

Task Team on Training, Competencies and Capacity Development

UN Big Data Global Working Group

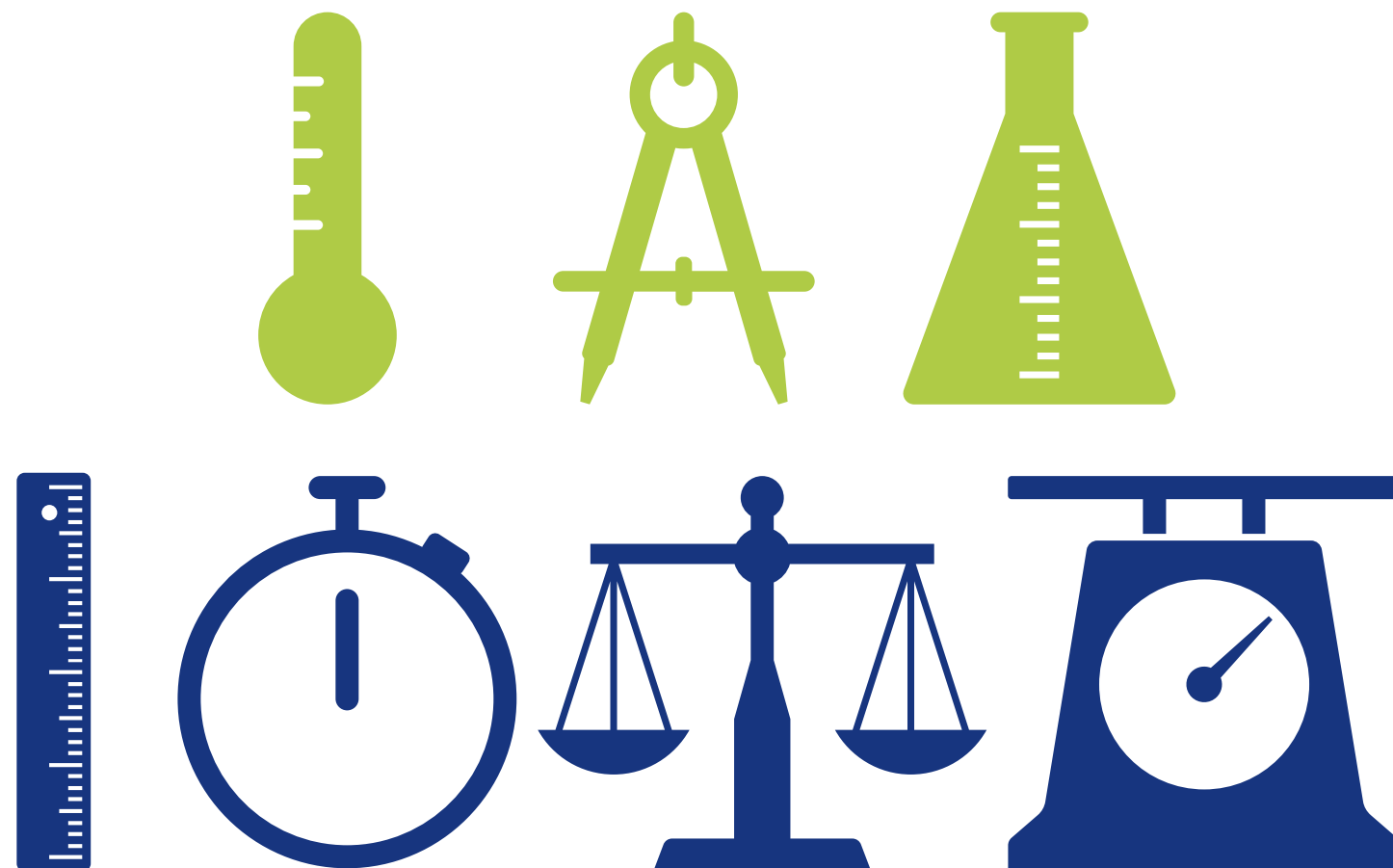
Dominik Rozkrut

Introduction

Official statistics needs to figure out how it should function in the modern world

- A need for better technical infrastructure
- Better skills
- Better public awareness of risks of misinformation
- The broader approach to cybersecurity, i.e. information security
- Better public awareness of the value of official statistics
- Involvement in cooperation at the international, supranational and global levels is becoming more and more important

Official statistics is about measurement



The only things that changes ...



