

Trusted smart statistics: new statistics for decision makers. Istat's experience

Massimo De Cubellis, Gerarda Grippo

1. Introduction

This paper describes the path followed by the European Statistical System and the Italian Statistical Institute to respond to changes due to the ongoing digital transformation. The digitalization of most daily activities has led to an enormous production of new data, prompting National Statistic Institutes (NSIs) to integrate these new sources within their production processes in order to: (i) enrich their information offerings; (ii) respond to the growing needs of stakeholders; (iii) support decision-making processes in a more efficient way. To achieve these results, NSIs must adapt their organizational, methodological and research paradigms to produce innovative outputs that make structured use of big data. These outputs, called Trusted Smart Statistics (TSS), represent NSIs' response to the changes taking place inside and outside the Institutes.

2. Trusted smart statistics: new statistics for decision makers

The last decades have been characterised by profound transformations that have led to significant changes due to the increasing availability and interaction of extraordinary technological innovations. Digitalization has given a strong boost to the data production and to the process datafication of society (Mayer-Schönberger, Cukier 2013). The spread of smart devices in many areas of daily life has led to the generation of increasingly granular data from a spatial and temporal point of view, which represent increasingly interesting sources for public and private organizations. The digital revolution has created a new environment and a new ecology; it has changed our culture in a profound and significant way. All these changes represent not only a digital but also a true ontological revolution (Floridi 2017).

The capillarity of information technology, through the spread of computers, smart devices and the development of the network and digital platforms, have consequences on people's behaviour and on the way in which they communicate, inform themselves, build their beliefs, redefine their behaviors. All these factors contribute to a change that goes beyond digitalization and brings about a transformation of meaning. This causes a significant process-change that requires a rethinking and radical redefinition of concepts, procedures, business, and management model in the sphere of social, political, and economic contexts, as well as within NSIs (Epifani 2020).

This rethinking involves statistical Institutes as institutional subjects, which make a significant contribution to the development of democracy, through official statistics to support public decisions. NSIs are essential actors within knowledge ecosystem because they offer a significant contribution to the strengthening of scientific communication, to the weakening of phenomena such as disinformation and infodemic. NSIs understood the potential of new data sources for statistical purposes: big data make our society measurable, represent a knowledge infrastructure and an opportunity to enrich its offer of information and to respond to the needs of a changing world. The world of NSIs's is mistakenly perceived as a static world with its own rules, specific processes characterized by strict quality criteria. In fact, the world changes the same as the way of doing statistics changes. We went from a world of pre-datification, where NSIs's efforts were focused on data collection, especially surveys and censuses - which have been the main source of data for years - to a world of huge data repositories (Ricciato 2020). In this data deluge,

Massimo De Cubellis, ISTAT, Italian National Institute of Statistics, Italy, decubell@istat.it, 0000-0002-8197-7280

Gerarda Grippo, ISTAT, Italian National Institute of Statistics, Italy, grippo@istat.it, 0000-0001-7568-6252

Referee List (DOI 10.36253/fup_referee_list)

FUP Best Practice in Scholarly Publishing (DOI 10.36253/fup_best_practice)

Massimo De Cubellis, Gerarda Grippo, *Trusted smart statistics: new statistics for decision makers. Istat's experience*, © Author(s), CC BY 4.0, DOI 10.36253/979-12-215-0106-3.22, in Enrico di Bella, Luigi Fabbris, Corrado Lagazio (edited by), *ASA 2022 Data-Driven Decision Making. Book of short papers*, pp. 125-128, 2023, published by Firenze University Press and Genova University Press, ISBN 979-12-215-0106-3, DOI 10.36253/979-12-215-0106-3

NSIs need to extract reliable information from many data sources, as Weinberger suggested in 2011 when he declared "Information represents to the data what wine represents to the vineyard: the delicious extract and distillate" (Weinberger, 2012) Official statistics perform a public service. Data are a collective asset, a public good; recent pandemic events have taught us that having good data allows us to arrive at more effective, timely and citizen-friendly decisions.

To make the choices of decision makers more effectively meet collective needs, new data governance and innovative business models are needed.

