Italian National Statistical system: Quality Reporting matters at every level

Andrea Bruni[1], Paola Giordano[2]

[1] Andrea Bruni, Istat, Rome, Italy, anbruni@istat.it

[2] Paola Giordano, Istat, Rome, Italy, pgiordano@istat.it

Abstract[1]

In recent years, Istat, in its legislative role of coordinating the National Statistical System (Sistan), carries out initiatives to support the improvement of the quality of statistics produced within the Sistan, integrated into the quality policy adopted by Istat and consistent with the international reference framework on data and metadata quality in official statistics. Regarding metadata quality, in 2018 the Steering and Coordinating Committee for Statistical Information (Comstat), Sistan's main governance body, released the Guideline Act No. 3 and in particular Article 4, which states that the clear and comprehensible dissemination of data requires metadata describing both the process of production and the characteristics of the statistical results obtained. Additionally, in 2021 both the Comstat and the Cogis, short for Commission on Quality Assurance of Statistical Information (Sistan's main vigilance body), suggested that Istat should implement two documents related to quality, the target recipients being the Not-Other National Authorities entities. The Italian version of the European Statistical System (ESS) Quality Assurance Framework has been published in October 2023, and the current work is aimed at defining an Italian "Handbook for Quality and Metadata Reports", version of the "ESS Handbook for Quality and Metadata". Such Handbook may be considered a strategic operational document to guide Sistan entities toward the production of standardized quality reports on statistical processes and outputs, for facilitating the comparability of the meta-information among statistical processes and among Sistan entities. Therefore, the Handbook will serve as a reference, within the Sistan framework, for the proper collection of the set of metadata, accompanying the dissemination of aggregated data for their proper use and interpretation. The drafting of the Handbook is going to be accompanied by the definition of the customized template serving the reporting purpose. Such template, based on the ESS standard SIMS (Single Integrated Metadata Standard), is planned to be used as a Metadata Structure Definition in the Istat SDMX Toolkit, so to publish quality reports as SDMX-compliant messages. For all these work strands, a Task Force has been set up. It is following several steps: 1) analysis of existing practices at Sistan entities on quality reporting, in particular structural and reference disseminated metadata; 2) mapping of the metainformation available at Sistan Entities against the SIMS concepts; 3) design of the items to be included in the standardized quality report; 4) preparation of the Handbook with the operational guidelines for the Italian quality report to be compiled.

Keywords: NSS, dissemination, SIMS, Handbook, SDMX.

---

[1] *This paper resumes the outcomes of a work jointly carried out by the authors. However, Section 1 and 2 is attributable to Andrea Bruni, and Section 3 and 4 to Paola Giordano.*

1. Introduction

Statistical Offices supply official statistics to international bodies, institutions, businesses, and citizens. The dissemination of statistical data necessitates providing readily available information on the sources, classifications, methodologies, and calculation procedures adopted. It entails ensuring simultaneous access to both data and metadata by all potential users, with maximum ease of accessibility. In essence, statistics should be disseminated in a clear and comprehensible manner, presented to offer the most impartial interpretation possible and to facilitate meaningful comparisons across time and space. To achieve this goal, users should be provided with a comprehensive set of concepts, definitions, classifications, and metadata, enabling them to navigate and interpret the data accurately.

The third round of Peer Review has taken place at Istat from 28th November to 2nd December 2022: in the final report (Eurostat, 2023), one of the improvement actions recommended by the team of independent international experts was "[…] *to refine the tools for the coordination of the National Statistical System. This includes the implementation and monitoring of standards, the provision of methodological advice, the provision of metadata, the dissemination of statistics produced across the National Statistical System* […]".

While the recommendation primarily targets National statistical authorities in their pursuit of compliance with the European Statistics Code of Practice (CoP), it is essential to note that the Istat approach, consistent with its overarching quality policy, extends beyond mere adherence. Instead, Istat interprets inputs from the European context in a broad sense, thereby embracing a comprehensive application of the recommendations, to the extent feasible, across the entire Sistan (the Italian statistical system).

By doing so, Istat aims to foster a culture of excellence and integrity in statistical practice across all levels and domains, thereby enhancing the reliability, credibility, and relevance of statistical information produced within the full Italian statistical landscape.

2. Italian quality framework

As of December 31, 2023, Sistan encompasses 3,305 offices, including 14 entities, namely Istat and the Other National Authorities (ONAs), which are dedicated to adhering to the European Statistics Code of Practice, the Quality Assurance Framework of the European Statistical System (ESS), and the reference metadata reporting standards set within the ESS, including the Single Integrated Metadata Structure (SIMS) v2.0 and the ESS Handbook for Quality and Metadata Reports.

