Official Statistics and Citizen Science

Monica Pratesi

Department of Economics and Management

University of Pisa

Definition(s)

- Official Statistics (OS) - that respects defined international principles* - are statistics published by governmental agencies or other public bodies such as international organizations as a public good.

- That is, the OS, the results from **collection and processing of data**, aims to provide a quantitative representation of the society, economy, and environment for **purposes of public interest**, for policy design, and evaluation, and as basis for informing the public debate.

*United Nations Fundamental Principles of Official Statistics (1994, 2011)

Definition(s)

- Citizen science (CS) can be defined* as the non-professional involvement of volunteers in the scientific process, commonly in data collection, but also in other phases, such as quality assurance, data analysis and interpretation, problem definition and the dissemination of results. CS is a growing practice in which scientists and citizens collaborate to produce **new knowledge** for science and society.

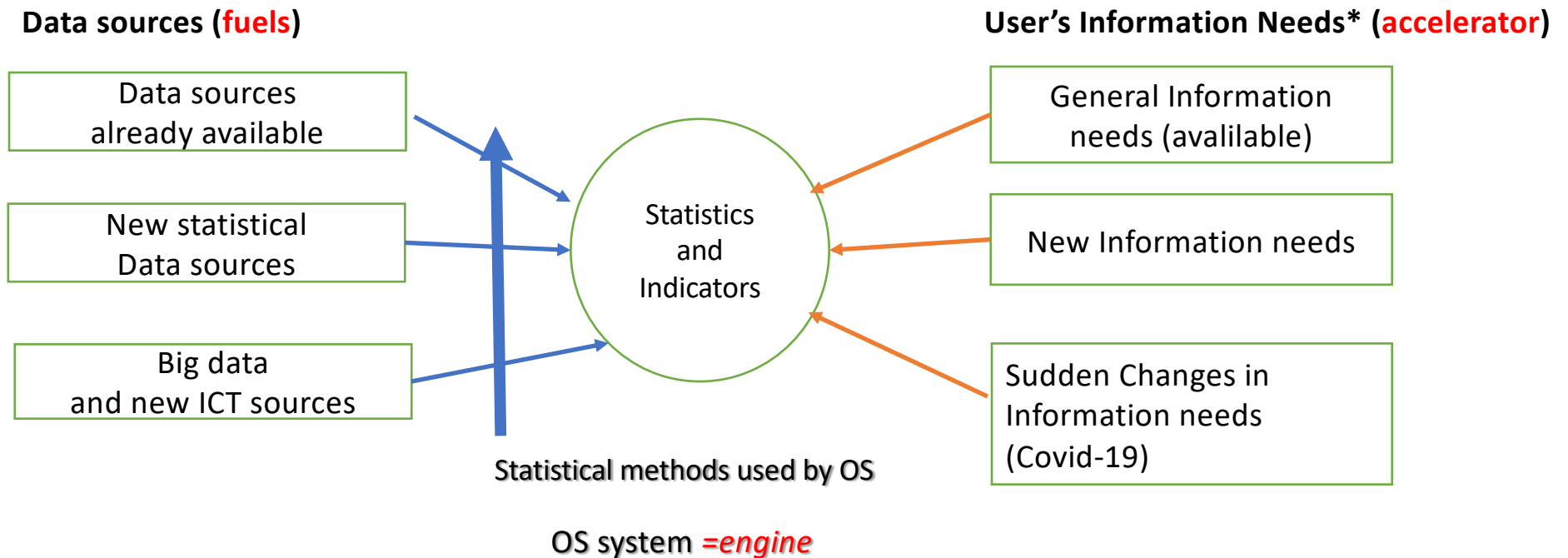*European Citizen Science Association Ten Principles of Citizen Science (2013-2015)

Official Statistics and Citizen Science

- The ultimate goal of this seminar is to investigate the interconnections and, specifically, to verify the possibility that CS can be used for OS, highlighting advantages and disadvantages.

- To do this, it is important to analyze and describe in detail the objectives, organizations, issues, both of OS and CS and their evolution over time

Part I - Official Statistics

- History: surveys from the origins to our times - **Italy**

- Consolidation **1850-1950**

- Official Statistical surveys and Sample Surveys under Threat **1950-1970**

- Increasing needs and demands by the users - **1970-2000**

Evolution– *engine, fuel(s), accelerators*

Data sources (fuels)

Data sources already available

New statistical Data sources

Big data and new ICT sources

Statistics and Indicators

Statistical methods used by OS

OS system *=engine*

User's Information Needs* (accelerator)

General Information needs (avalilable)

New Information needs

Sudden Changes in Information needs (Covid-19)

- *Users: **Citizens, Stakeholders, Companies, Institutions, Government**

- **NB** Need for timely reactions!!!!

History - evolution

Statistics as **Science of the State**, counting and listing populations and State resource.

Definition of Statistics by Gioia , 1826* (see Biggeri L., 1989)

"quella somma di cognizioni, relative ad un paese, che nel corso giornaliero degli affari possono essere utili a ciascuno ed alla maggior parte dei suoi membri od al governo, che ne è l'agente, il procuratore o il rappresentante" e scriveva ancora che la filosofia della statistica esamina "l'influsso, sia in più che in meno, delle cause fisiche e morali sulla popolazione, sulle produzioni, sulle arti, sul commercio, sulle abitudini intellettuali, economiche e morali"

Italy 1850-1950: 1926

- Few topics: population, economic production in agriculture, industry and commerce, exports and imports etc.

- Methods of data collection: essentially Census (of population and industry) and Registers of the different administrations of the State.

- **Istat** turned 90 in July 2016**.**

- It was established in 1926 by Law No. 1162, which gives the new **Central Statistical Institute** the functions carried out until then by the General Statistical Division of the Ministry of Agriculture (Divisione di statistica generale del Ministero dell'Agricoltura).

Italy 1950-1970

- After the II world war, the governments, policy makers and all the stakeholders **need of a lot of statistical information to reconstruct (Piano Fanfani)**

- **The responses by OS**. Do surveys in different domains implementing new surveys: In Italy in particular for the construction of the **National account** and for the implementation of **Sample surveys**.

- The introduction of the sample surveys was due essentially to problems of cost and timeliness in obtaining the results.