The traditional paradigm based on survey data statistics model does not fully meet the new information needs because, concretely, it is neither adaptable to a changing environment nor compatible with the new social infrastructure represented by digital. In this sense, we recall the words with which Mariana Kotzeva, Director General of Eurostat, opened the 13th National Statistics Conference: "Statistics follows life"; NSIs must be in the world if they want to follow and tell life through data. The so-called datafication of society has led to the spread of new data players in both the public and private sectors outside the Official Statistics system. In the pre-datafication world, the NSIs held a monopoly on statistical activity; the only alternative to official statistics was the absence of statistics. Nowadays, NSIs are one of many data-producing entities within the complex information ecosystem. Various actors in the public and private sectors are producing new data and offering alternative statistical viewpoints on emerging phenomena. Official statistics, in competing with other producers of statistics, must keep its institutional role. They must continue to produce official statistics, even with the help of new data, ensuring the same levels of quality, relevance, accuracy, and reliability the same as traditional statistics. In other words, NSIs face a twofold challenge: on the one hand, they must take advantage of the enormous availability of externally produced and collected data, and on the other hand, they must maintain the same high degree of quality in the statistical information they produce. In a context where the amount of information available to users is increasing, it is only the recognition of the quality of the data, and the institutional role of those who produced it, that can enable users and decision makers to navigate an increasingly crowded information ecosystem. The relevance of statistical information, its timeliness and usability, are crucial for building a relationship of trust with users. The response that official statistics gives to all these questions to stakeholders, decision makers, and users is Trusted Smart Statistics, an expression coined by Eurostat to indicate the mature stage of producing statistics with big data. The new model for European statistics involves greater integration of the information produced and strong use of statistical registers and big data—a holistic approach that aims to provide new, more effective and efficient tools to support decision makers. The starting point of this strategic path is the Scheveningen Memorandum, sanctioned within Eurostat in 2013. This memorandum, formalized the need for all European Statistical Institutes to consider big data sources as new sources for official statistics, launching experimental projects aimed at understanding how to exploit the big data potentiality. In fact, the European Statistical System network (ESSNet) has implemented several experimental projects such as: Essnet Big Data I, Essnet Big Data II, ESSnet Towards Trusted Smart Statistics, and Essnet Smart Surveys. The use of new data sources has been for several years at the centre of the European NSIs agenda. It has required in all NSIs an experimentation phase for studying appropriate methodologies to exploit the use of big data sources, considering the issues related with privacy constraints. Eurostat enables and contribute to these activities in both the design and execution phases, within the framework of official statistics innovation. The term Trusted Smart Statistics was proposed by Eurostat to represent the evolution of traditional statistics and was officially adopted by the European Statistical System in the Bucharest Memorandum on October 12, 2018 during the 104th Directors General of the National Statistical Institutes (DGINS) conference. The Bucharest Memorandum helped to enhance and formalize the contribution of big data in terms of validity, accuracy, and reliability of outputs. The term Smart Statistics refers to multi-source and multi-output statistical production systems that use innovative technologies aimed at flexibly integrating big new data sources into statistical production. The reliability of statistics, to which the term trust refers, is closely linked to the reliability of the institution that produces them. It is based on: (i) compliance with standards for data processing and privacy; (ii) infrastructure that enables data processing (iii) methodological characteristics; (iv) quality guarantees of the entire

production processes. Historically, NSIs have always had the full control of the entire statistical production process because it took place totally in-house. The in-house management of the statistical production process from the direct collection of data from respondents to the dissemination of the statistics produced enabled NSIs to be able to guarantee the reliability, quality, and relevance of the data collected and the methodologies applied, and all the standards necessary for the statistics produced to be called “official”. The production of statistics with the new data sources, often require the use of data collected and held by third parties (e.g.: mobile phone operators). However, it is necessary to maintain the same levels of quality and the same characteristics that make it possible to be able to ensure the official nature of the statistics produced and the trust that users have with respect to the institutional role of NSIs and their statistics. To maintain this level of confidence and ensure the same levels of quality and relevance as traditional surveys, an adjustment of the entire statistical production process is essential. If some steps in the process (data collection and processing) are external to the NSIs, they must still be designed and controlled by the NSIs themselves.

The release of the first outputs with the use of big data, experimental statistics, and the comparison between different NSIs within the EssNet projects, has made NSIs aware on the potentiality of new data sources. The use of these data not only requires strictly technical capabilities and more powerful IT infrastructure, but also requires investment in the different areas of which individual organizations are composed (methodological, organizational, legal). For example, new data sources require the following new methodological approaches: (i) to transform raw data into statistical information and concepts; (ii) to use data that were not designed and collected for statistical purposes; (iii) to overcome coverage issues; (iv) to integrate new data sources with traditional ones. The character of timeliness and temporal and spatial granularity of TSS will enable policy makers to make decisions based on much richer data than those produced with traditional statistics. New data sources have an impact both within individual organizations, in terms of organizational adjustment, and externally, in terms of the ability to produce new products to support public decision makers more effectively. TSS enable decision makers to have more timely access to data and statistics in different sectors; up-to-date statistics also enable decision makers to implement government policies with more accurate spatial detail. The official nature of TSS would give decision makers an opportunity to put new phenomena on their policy agenda. In addition, TSS helps to give a new role to citizens. We said that big data make society measurable, put humans at the centre, create a new digital humanism, and can give rise to citizen statistics as new processes institutionalized by the new social infrastructure represented by digital. TSS become the product of a trust-based exchange between citizen and NSIs. Citizens become, through their daily online and offline actions, "measurable": they become data producers and statistical users, at the same time. Through their active participation in smart surveys, they can provide smart data to support the production of TSS. In order to enhance the role of citizens, it is appropriate for NSIs to establish a "social pact" with the citizens themselves, enabling the NSIs to collect data from citizens and return it to them in the form of useful information.