... is the toolset



PostgreSQL



Statistics Poland the winner of the EU Big Data Hackathon organised by Eurostat



Open data... inside and outside



Portal
API GUS

Portal Informacyjny

English



A

A

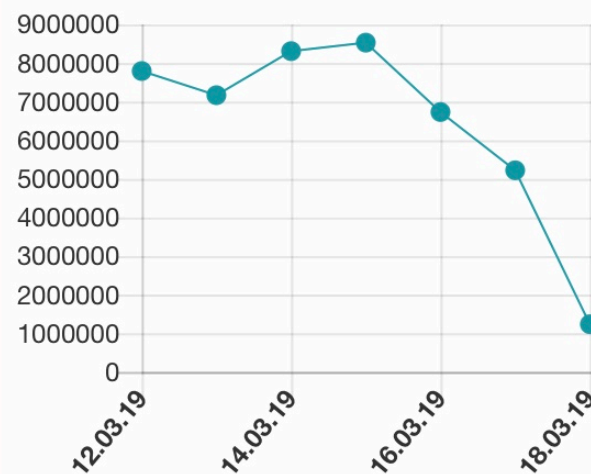
A

API REGON

3 767 514 252
zapytań od 01.01.2015

3518 zarejestrowanych użytkowników

Dzienne wykorzystanie usługi - ostatnie 7 dni

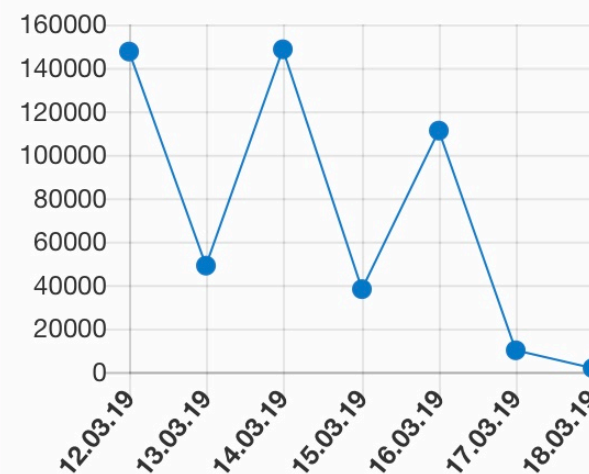


API TERYT

61 875 167
zapytań od 01.01.2015

512 zarejestrowanych użytkowników

Dzienne wykorzystanie usługi - ostatnie 7 dni

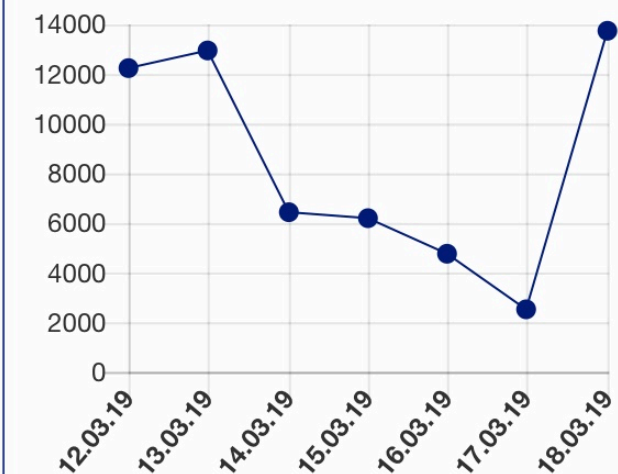


API BDL

851 934
zapytań od 04.12.2018

484 zarejestrowanych użytkowników

Dzienne wykorzystanie usługi - ostatnie 7 dni



Strategy forward

- We must be (even) more goal-oriented
- We must look for employees not for a specified range of activities, but based on specific hard and soft skills
- We need to rethink how
 - to attract data science talent
 - to train our own data science specialists
 - to educate our stakeholders

Task Team on Training, Competencies and Capacity Development

Objectives

- A. Develop methods and tools (including coordination and facilitation) for
 - a baseline need identification of “Big Data” competencies in NSSs, and
 - an assessment of institutional readiness of NSSs in using Big Data
- B. Build a competency framework for the Big data acquisition and processing in the current data landscape of NSSs

Objectives

- C. Identify and analyze the existing supply of training of data scientists
 - in academic and other centers
- D. Facilitate establishing a global network of institutions for training and capacity building on Big Data in official statistics,
 - under the umbrella of GWG
- E. Review of experiences and lessons learned from different organizational structures and data science governance
 - across national statistical systems

Objectives

- F. Provide guidance on the development of a training program
 - that addresses the gaps and mismatches
 - identified in the competencies analysis
- G. Coordinate training plans in line with
 - the Global Survey on Big Data for Official Statistics and
 - other outcomes from tasks teams formulated within the GWG
 - or other relevant stakeholders

Deliverables and Time Schedule

- 1. Assessment of institutional readiness (2019).
 - Output: Report of the skills assessment survey and data science governance across NSSs.
 - Method for needs identification and analysis of Big Data skills.
 - Assessment framework for national institutional 'readiness' in terms of governance and infrastructure (as part of Global Survey on Big Data for Official Statistics).

Deliverables and Time Schedule

- 2. Competency framework for new data sources in official statistics (2019).
 - Output: Report with the proposition of the competency framework.
 - Identification and definition of knowledge, skills and attitudes for work with new data sources.