.

Italy 1950-1970

- These works were developed also with the **contribution** of many scientists **from the academia** (above all statisticians, but also of the field of sciences involved in the surveys). In Italy, the cooperation between OS and Academia and the exchange of positions of responsibility were always existed, also due to the fact that the President of Istat must be a full professor at the university.

- In Italy in the **1952**, Istat (at that time the Italian central statistical office) appointed a scientific commission for the implementation of a regular **sample survey on Labour Force**, of which **Boldrini** and other academic statisticians were called to be members. In 1954 the first results of this sample survey were disseminated!

1970-2000

The responses by OS: Methodological and Technical (see: Groves, Kalton and Rao, 2009; and O' Muircheartaigh, 2017)

Structured and Standardized Survey Methodology

- Inference and error in surveys;
- Sample design and sampling error
- Methods of data collection
- The role of interviewers and of the respondents
- Nonresponse in sample surveys and their treatment
- Towards the Total Survey Error (TSE) model

Italy 1950-1970

Investment to improve the **quality of surveys**, especially in the social field.

Error Profile for each statistical information (Bailar, 1988)

- Use of **multipurpose sample surveys**
- Use of **systems of surveys** on the families and their components (on the occupation, expenditures, living conditions, quality of life, etc.); and on enterprises and institution (production, value added, costs for raw materials and labor, investments, import and export)
- Investment in the **construction of specific statistical registers** on families, enterprises and institutions (administrative registers)

Italy 1950-1970

- Use of **provisional estimates** to **guarantee the timeliness for the economic indicators** (this introduce confusion on the users and media, that do not understand clearly the estimations)

- Preparation of methodological **Manuals** (in Italian **Metodi e Norme**) and providing of **Meta-Data** where all the information on the characteristics of each statistical information disseminated are provided

- **Preparation of Micro data files**

Italy 1970-2000 ...and beyond

The citizens are involved too much as respondents

- For different samples surveys and for the contacts with the public administration agencies

- Big burden for Citizens and no clear return for them, and no involvement in the definition of their information's needs

- Citizens' lack of an organized and detailed demand of statistical information with strong frustration and a sense of mistrust (see Citizen Science)

Italy 1970-2000 …and beyond

Implementation of the National Statistical System (Sistan)

- Since 1989 ([Legislative Decree 322/89](#)) Istat has been performing the role of directing, coordinating, and providing technical assistance and training within the [National Statistical System](#) (Sistan). in order to rationalise the production and publication of information and to optimise resources allocated to official statistics.

- Sistan is **made up of Istat**, central and branch statistical departments of Public Administrations, of local and regional bodies, Chambers of Commerce, other public bodies and administrations providing statistical information

- The Sistan prepares a 3 years **National Statistical Plan** approved by the Parliament

- It is important to point out that for the updating and control of the plan, **Quality Circles** have been established and functioning. The circles include many representative of the Sistan's bodies and of other institutions, but there are no representatives of the citizens

- A Commission for the guarantee of statistical information (**COGIS**) have been established and functioning

- CUIS– Conferenza Utenti Informazione Statistica

Italy 1970-2000 ...and beyond

.... in 90s the trust in OS diminished

 Essentially for the **reduction of the response rate** in the (sample) surveys and for the interference of some governments in the dissemination of the results

- For a complete and clear picture of the situation see: **Statistics, a Matter of Trust** (Green Book, UK, 1998)

The responses by OS

- During the second part of 90' the International OS Organization prepared and then published the **UN Fundamental principles of Official Statistics** and the EU **Code of Practice**, with the dual purpose of : **improving the confidence in OS** by setting up **minimum quality standards** and identifying checking procedures for monitoring statistical processes; serving as a self-regulatory instrument to enhance the quality of the statistics, produced and disseminated

- **Istat had an important role** in both the activities

Italy 1970-2000 ...and beyond

Many surveys and too many data difficult to interpret by many users, essentially for their lack of statistical literacy

- Many expensive surveys and not easy to interpret results
- The available **statistical information at the same time deficient and overabundant**
- Information: Rice, 1977: The information you have is not want you want; the information you want is not what you need; the information you need is not what you can obtain

2000…and beyond

BES (Equitable and Sustainable Well-being) and for **SDGs (Sustainable Development goals of the UN)**, and on the improvement of the preparation and dissemination of the **micro data files**

- The BES aims of evaluating the progress of society not only from an economic, but also from a social and environmental point of view. To this end, the traditional economic indicators, GDP first of all, have been integrated with measures of the quality of people's life and of the environment

- The set of indicators, illustrates the **12 domains* relevant for the measurement of well-being**, is updated and illustrated annually in the Bes report.

- In 2020, the **set of indicators has been expanded to 152** (it was 130 in previous editions), with a deep revision that takes into account the transformations that have characterized Italian society in the last decade, including those linked to the spread of the COVID-19 pandemic.

CNEL/Istat: consultation to define domains, inclusive of citizens

2000...and beyond

Istat modernization:

- **The implementation of a new model of official statistical production based on integrated use of different data sources**

- Istat worked on the implementation of Integrated system of registers (SIR), sample surveys and Big Data and Internet Data, both for Population and Enterprises – Permanent Census

- Experimental Statistics: innovation

- Valorization of Social and Economic Statistics: *Antennas*!

SIR – Sistema Integrato Registri

R. Monducci (2019)

9

Data Today

- Bloomberg financial market data
- City agency household data
- Retail orders and delivery
- Airline passenger data
- Phone connections
- Email connections
- Smart-grid data
- Smart-car data
- Phone content
- Email content
- Google maps
- Facebook
- Twitter
- National Statistical Agencies data

Risk: "You are treated according to your data, not your reality"
Colm O'Muirchetaigh, 2017

Part II - Citizen Science

- Definition
- History: first experiments, problems and solutions, Associations
- What's up? Commitment of actors, Big data and CS: two examples
- From here where do we go? CS formonitoring SDGs; CS for Official Statistics
- Conclusions: towards Trusted Smart Statistics

Definition

- Citizen science (CS) can be defined as the non-professional involvement of volunteers in the scientific process, commonly in data collection, but also in other phases, such as quality assurance, data analysis and interpretation, problem definition and the dissemination of results.