Istat, acting as the coordinator of this quite complex system, has progressively intensified its efforts to implement support initiatives aimed at enhancing the quality of statistics produced by

the other statistical offices within Sistan, excluding those involved in producing European statistics. Consequently, a strategy has been developed for these entities, referred to as non-Other National Authorities (not-ONAs), with the objective of adapting the tools from the ESS quality framework to suit the specific needs of the local context in Italy.

In accordance with the National Statistical Law, the coordinating entity of Sistan is the Italian National Institute of Statistics (Istat), overseen by the Steering and Coordinating Committee for Statistical Information (Comstat), chaired by the President of Istat. Given its pivotal position, Istat assumes an indispensable guiding role, particularly in providing technical assistance and supervision to ensure the compliance of produced and disseminated statistics with quality standards. Additionally, the Commission on Quality Assurance of Statistical Information (Cogis), serving as an advisory body, oversees all activities within Sistan.

Through these mechanisms, Sistan benefits from centralized leadership, expert guidance, and rigorous quality assurance protocols, ultimately enhancing the credibility and reliability of official statistical outputs.

Ensuring the quality of statistics stands as a cornerstone for fostering trust in public statistics. An effective governance system for official statistics must establish standardized and recognized requirements, facilitating the identification of mechanisms to control statistical processes, measure quality, and communicate it to external users.

Aligning with the ESS standards entails leveraging established mechanisms.

The strategy adopted involves extending the European framework to encompass the "Sistan-not-ONA" sphere, thereby customizing the European quality architecture to domestic requirements. This concerted effort towards convergence has culminated in the development of various documents, as outlined in the following table.

Table 1: ESS documents and parallel Italian framework.

ESS reference document	Corresponding Italian document	Release Year (Italian document)
European Statistics Code of Practice	Codice italiano delle statistiche ufficiali (Italian Code of official statistics)	2010
European Statistics Code of Practice - revised edition 2017	Codice italiano per la qualità delle statistiche ufficiali (Italian code for the quality of official statistics)	2021
Quality Assurance Framework - revised edition 2019	Guida per l'implementazione del Codice italiano della qualità delle statistiche ufficiali (Guide for the implementation of the Italian Code for the quality of official statistics)	2023

Moreover, in 2018, the Comstat issued Guideline Act No. 3, a significant directive within the system (Comstat, 2018). Of particular relevance is Article 4, which underscores the importance of metadata quality in the dissemination of statistical data. As highlighted in the article, the emphasis on metadata and quality reporting underscores the importance of providing clear and comprehensive data dissemination. By including metadata that describe not only the production process but also the characteristics of the statistical results obtained, the Committee demonstrates its commitment to transparency in official statistics across the entire Sistan. By prioritizing clarity through robust metadata and quality reporting practices, the Committee promotes trust and confidence in the integrity of official statistics, ultimately enhancing their utility and value to users.

As part of the ongoing effort to replicate locally the ESS quality framework, both the Comstat and the Cogis have recommended a significant step for Istat: the development of an "interpretative document," to be called the Italian Handbook for Reports on Quality and Metadata. This document, of operational nature, seeks to provide guidance to not-ONAs entities within Sistan to produce standardized reference metadata reporting on the quality of processes and statistical outputs.

Specifically, drawing upon the ESS standard "Handbook for Quality and Metadata Reports" (revised edition 2021), the objectives of this Italian document are threefold:

- to harmonize the documentation produced for different statistical processes and by different Sistan bodies;
- to facilitate use and comparison pertaining to both statistical processes and outputs;
- to ensure clear and standardized dissemination of data, enabling their proper utilization and interpretation.

By aligning documentation practices and ensuring consistency in reporting, this handbook seeks to enhance the quality and usability of statistical information produced within the Sistan framework, thereby promoting better adherence to the Italian Code for the quality of official statistics. In addition, in its efforts to enhance the tools for disseminating official statistics, Istat has made significant strides, notably through the development of a native SDMX open-source software. This innovative framework has enabled the implementation of the "Public Statistics Hub" project, a centralized platform for the dissemination of aggregate data produced by the Sistan. The web portal offers users convenient access to a wide array of statistical information, meticulously organized and harmonized with respect to structural metadata.

Regarding the management of reference metadata, the software provides a dedicated module for handling Metadata Structure Definitions (MSDs), enabling the compilation and publication

of quality reports in SDMX-ML v2.1 format, according to the functional flow shown in the following diagram.

