

Harnessing the full power of Timely Data for Official Statistics in Africa

Big Data and Official Statistics in Africa
Maximizing the Opportunities

Webinar October 6th 2022
Carlo Vaccari - ISTAT

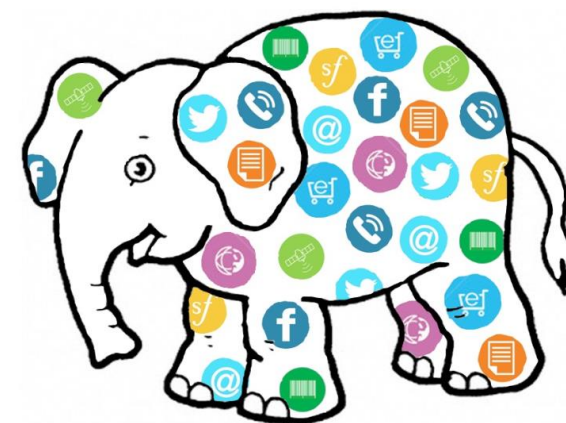
UNECE HLG Experience

Share Italian and European experience

High Level Group to coordinate groups working on Statistical Standards: UNECE, OECD, Eurostat, NSIs

2013: task team

- identify the possibilities and issues that Big Data poses Official Statistics
- analyze the feasibility of production of Official Statistics Big Data sources
- facilitate the sharing across NSOs of knowledge, expert tools and methods



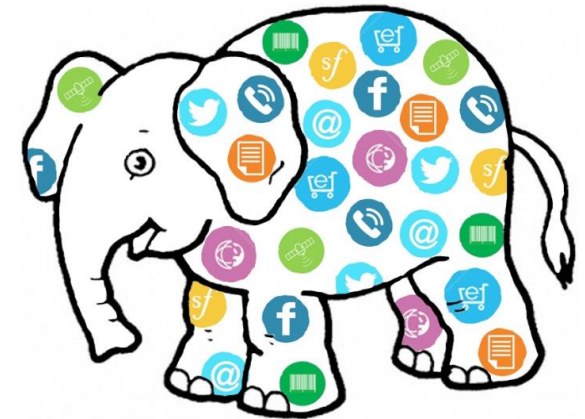
UNECE HLG Experience

2014-2015 [projects](#)

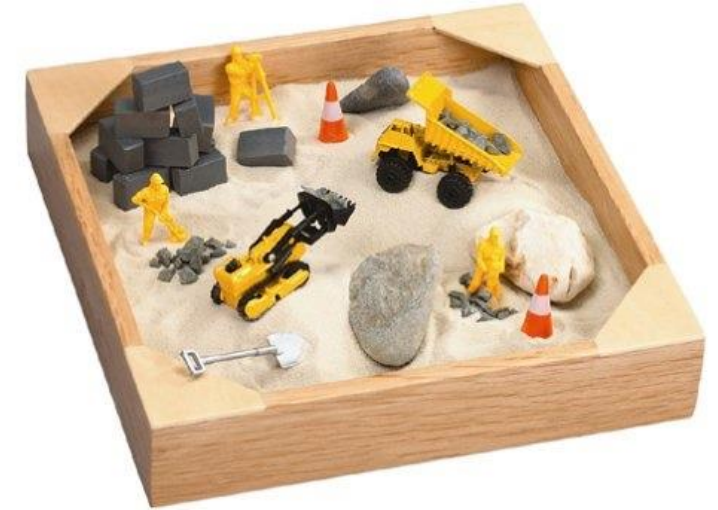
More than 70 people from 22 organisations

Focus on

- Shared computing environment ('sandbox')
- Partnership
- Quality



Sandbox



Sandbox

- Practical [experience](#) on high-end computing center
- Partnership with privates for environment and software tools
- Collaboration between NSOs
- Testing and sharing methods, tools, environments, solutions

Partnership

Providers of data - access to data, privacy and confidentiality with Government, social media, private

Design - research design and development with Academia, research institutes, international organizations

Technology - tools, data and infrastructure with IT companies, data providers, Academia, open-source communities

Users – Coordination and communication with relevant stakeholders: institutional, ethical groups, consumer, business, media, ...





Quality

- Input - input quality framework with indicators:
 - Source: reliability, privacy, availability, costs, procedures, ...
 - Metadata: representativeness, completeness, unique id, ...
 - Data: collection, coverage, efficiency, integrability, ...
- Output - output quality framework with indicators:
 - Metadata: clarity, accessibility, completeness, ...
 - Data: relevance, accuracy, timeliness, accessibility, coherence,
- Throughput - process quality with indicators :
 - Editing and cleaning: unambiguous, objectivity, granularity, ...
 - Conversion and transformations: compliance, categorization, precision, ...
 - Linking: completeness, selectivity, accuracy, unique_id, time_related
 - Aggregation: quantity, confidentiality, integration, validity, accuracy



Eurostat ESSnet on Big Data

- Eurostat ESSnet inherits the experience of the UNECE Sandbox: duration first step 2016 – 2018
- Second step 2018 - 2021
- Wiki available on [Webgate site](#)
- 22 partners: 20 NSIs and 2 statistical organisations
- 8 work-packages (WP) active: for each WP see Documentation and Working Area



Big data pilots in the ESS

- Web scraping on job vacancies
- Smart meters for electricity consumption
- Automatic Identification System for Ships identification
- Mobile phone data
- Scenario for using multiple inputs
- Nowcasting of indicators for short term statistics



For each pilot: documentation, methodology, experimental statistics, software tools (Github! See [example](#))



Big Data other links

European CROS: [Big Data events](#)

European CROS: [Documents and resources](#)

UNSD [Big Data group](#)

OECD [Data Driven Innovation](#)





Africa specificities

- Mobile subscription rate: data mined from mobile phones will be at the hearth of Africa's big data phenomenon (only 5% fixed broadband vs. 28% mobile broadband)
- Mobile data critical for developing countries that do not have sophisticated national statistics collection methods or usable administrative data
- Open data from governments –more than 17 open government data portals – can help push the spread of big data



Africa specificities

- Agriculture central for African countries: Data Science for Precision Agriculture (satellites + mobiles)
- Connection between electricity and internet connection
- About 700,000 professional developers in Africa, more than 50% in five key markets: Egypt, Kenya, Morocco, Nigeria, and South Africa. About 50% unofficial trained (self, online, bootcamp)
- Importance of Hubs and Accelerators

Recommendations

Focus on:

- Need for networks! In networks you can share rare skills and scarce resources
- UNECE ➡ UNECA! African Center for Statistics
<https://www.uneca.org/data-and-statistics>
- Training: regional and sub-regional training schools
- Collaboration: South-South / North-South
- Partnership
- Practice and SHARING



Thank you for your attention!

vaccari@istat.it

vaccaricarlo@gmail.com