- Citizen science is a growing practice in which scientists and citizens collaborate to produce **new knowledge** for science and society.

- Until recently, citizen science has been recognized mainly in the natural sciences and local history. Developments are foreseen in policy formulation and data politics. European Commission, 2020

History

Example: Professor James Murray, who in 1879 became director of an Oxford University Press project, The New English Dictionary on Historical Principles (now known as the Oxford English Dictionary). Enlist volunteers from everywhere English is spoken. He writes an appeal to English-speaking people around the world, asking them to send their contributions on slips of paper. Booksellers, librarians and newsagents distribute it. The slips pile up

1900 – NO traditional Christmas "Side Hunt,"

December 14, 2020 - January 5, 2021

There is a specific methodology to the Christmas Bird Count, and all participants must make arrangements to participate in advance with the circle compiler within an established circle, but anyone can participate. Each count takes place in an established 15-mile wide diameter circle, and is organized by a count compiler.

Source: Video still taken from "Christmas Bird Count" by Chan Robbins. http://www.audubon.org/news/11-reasons-christmas-bird-count-rocks

2020

- What is the difference, if any, between citizen science, participatory science, post-normal science, civic science, and crowd science? Is citizen science just a new political term in order to obtain funding? We do not enter here in the debate.

- A search of the ISI Web of Knowledge topic 'citizen science' revealed 2625 publications of which 1028 could be attributed to European first authors (8 nov 2019 – Cost action CA15212)

2021
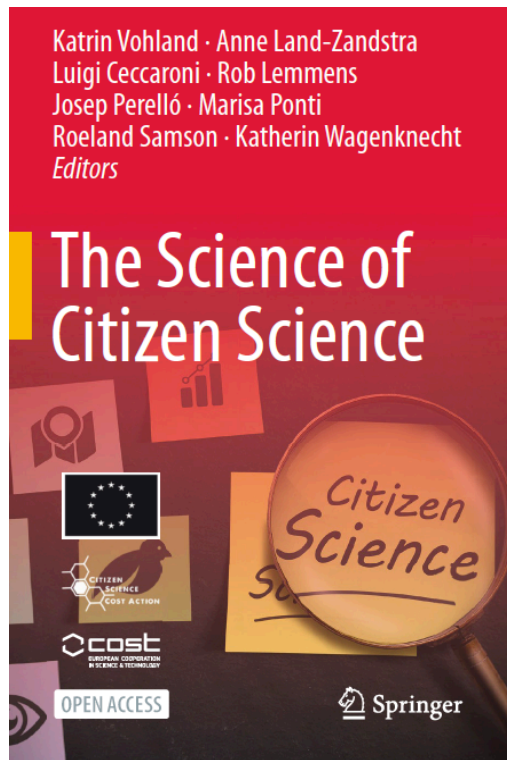
COST - Cooperation in Science and Technology

This COST action was concluded in September 2020

Fig. 1.1 European Union member countries during the key phase of COST Action CA15212 (the UK left the EU on 1 February 2020) and member countries of COST. Except for Moldova and Iceland, all COST countries are members of CA15212. Country data: World Bank Official Boundaries; COST data from www.cost.eu

Problems, solutions

- In the majority of cases, citizens contribute **data** to an established research question, which leads to statements from scientists such as 'you don't get eureka moments' (Riesch and Potter 2014, p. 8). In fact, science does not only mean contributing to a specific question, but a deep knowledge of the whole field, its methods, its history, its literature, its discourses. This takes time, for which scientists are paid, and citizen scientists are not.

- **knowledge** as a commons is seen as a **public good**

Example: smog traker

Problems

- participant retention and the adaptation of new digital technologies
- interaction of human users with technology and citizen-generated information
- data quality
- a conceptual model is needed to achieve a common understanding and representation for citizen science projects, their participants, and their outcomes. Based on international standards of data interoperability, this model is designed for information sharing amongst citizen science projects

Associations

- They want to establish citizen science as a recognized, promoted and funded approach, one that fosters scientific literacy and the democratization of science
- ECSA - German Law on April 1st 2014. European Citizen Science Association
- http://www.lteritalia.it/ 1995
- http://www.citizenscience.enea.it/ 2016
- 2015 "Citizen Science Italy" https://www.globe.gov/web/italy-citizen-science
- http://www.scienzacollaborativa.it/

Ten Principles of Citizen Science

These principles are also available in other languages ⬀.

ecsa | European Citizen Science Association

1. Citizen science projects actively involve citizens in scientific endeavour that generates new knowledge or understanding. Citizens may act as contributors, collaborators or as project leaders and have a meaningful role in the project.

2. Citizen science projects have a genuine science outcome. For example, answering a research question or informing conservation action, management decisions or environmental policy.

3. Both the professional scientists and the citizen scientists benefit from taking part. Benefits may include the publication of research outputs, learning opportunities, personal enjoyment, social benefits, satisfaction through contributing to scientific evidence, for example, to address local, national and international issues, and through that, the potential to influence policy.

4. Citizen scientists may, if they wish, participate in multiple stages of the scientific process. This may include developing the research question, designing the method, gathering and analysing data, and communicating the results.

5. Citizen scientists receive feedback from the project. For example, how their data are being used and what the research, policy or societal outcomes are.

6. Citizen science is considered a research approach like any other, with limitations and biases that should be considered and controlled for. However unlike traditional research approaches, citizen science provides opportunity for greater public engagement and democratisation of science.

7. Citizen science project data and metadata are made publicly available and where possible, results are published in an open-access format. Data sharing may occur during or after the project, unless there are security or privacy concerns that prevent this.

8. Citizen scientists are acknowledged in project results and publications.

9. Citizen science programmes are evaluated for their scientific output, data quality, participant experience and wider societal or policy impact.

10. The leaders of citizen science projects take into consideration legal and ethical issues surrounding copyright, intellectual property, data-sharing agreements, confidentiality, attribution and the environmental impact of any activities.

The Principles

- The dominant method for engaging the public in scientific research is the 'contributory' method, where the public solely **collect and submit data** to research projects

- little is published on the practice and impacts of **collaborative and co-created citizen science**, and additional research and sharing of in this area would be welcome

- There is evidence that **feedback** is a motivator for more participation

- Too few citizen science projects give participants **direct access to the resulting dataset**, and few project websites clearly describe if/how data will be shared with national and international databases. Cleaning, formatting and archiving data requires resources and infrastructure, and this vital step must be planned into project timescales and funding at the outset.