Figure 1: Functional flow for producing quality reports
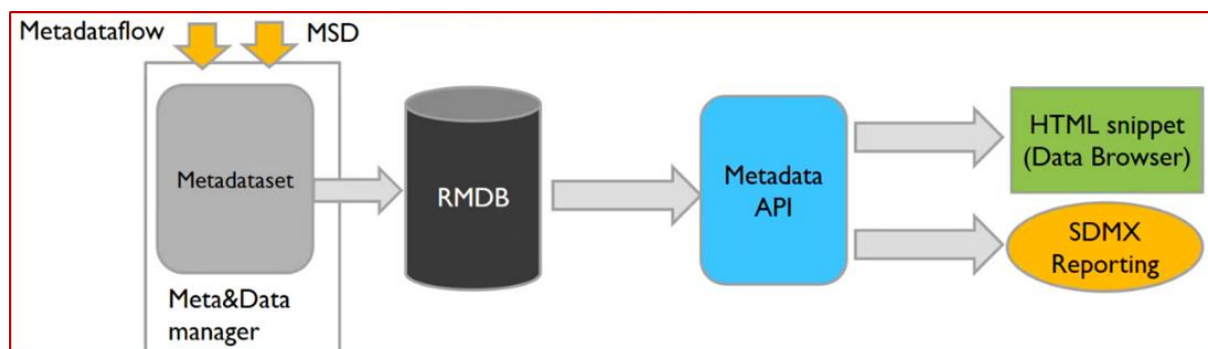
In detail, when a metadata flow referencing an MSD is created, it becomes feasible to generate a metadata set, which is then stored within a designated Reference Metadata Database (RMDB). Within this service-oriented architecture, the RMDB is consumed by a Metadata API provided by the software, offering two distinct functionalities: retrieving SDMX messages for various purposes (such as quality reporting), or visualizing/disseminating quality reports through an HTML page using a specific template via an associated web client, or as a standalone object representing meta-information.

3. Methods

Building upon the components delineated in the preceding paragraphs, a Task Force has been instituted at Istat in September 2023, with the purpose of drafting the Italian Handbook. The Task Force is responsible for organising and following the activities listed below, taking into account the relevant deadlines:

- Analysis of existing practices among entities within Sistan concerning quality reporting, with a focus on structural and referential metadata – March 2024.
- Mapping of available meta-information from Sistan entities related to concepts outlined in the European standard "SIMS - Single Integrated Metadata Structure" – March 2024.
- Designing contents to be incorporated into the standardized quality model (quality report) – July 2024.
- Preparing the "Handbook for Quality Reporting in Sistan", providing operational guidelines for compiling the quality report in accordance with the standardized format – December 2024.

Additionally, the Task Force will assess solutions to standardize metadata dissemination across the "Hub of Public Statistics" platform, aiming to identify methods that minimize the reporting burden on Sistan entities while aligning with solutions already implemented for quality reporting within IstatData, the Istat corporate dissemination datawarehouse.

The first two activities have been completed. The analysis began with the National Statistical Programme (NSP), comprising 805 statistical works of various typologies:

- 714 Statistics (274 from Istat and 440 from other Sistan entities);
- 62 Design studies (36 from Istat and 26 from other Sistan entities);
- 29 Statistical Information Systems (10 from Istat and 19 from other Sistan entities).

Istat experts have been analysed 379 out of the 440 statistics (focusing on the most relevant non-Istat ones), starting from the data sheet that each Entity must compile to update its records within the NSP online system. Subsequently, experts searched for these surveys on the official websites of the Sistan entities responsible for each survey, to gather all available existing metadata. Furthermore, for each survey, metadata were mapped with SIMS concepts, to enhance understanding of the information needs and relevance for each entity.

4. Results and conclusions

As stated above, starting from 379 statistics, Istat experts have been able to find metadata for 213 of them (56.2%) on the scouted official websites. The mapping activity with the SIMS concepts is reported in Table 2 below. It's important to note that only SIMS items having 3 or more occurrences are reported.

Table 2: Occurrences of SIMS concepts in Sistan disseminated metadata.