Deliverables and Time Schedule

- 3. Analysis of the supply and profiles of specialists in areas related to the analysis of new data sources, such as Big Data (2019).
 - Output: Report on the training market for data scientists.
 - Identification and analysis of data scientists training programs in academic institutions and other relevant centers.

Deliverables and Time Schedule

- 4. Development of training courses (2019-2020).
 - Output: Guidelines on training program; technical seminar conducted.
 - Providing guidelines on the development of a program on training (including 'curriculum' development) based on skills needs analysis.
 - One technical seminar on Big Data in ... to discuss the results of the skills analysis, develop the concept for the guidelines and validate and finalize draft guidelines on program of training (if possible, the first seminar as a side-event of GWG Big Data conference).

Deliverables and Time Schedule

- 5. Capacity building and sharing experiences through innovation centers and global network (2019 onwards).
 - Output: TOR on global network of institutions; Support for establishment and operation of innovation centers.
 - Formulate TOR and operating mechanisms for a global network of institutions for training and capacity building on big data.
 - Study on organizational change through innovation centers.

Deliverables and Time Schedule

- 5. Capacity building and sharing experiences through innovation centers and global network (2019 onwards).
 - Output: TOR on global network of institutions; Support for establishment and operation of innovation centers.
 - Providing technical advices and guidelines in collaboration with relevant stakeholders for replicating/adapting 'experiments' as well as moving towards their implementation.
 - Developing a repository for good practices and experiences (including training materials and references) and encouraging country to perform study visits and knowledge sharing.

Key Performance Indicators

- Deliverables are produced on time
- Methods and tools are accepted by the GWG oversight group
- Guidelines on training programs receive positive feedback from stakeholders

Competencies

Big data team (leaders) competencies

- Technical knowledge and expertise
- Analytical competencies
- Judgement and decision-making
- Management and delivery of results
- Building relationships and communication
- ...
- (Credibility and influence)
- (Leadership)

Big data team knowledge

- Statistical (econometric) theory
- Data analysis and classification
- Computer science
- Cybersecurity
- ...

Big data team skills

- Storing large datasets
- Processing live streams of data
- Properly formulating hypotheses
- ...

Big data team attitudes

- Creativity
- Curiosity
- Persistence
- Passion
- Goal-orientation
- ...

Big data roles

- IT infrastructure specialists
- Data engineers
- Researchers
- Team leaders
- ...

Related materials ... (on wiki)

- UNECE High-Level Group for the Modernization of Official Statistics (2016)
- Eurostat, The updated ESS Vision 2020 skills framework* (2017)
- UK Office for National Statistics
- ...

Issues to tackle

- Differences
 - regions, countries, states of affairs
 - governance, institutions, organizations
 - needs, challenges
 - infrastructure, resources

Human resources

Curriculum related problems

- Graduate and postgraduate programs tend to be biased towards either statistics or computer science
- Many vital areas important for official statistics not covered
 - integration of different sources, linkage techniques, advanced sampling and estimation methods, national accounts, time series analysis and seasonal adjustment, ...
- The contents of the courses sometimes outdated or led in an old-fashioned way

Major challenges

- Where to get the right specialists
 - Are universities preparing students to work at NSIs?
- How to attract them to work for public statistics
 - Strong competition on a labor market
- How to develop your own human resources
 - Specialized training programs
- How to retain your staff
 - Attractiveness of the job

How to attract and retain the workforce we need for the future?

- Opportunity to do challenging work
- Availability of data
- Opportunities for professional development
- Impact of the work
- Effective recruiting and hiring process
- Flexibility of work arrangements
- Leadership and mentoring

Implementation cases

New specialized units within NSIs (internal)

- UK
 - ONS's Data Science Campus
- NL
 - CBS's Center for Big Data Statistics
- FR
 - INSEE's SSP Lab (Official Statistics Service Lab)
- It might be useful to review experiences and lessons learned from different organizational structures and data science governance across NSIs

New specialized units within NSIs (internal)

- PL
 - Change of the organizational structure of the Central Office
 - Reorientation of responsibilities
 - ICT Systems, Geostatistics and Census Department
 - Statistics Computing Center
 - Data Engineering Centers in Regional Offices
 - Priorities
 - satellite imagery, traffic sensors, scanner data, web scrapping...

The way forward

Organization readiness for big data innovations

- (Goal oriented) organization culture
- Focus on opportunities
- Organization-wide understanding
- Education, trainings
- Team leaders
- Infrastructure and technical talent
- Start small, make quick progress, achieve quick wins

Implementation measures

- Specialized internal units
 - Internal programs of advanced training
- Strategic partnerships
 - Involvement of external partners (including universities)
- Research grants
 - Strategic methodological advances
- Hackathones

Thank you

for your attention!