- Imposition of a top-down set of standards for citizen science would be incongruent with its naturally bottom-up, flexible nature, but the Ten Principles may nonetheless serve the same aim of promoting excellence in science research, environmental protection, and public engagement and active involvement in the scientific and policy processes.

What's up?

Committment of actors

- Citizen science fosters an open and participatory approach to science, **reducing the distance between science and society**, and contributing to the goal of an inclusive society. Together with public and private actors, citizen scientists can play a role in developing society, improving communities, and promoting public participation.

- Each activity requires a certain type of competence, a certain degree of skill and a defined level of commitment or "engagement", which has been quantitatively estimated only by a few research

- **Why do People Engage in CS?**

- The lack of surveys on the level of engagement of participants is a limit for the development and dissemination of projects, as having a clear vision of how citizens behave in the various phases of research and analysis is essential to be able to propose and evaluate new engagement strategies that guarantee their active participation. Rotman et al. (2014), Aristeidou, Scanlon & Sharples (2017),

- On Twitter, 336 personal accounts contain "citizen science" in their description. Researchers (PhD students, post-docs, or faculty) form the largest category in the dataset (36%, Figure 2A). Two-thirds of these researchers work with data produced by citizen science projects in various fields: environment (42%), health (9%), astronomy (8%) and other fields (10%). The last third consists of researchers who study participation in citizen science projects. These individuals work in the fields of user experience, design studies, gaming studies, science and technology studies, or science education. Elise Tancoigne (2019) Invisible brokers, Citizen Science on Twitter, JCom

- **Personal interest**: Expanding your knowledge; promote their hobbies; spend timewith friends or make new friends. **Self-promotion**: Promoting new opportunities for oneself; participate in an experience useful for the Future. **Self-efficacy**: Taking control of the scientific process; take part even in the mostcomplex project as well as scientists. **Social responsibility**: Strongly related to the community to which they belong and the role of the system. support of local institutions.

- "**long term participation**". By participating in the projects, volunteers are exposed to the effects that the citizen science has on their environment and this can push them to behave like mediators between local communities and scientists. The role of education is also important, considered a tool to empower local populations.

Big data and CS

- engagement by gamification, social interactions and rewards
- **widespread use of smartphones** with a wide variety of sensors (such as proximity sensors, camera, microphone, GPS sensors) allows. citizen scientists to detect not only data closely related to scientific research and the object of the project, but also a huge amount of information generated "implicitly" during social interactions which, if properly integrated, can lead to unexpected discoveries. This possibility represents a new opportunity for citizen science.

- **Passive sensing**: participants who offer the project a resource that belongs to it (for example, for example the mobile phone) for automated perception. The information collected through the sensors are then used and analyzed by scientists

- **Volunteer computing**: participants share computational resources (on their computer, tablet or smartphone), allowing scientists to operate complex computer models when the device is not being used

- **Volunteer thinking**: volunteers offer their ability to recognize models or to analyze information that will be used in a scientific project

- **Environmental and ecological observation**: pollution monitoring environmental protection and observations of flora and fauna

- **Participatory sensing**: participants download an app that allows you to collect data using their mobile phone sensors (such as Wi-Fi, Bluetooth, GPS) and in some cases they also provide behavioral information

- Community/Civic science: "bottom-up science", this type of activity is started and led by a group of participants who identify a common problem and address it using scientific methods and instruments

https://envirocar.org/index.html?lng=en

https://ebird.org/home

Part III – Official Statistics and Citizen Science

- From here where do we go? CS for monitoring SDGs; CS for Official Statistics

- Conclusions: towards Trusted Smart Statistics

From here where do we go?

- Is it time for a new regime of data collection through different devices? How can we capture citizens' meanings and intentions when they produce data? Can we develop 'smart' methods that do not rely on cooperating with, and data generated by, large tech companies, but by developing methods and data co-produced with citizens?

CS for monitoring SDGs

Mapping citizen science contributions to the UN sustainable development goals
Fraisl et al, Sustainability Science (2020) 15:1735–1751

Fig. 3 Citizen science contributions to SDG monitoring by tier classification. The green shading denotes direct contributions, the yellow are supplementary contributions, and the orange shows indicators that have both direct and supplementary contributions. The values within each box are the SDG indicator numbers.
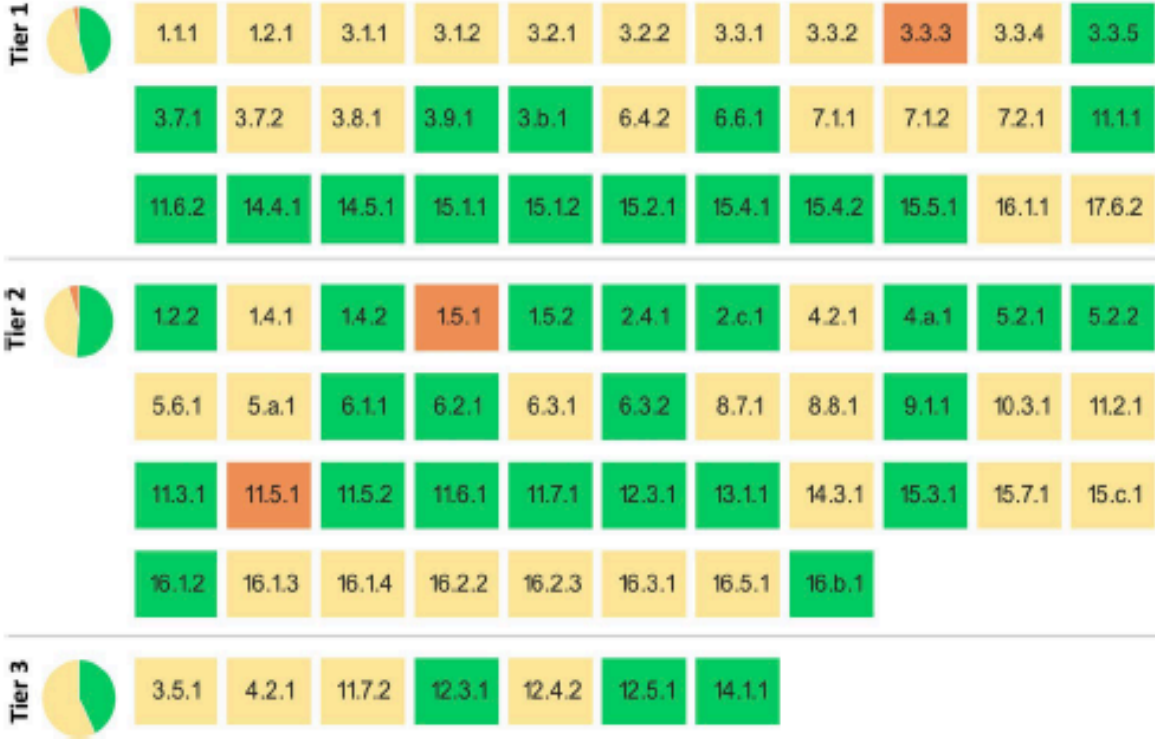
- The most remarkable finding from this review process is that citizen science has the potential to contribute to all 17 SDGs, since it is already contributing or could contribute to at least one indicator per goal.

