individual basis and subject to availability of places
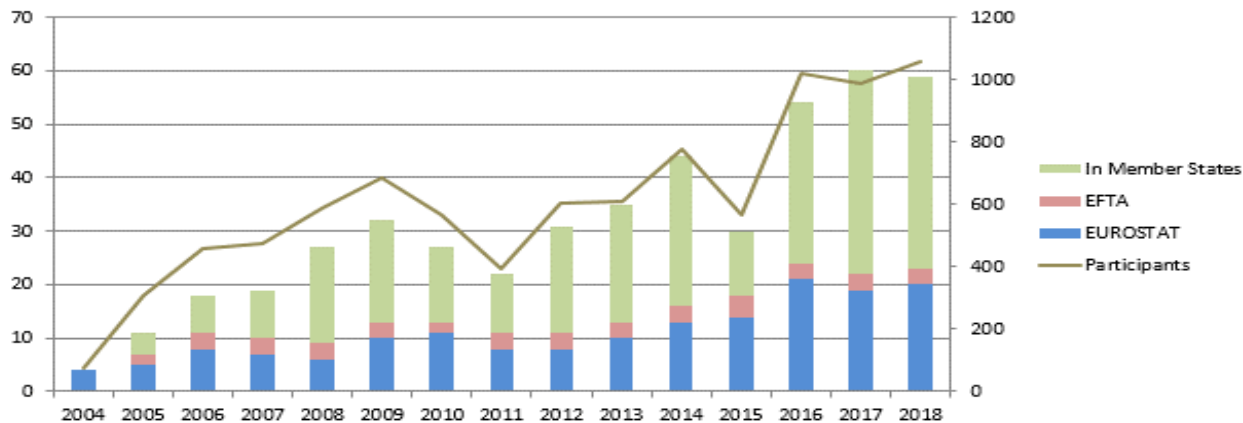
ESTP Programme 2019

- *Statistical Methodology*
- *Economic Statistics*
- *Environmental Statistics*
- *Metadata and Classifications*
- *Quality and Dissemination*
- ***New Data Sources***
- *Business Statistics*
- *Social Statistics*

ESTP courses supporting big data (2019)

19–21 Feb
Mobile Network Data

18-21 Jun
Introduction to big data and its tools

5-7 Jun
Statistics production with the use of GIS

23–26 Sep
Webscraping

18 – 20 Sep
Making statistics by combining data from different sources

3–6 Jun
The use of R in official statistics: model based estimates

25–27 Jun
Machine learning econometrics

16–17 Sep
Social data mining

14–16 Oct
AIS vessel data

Big data courses   **Methodology courses**

European Commission

ESTP more information

ESTP National Contact Points

ESS Website

http://ec.europa.eu/eurostat/web/ess/about-us/estp

CIRCABC

https://circabc.europa.eu/w/browse/6ade1ca8-6a06-44bd-bff0-498217d0ec05

European Commission

European Master in Official Statistics (EMOS)

- *Label awarded by ESS Committee to Master programmes fulfilling EMOS requirements– valid 4 years*

- *Coordinated by Eurostat, universities deliver on learning outcomes and mandatory internships with statistical institutes*

- *Current challenges: recruitment, training & development, long-term retention of employees.*

- *EMOS graduates – both new students and existing NSI staff - better prepared for work in statistical offices*

EMOS - Label based on learning outcomes

1. System of official statistics	• Official statistics: relevance and sources • International and European institutions • European Statistics Code of Practice
2. Production models and methods	• Different kinds of data sources! • Quality dimensions
3. Specific themes	• Economy & finance, population & social conditions, international trade, environment & energy, science & technology …
4. Statistical methods	• Statistical methods: sampling & estimation, index theory, econometrics … • Statistical programmes
5. Dissemination	• Effective communication with different audiences • Confidentiality and disclosure control

EMOS- 33 programmes in 19 countries

16

European Commission

EMOS
Focus of 3ʳᵈ call 2018-2019

- *Production of statistics by **integrating data** from different kinds of data sources*
- ***Model-based** and **algorithm-based estimation** techniques*
- *Understanding of **institutional environment***
- *Dissemination – **visualisation techniques***
- ***Communication skills***