3. Istat’s experience

After the adoption of the Bucharest Memorandum, a reflection began in Istat on how to govern this innovation process. The release of the first outputs with the use of big data, experimental statistics, the debate between the different NSIs in the European Statistical System within the EssNet projects, has made it possible to acquire the awareness that the use of new data sources, it not only requires strictly technical skills and more powerful IT infrastructures, but requires investments in the various sectors that make up individual organizations. For Istat, the production of TSS represents a highly innovative strategic objective, both scientifically and organizationally for the following purposes: (i) enrich the supply of information in terms of timeliness and spatial granularity; (ii) efficiency, due to the automated integration of data sources and flows; (iii) the ability to capture new phenomena that cannot be measured by surveys or administrative sources alone; (iv) provide answers to stakeholders; (v) reduce the statistical burden on respondents; (vi) train or integrate new skills needed to extract information from the new data sources to contribute in a coherent way to the building of a new organizational model. All these factors are crucial to enhancing the relevance and reputation of Official Statistics through recognition of the unique role of

Istat in terms of the reliability of the statistics produced and the transparency of the production processes. Istat has established a specific Center with the purpose of guiding statistical production activities toward the Trusted Smart Statistics production system. The Center is an agile organization, whose interdepartmental character makes it possible to overcome organizational fragmentation. It represents the point of connection and monitoring of all activities aimed at building the new production system. An internal Steering Committee, consisting of the Institute's top management and responsible for the process of TSS strategic analysis, heads the Center. The strategic decisions on activities and investments, are formalized in a Roadmap. The Roadmap is a strategic document containing the planning of activities aimed at building the TSS production system. At the organizational level, the TSS Center has the task of designing a sustainable organizational structure to help and support the individual company components in working in a systemic, coherent, and synergistic way with the aim of creating the new Trusted Smart Statistics production system. This means that each organizational dimension is involved in the changing of the business model process: legal, methodological, communication, strategic planning, human resources, and skills development. All "cross-divisional" directorates must support the TSS production system to ensure its functioning and to ensure the release of new products with the same level of reliability and quality as traditional statistics. In Istat, to monitor the adjustment of individual organizational dimensions, a monitoring and guidance framework aimed at measuring the organizational maturity status of individual components of the "new" Trusted Smart Statistics has been implemented as a new production system. This tool will support the monitoring phases of the actions, implemented by the Istat individual organizational structures, on the strategic and operational levels, aimed at building the TSS system. The results of the first monitoring revealed how important the organizational component is. In addition to highly technical factors such as IT infrastructures and sound methodological systems, it is important that the new paradigm is supported by organizational changes at various levels. Communication, the legal sector for the definition of aspects related to data access and the ethical use of data, the human resources sector, the planning of objectives are all dimensions involved in the implementation of the new production system.

At the European level, there is an intense debate on this issue.

Statistical Institutes have already achieved promising results but are now facing new challenges, which require ever stronger interactions and collaborations with other public and private actors. These are paths already undertaken, but which must necessarily be followed to the end in order to ensure that the wealth of data produced that we all produce daily can be transformed into statistical information that can be trusted and become a common good.

References

- Epifani S. (2020). *Sostenibilità digitale*, Digital Institute, Book
- European Statistical System Committee - Eurostat (2018). Bucharest memorandum on official statistics in a datafied society (Trusted Smart Statistics), Memorandum
- European Statistical System Committee - Eurostat (2013). Scheveningen memorandum on big data and official statistics, Memorandum
- Floridi L. (2017). *La quarta rivoluzione*, Raffaello Cortina Editore, Book
- Mayer-Schönberger V., Cukier K. (2013). *Big data. A revolution that will transform how we live, work and think*, Garzanti, Book
- Ricciato F., Wirthmann A., & Hahn M. (2020). Trusted Smart Statistics: how new data will change official statistics. *Data & Policy*, 2, E7, Journal Article
- Weinberger D. (2012). *Too big to know*, Basic Books, Book