- For example, **household surveys**, as the main data source for one third of all SDG indicators (ISWGHS 2019), can benefit from citizen science approaches. Citizen science could complement household surveys at a local level by providing more granular data to inform local decision making processes through involving community members to collect data.

- Instead, our goal was to identify at least one project in which the data are supporting, or could support, the SDG monitoring process. Hence, we propose that this work could serve as the basis for more elaborate research focusing on one or more specific goals or targets, studying the methodologies and approaches used by identified citizen science initiatives and providing guidelines on including them in official statistics.

-

CS for Official Statistics

- Big data in Official Statistics:

  data generated by companies in the private sector with all the aforementioned **drawbacks**

Official and Unofficial Statistics for the Public Good?
David J. Hand 2017

Example 1: Supermarket scanner data for price indices

"A fifth of EU countries use [supermarket scanner] data already"
HICP, Practical Guide for Processing Supermarket Scanner Data, Eurostat, Sept 2017

- *The Netherlands CPI: 2002*
- *Norway CPI: 2005*

- Gives both prices and quantities
- Not merely a sample (by collector) of goods sold
- Extended periods
- Only items actually sold (not just on shelves)
- Includes real prices (discounts etc)
- Cheaper to collect
- Bigger data sets
- Churn is visible

But:

- Are the scanner data suppliers representative of the population you want to cover?

- Different types of outlet (even within one chain) may have different pricing strategies;

- Do the scanner data suppliers change internal strategies (special offers; 80/20 → product line rationalisation)?

- Will the supplier be in business in a year's time?

- Is there also a possible feedback/gaming danger?

EUROPEAN STATISTICS DAY
20.10.2017
BETTER DATA.
BETTER LIVES.
Lisbon, PORTUGAL

ASESD Camilo Dagum Centre

Official Statistics in evolution? paradigm shift

	Predigital world (2 decades ago)	Digital world (now)
Inputs	Surveys (Census) + Administrative archives	New digital data (Big data)
Quality	Timeliness, Relevance, Accuracy, Costs	Improvements? Potential gains?
Operational conditions	Monopoly	Data players, ecosystem, disintermediation
Processes, organization, regulations, practices	Hardware, software for Inputs	Hardware, software, *humanware*

The term "Trusted Smart Statistics" (TSS) was put forward by Eurostat to signify this evolution
(Ricciato et al., 2020) and officially adopted by the European Statistical System (ESS) in 2018
in the so-called Bucharest memorandum (European Statistical System Committee, 2018).
Official Statistics: credible, accurate, trustworthy

Survey data?

Survey data

- Well developed theory
- Uncertainty bounds (from sampling variation)
- *Concern about decreasing response rates*
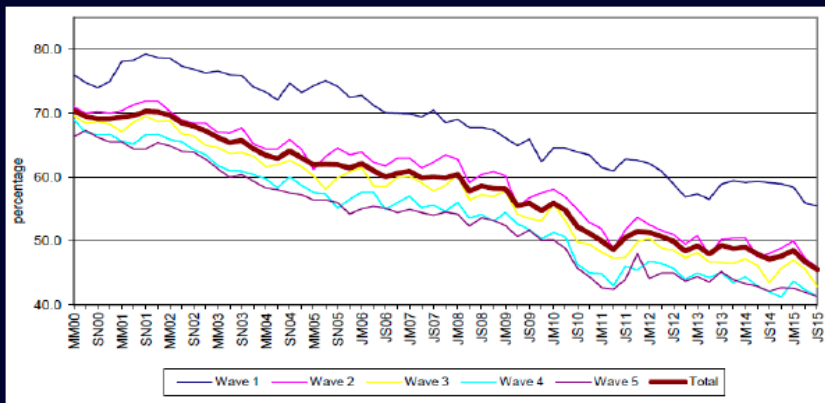
UK's LFS quarterly survey wave-specific response rates:
March-May 2000 to July-Sept 2015

Source: http://www.ons.gov.uk/ons/guide-method/method-quality/
specific/labour-market/labour-force-survey/index.html

EUROPEAN
STATISTICS
DAY
20.10.2017
BETTER DATA.
BETTER LIVES.
Lisbon, PORTUGAL

Italy - LFS - Response rates
2014-2020

nord ovest — nord est — centro
sud — isole — italia
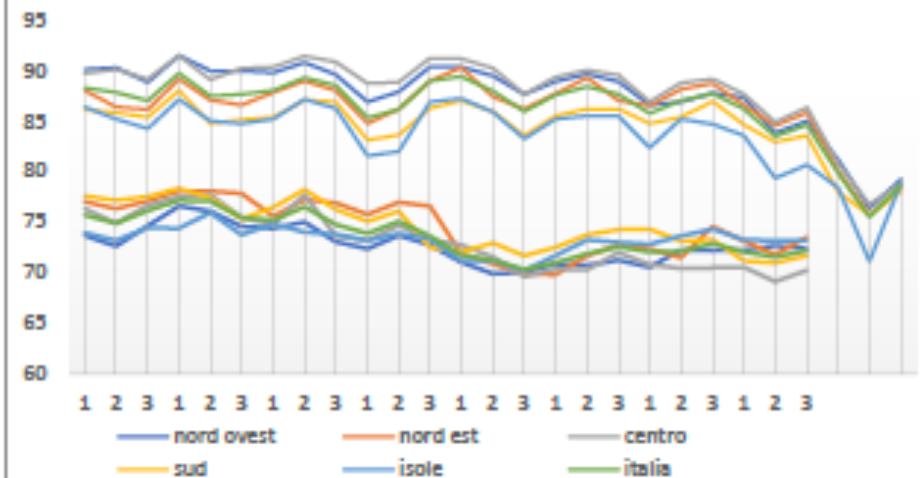
Istat SIQual, link http://siqual.istat.it
TASSO_MRT: Tasso di Mancata Risposta Totale. Viene utilizzata la formula consigliata da Eurostat
(Eurostat (2020 European Statistical System (ESS) handbook for quality and metadata reports —
2020 edition, pag.271-272 https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-
/KS-GQ-19-006)

Tasso di mancata risposta totale

$$= \frac{\text{Unità non rispondenti} + \alpha \cdot \text{Unità non risolte}}{\text{Unità rispondenti} + \text{Unità non rispondenti} + \alpha \cdot \text{Unità non risolte}} \cdot 100$$

Citizen Data and Trust in Official Statistics
Ruppert et al, ECONOMIE ET STATISTIQUE / ECONOMICS AND
STATISTICS N° 505-506, 2018

- **Big data and citizens are inseparable:** from smartphones, meters, fridges and cars to internet platforms, the data of most digital technologies is the data of citizens.

- In addition to raising political and ethical issues of privacy, confidentiality and data protection, the repurposing of big data calls for rethinking relations to citizens in the production of official statistics if they are to be trusted.

Data and statistics
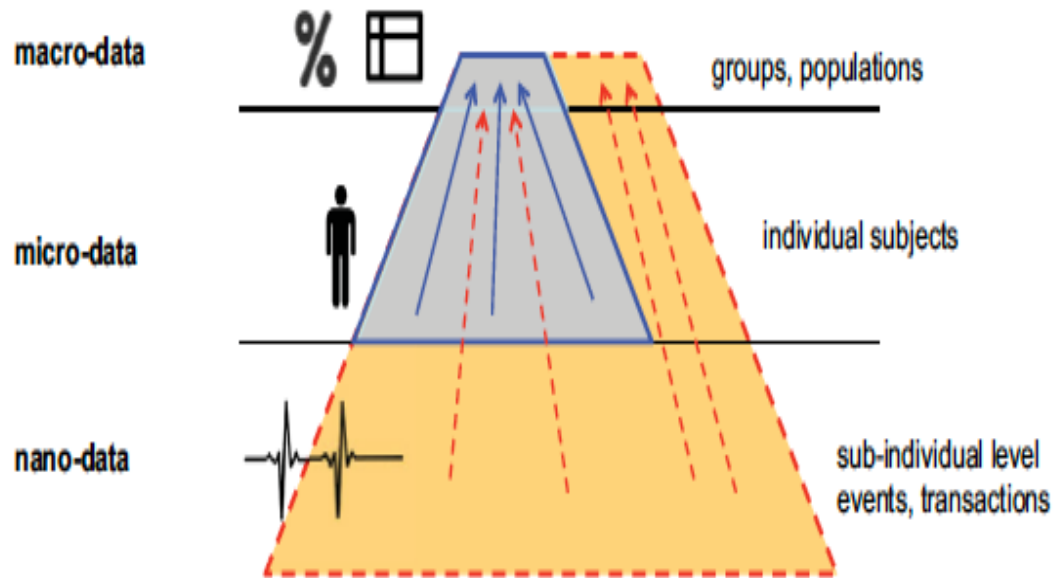
Main role of **official statistics** remains the production of aggregate indicators at the **macro-level** to represent parsimoniously the whole society.

Reduction process from large amount of micro-/nano-data toward a parsimonious set of macroscopic statistics

Nano data or granular data, behavioral data
Ricciato et al (2020) Data and Policy

Nano-data

- Potential breach of individual privacy. It encompasses also the possibility to exert mass surveillance, to enable different dystopian forms of social control, and to generate more subtle, less evident forms of power asymmetries

- Cambridge Analytica: "A main lesson to draw is not that an academic, an internet platform, and a data company are culpable. Rather it is that data and politics are inseparable such that academics, statisticians or app developers cannot be naïve but must be reflexive about how they may be implicated in the ways data is part of emerging forms of power relations". Ruppert, Statistical Journal of the IAOS 35 (2019) 633–641

From CS to citizen data and to citizen statistics

- main principles for a citizen data app for official statistics
- extend beyond Big Data and include methods and data of statistical production such as **questionnaires and registers**
- there are examples of Apps performing Smart Surveys

Michela Franchini [1], Stefania Pieroni [1,*], Nicola Martini [2], Andrea Ripoli [2], Dante Chiappino [2], Francesca Denoth [1], Michael Norman Liebman [3], Sabrina Molinaro [1] and Daniele Della Latta [2]

[1] Data Learn Lab, Institute of Clinical Physiology of the National Research Council, 56124 Pisa, Italy; michela.franchini@ifc.cnr.it (M.F.); francesca.denoth@ifc.cnr.it (F.D.); sabrina.molinaro@ifc.cnr.it (S.M.)
[2] Data Learn Lab, Gabriele Monasterio Foundation, 1, 56124 Pisa, Italy; nicola.martini@ftgm.it (N.M.); ripoli@ftgm.it (A.R.); radio1@ftgm.it (D.C.); dellalatta@ftgm.it (D.D.L.)
[3] IPQ Analytics, Kennett Square, PA 19348, USA; michael.liebman@ipqanalytics.com
[*] Correspondence: stefania.pieroni@ifc.cnr.it

check for updates

@DRESSbot by D2L

4 aprile, 2020 da D2L Team

Questo bot per telegram, sviluppato dal Data Learn Lab (laboratorio congiunto tra l'Istituto di Fisiologia Clinica del CNR e Fondazione Toscana Gabriele Monasterio) serve per il monitoraggio attivo del rischio di contagio da SARS-Cov-2 e per supportare le istituzioni preposte (Ministero della Salute, Protezione Civile, Regioni, Province, Comuni e ASL).

Per attivarlo, basta che in Telegram inizi una nuova chat con il bot @D2LDress_bot.

Una volta avviato il bot ed accettati i termini per l'utilizzo, entrerai a far parte della piattaforma. Nelle ore successive ti verrà chiesto di condividere la tua posizione per personalizzare il tuo rischio di contagio. Nei giorni successivi DRESSbot ti contatterà facendoti alcune domande di vario genere. In funzione delle tue risposte, il sistema di intelligenza artificiale (AI) che controlla la piattaforma, sarà in grado di aggiornare il tuo rischio tenendo in considerazione oltre alla posizione anche il tuo stato di salute.

Infine riceverai nuovi suggerimenti, scientificamente validati, per la promozione della salute e gestione dello stress e dell'isolamento.

NB. DRESSbot non monitora costantemente la tua posizione ma ti chiede di geolocalizzarti una tantum e calcola il tuo rischio in funzione della posizione che hai fornito in quel momento.

Clicca qui per collegarti direttamente:

https://telegram.me/D2LDress_bot

oppure cerca direttamente il bot @D2LDress_bot in telegram.

DRESSbot bot

attivo del rischio di contagio da SARS-

⚡ APERTURA RAPIDA

22:27

DRESSbot
Per piacere puoi inviarmi la tua posizio...

22:28

GRAZIE della risposta! Stiamo lavorando per fornirti un modello di rischio personalizzato!

22:28

17 marzo

DRESSbot bot

17 marzo

Ciao! Queste sono le domande previste per te oggi:

12:00

Quanti anni hai? 55-59

12:00

Sei uomo o donna? Donna

12:00

Sei un lavoratore? Si

12:00

Sei uno studente? No

12:00

Sei un pensionato? No

12:00

Sei disoccupato? No

12:00

Sei casalinga(o)? No

12:00

daily DRESS energy

DRESSbot bot

👍

12:00

Contagi totali (incremento giornaliero)

RE 37931 (+221)
MO 52022 (+201)
BO 75093 (+479)
RA 23550 (+216)
LU 18241 (+19)
PT 9826 (+154)
PO 15159 (+69)
FI 47448 (+258)
PI 22451 (+159)
SI 9581 (+107)
AR 16197 (+121)

70000
60000
50000
40000
30000
20000
10000

DRESS: Cosa accade intorno a Te

18:47

L'adesione a DRESS è **assolutamente gratuito** e non comporterà:
- alcun rischio per la tua salute
- nessuna violazione della tua privacy
- nessun uso illecito dei tuoi dati sensibili, in accordo con il **D.Lgs. 10 agosto 2018, n. 101 – Disposizioni per l'adeguamento della normativa nazionale alle disposizioni del Regolamento UE 2016/679**

L'adesione al sistema DRESS ti potrà fornire i seguenti **benefici:**
- ottenere, in tempo reale, la stima del tuo rischio di contagio
- ottenere informazioni, scientificamente validate, circa i comportamenti più efficienti per mantenere il tuo stato di salute
- consentire l'accesso a documentazione, opportunamente vagliata e aggiornata, circa l'emergenza sanitaria da Covid-19, gli obblighi normativi e la modulistica necessaria per la certificazione degli spostamenti

Infine, aderendo a DRESS fornirai il tuo prezioso aiuto nell'identificare nuove soluzioni per far fronte all'attuale emergenza sanitaria.

Privacy by design

Caro cittadino,

aggiungendo tra i tuoi contatti @D2LDress_bot ed avviando il bot, entrerai a far parte del sistema DRESS – Autovalutazione Del Rischio e Supporto Socio-Sanitario che il Data Learn Lab (laboratorio congiunto tra l'Istituto di Fisiologia Clinica del CNR e Fondazione Toscana Gabriele Monasterio) ha predisposto per attuare il monitoraggio attivo del rischio di contagio da SARS-Cov-2 e per supportare le istituzioni preposte (Ministero della Salute, Protezione Civile, Regioni, Province, Comuni e ASL) nell'attivare servizi dedicati ai cittadini, sulla base dei loro bisogni sanitari o sociali.
DRESS è uno strumento di Medicina Partecipativa, predisposto per consentire agli utenti di essere parte attiva dello studio, nell'idea di costruire una sanità CON il cittadino e non PER il cittadino.

Tutte le informazioni che ci renderai disponibili **saranno soggette** **segreto statistico e trattate in forma rigorosamente anonima e aggregata, in conformità a quanto disposto dal D.Lgs. 10 agosto 2018, n. 101 – Disposizioni per l'adeguamento della normativa nazionale alle disposizioni del Regolamento UE 2016/679.** I tuoi dati saranno utilizzati esclusivamente per gli scopi previsti dal progetto o comunque per finalità di ricerca ad esso correlate.

Il titolare del trattamento dei tuoi dati è Fondazione Gabriele Monasterio (FTGM) in quanto tale, mentre il Responsabile del trattamento dei tuoi dati è individuato nel Direttore Generale di FTGM.

	Traditional data sources Surveys (Census) + Administrative archives	Smart surveys New digital data (Big data)
Interview time	Static, time scale (month, quarter, year)	Continuos interaction
Interaction model	Single survey session Batch of questions CATI, CAPI, CAWI	Respondent burden is diluted Continuous flow of questions, low rate over long interval, long two-way dialogue
Unit	Single record: individual data subject (person, household, company) collection of variables	Collection of events at «sub-individual» level: data points: transactions, encounters, movements..
feedback	Collective feedback: survey report	Individualized reports, selected figures, better communication of the importance of OS

Official Statistics for Citizen Statistics

• To do list:

1. to identify the possible design elements of a 'citizen data app' or a platform that allows different apps to communicate and share data for statistics for government and research.

2. privacy by design; be based on a form of 'citizen science' and should be open source. The resulting design (and if successful, app development) would ideally be applicable to a wide range of NSIs in Europe and beyond.

3. to think of citizens as not mere research subjects, but as actively involved in the production of data as opposed to traditional methods where they have been otherwise 'passive' subjects – data literacy, statistical literacy

- OS have the opportunity to pioneer new models of data collection and data use based on their constituent principles of transparency, openness, independence, and democratic control (United Nations, 2014)
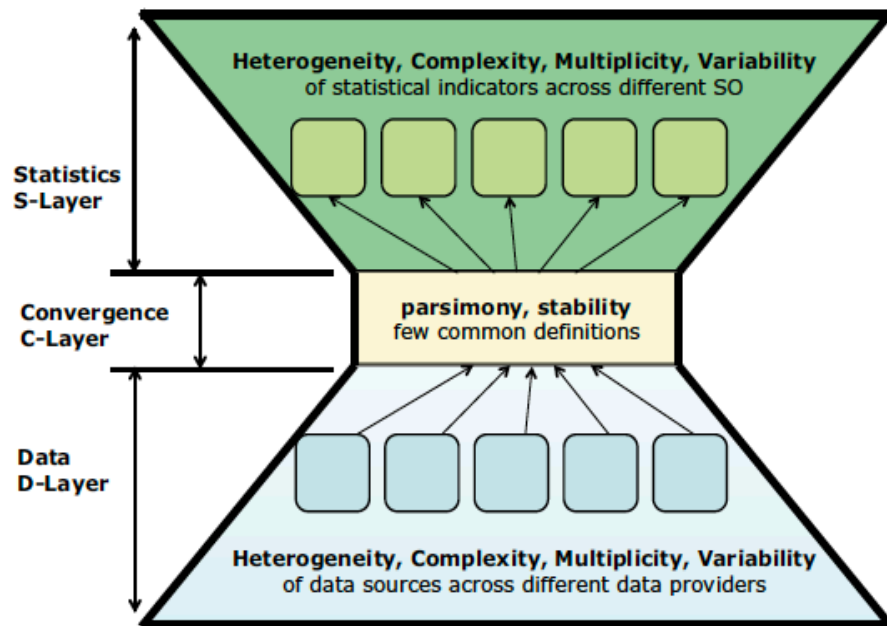
Trusted Smart Statistics -1

TSS principles:

- Sharing computations
- Sharing control over computation execution – distributed computation
- Multidisciplinary teams
- Engineering trust
- From methodologies to code

- Data scouting
- Data stewardship

Trusted Smart Statistics -2

First lower layer of data processing: from collected data (possibly unstructured and/or rich of technology-specific information that is not relevant for official statistics) to intermediate data (and associated meta-data) that can be more easily interpreted and further processed by statisticians, possibly in combination with other data sources, following the **multi-source paradigm** discussed below. Close cooperation between statisticians and domain-specific technology specialists is required only at this first (lower) layer, in order to build functions with technology-specific logic for selecting and transforming the

Third layer of statistical analysis: data components that are relevant for further statistical purposes.

Multi-purpose sources and multi-source statistics
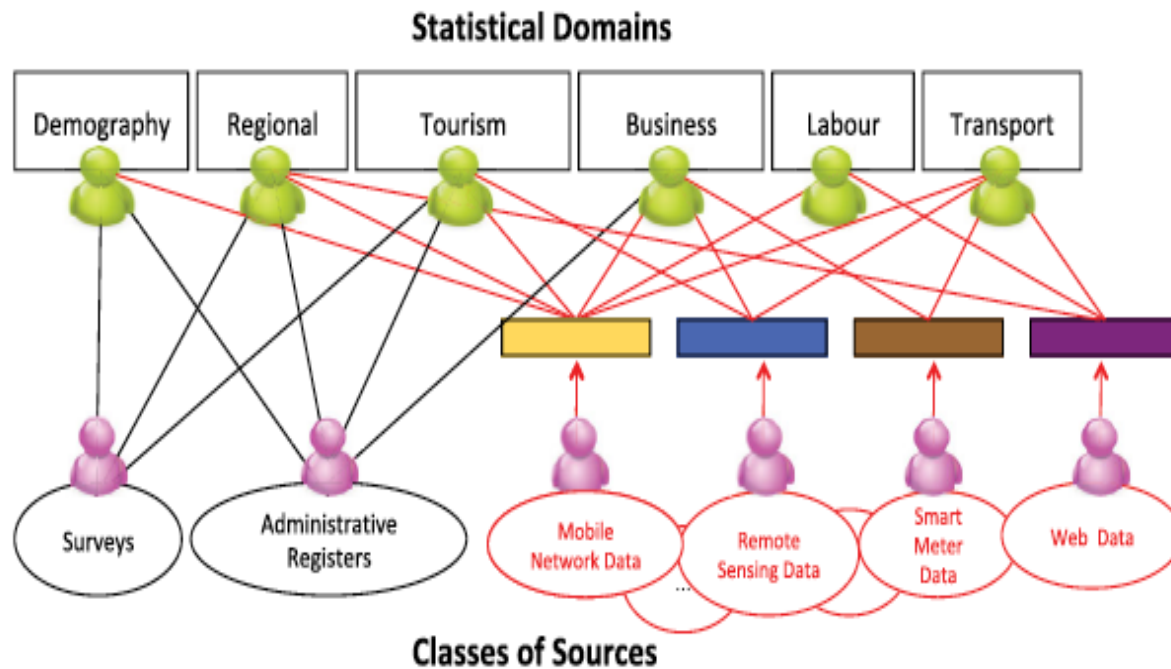
Each class of data serves multiple statistical domains (multi-purpose sources) and each statistical domain can benefit from different sources of data (multi-source statistics)

- novel statistics and indicators can be developed integrating multiple data sources, including combinations of traditional and nontraditional data sources (Marchetti et al, *Journal of Official Statistics*, 31 (2): 263-281

- surveys can be used to calibrate indicators computed from new data sources, for example, against selectivity bias and/or under-coverage errors, similarly to what is done today in those countries where census is based on administrative data integrated by specifically designed sample surveys (see, e.g., ISTAT, 2014).

- different competencies in the teams of human experts that are called to develop the data processing methods (multi-disciplinarity)

- a sound understanding of data generation systems and processes that are at the same time social and technical, or socio-technical

- sound understanding of both behavioral aspects and technological aspects

- …that is Official Statistics and Citizen Science!

- THANK YOU!