ID	Concept Code	Concept Name	n.
S.1.1	CONTACT_ORGANISATION	Contact organisation	40
S.1.2	ORGANISATION_UNIT	Contact organisation unit	34
S.1.3	CONTACT_NAME	Contact name	26
S.1.4	CONTACT_FUNCT	Contact person function	17
S.1.5	CONTACT_MAIL	Contact mail address	17
S.1.6	CONTACT_EMAIL	Contact e-mail address	28
S.1.7	CONTACT_PHONE	Contact phone number	21
S.2.3	META_LAST_UPDATE	Metadata last update	3
S.3.1	DATA_DESCR	Data description	202
S.3.2	CLASS_SYSTEM	Classification system	92
S.3.3	COVERAGE_SECTOR	Sector coverage	39
S.3.4	STAT_CONC_DEF	Statistical concepts and definitions	148
S.3.5	STAT_UNIT	Statistical unit	152
S.3.6	STAT_POP	Statistical population	119
S.3.7	REF_AREA	Reference area	116
S.3.8	COVERAGE_TIME	Time coverage	113
S.3.9	BASE_PER	Base period	7
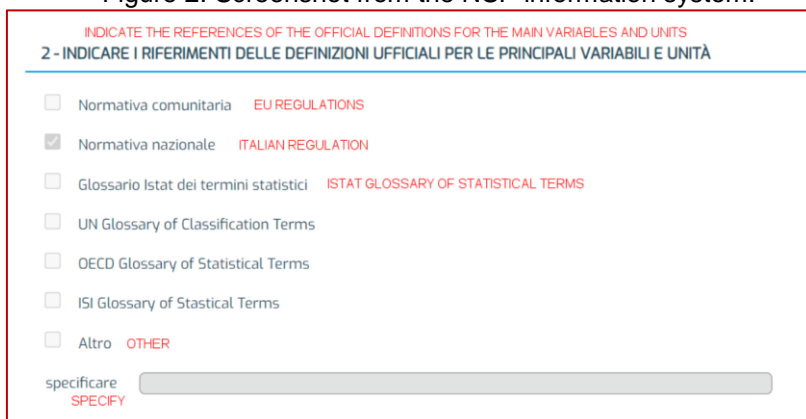
ID	Concept Code	Concept Name	n.
S.4	UNIT_MEASURE	Unit of measure	48
S.5	REF_PERIOD	Reference period	46
S.6.1	INST_MAN_LA_OA	Legal acts and other agreements	92
S.6.2	INST_MAN_SHAR	Data sharing	4
S.7.1	CONF_POLICY	Confidentiality - policy	11
S.7.2	CONF_DATA_TR	Confidentiality - data treatment	5
S.8.1	REL_CAL_POLICY	Release calendar	40
S.8.2	REL_CAL_ACCESS	Release calendar access	8
S.8.3	REL_POL_US_AC	User access	3
S.9	FREQ_DISS	Frequency of dissemination	52
S.10.1	NEWS_REL	News release	71
S.10.2	PUBLICATIONS	Publications	69
S.10.3	ONLINE_DB	On-line database	102
S.10.4	MICRO_DAT_ACC	Micro-data access	10
S.10.5	DISS_OTHER	Other	8
S.10.6	DOC_METHOD	Documentation on methodology	29
S.10.7	QUALITY_DOC	Quality documentation	4
S.11.1	QUALITY_ASSURE	Quality assurance	10
S.11.2	QUALITY_ASSMNT	Quality assessment	12
S.13.1	ACCURACY_OVERALL	Overall accuracy	11
S.13.3.3	NONRESPONSE_ERR	Non response error	4
S.13.3.3.1	UNIT_NONRESPONSE_RATE	A4. Unit non-response - rate for P	3
S.14.1	TIMELINESS	Timeliness and TP2. Time lag - final results for U	42
S.15.1	COMPAR_GEO	Comparability - geographical	3
S.18.1	SOURCE_TYPE	Source data	132
S.18.2	FREQ_COLL	Frequency of data collection	58
S.18.3	COLL_METHOD	Data collection	162
S.18.4	DATA_VALIDATION	Data validation	56
S.18.5	DATA_COMP	Data compilation	35
S.19	COMMENT_DSET	Comment	8

The findings of this prospectus will serve as pivotal factors influencing forthcoming decisions of the task force, towards conclusively defining the establishment of the Italian MSD, a process driven by the meticulous selection of SIMS concepts. Simultaneously, there is an exploration underway to develop national concepts that complement those within SIMS.

For instance, considerations include the potential inclusion of dedicated entries addressing mandatory response obligations and the role of the data protection officer, recognizing the criticality of these aspects. Additionally, further ongoing hypotheses under scrutiny involve the extraction of pertinent insights from SIMS Section 18 "Statistical Processing", aiming to amplify their visibility and strategic significance within the framework.

One operational solution under consideration involves leveraging the information already collected by Istat within the NSP information system. This platform, exemplified in Figure 2, serves as a repository where managers of Sistan statistical offices furnish comprehensive data crucial for the definition of the National Statistical Programme.

Figure 2: Screenshot from the NSP information system.

Implementing this practice would alleviate the burden for quality report compilers. It entails providing a pre-compiled template where a significant portion of the required information is already included, requiring only confirmation from the compilers. This streamlined approach not only enhances efficiency but also contributes to the sustainability of the Sistan quality reporting project in the long term.

Acknowledgment

References

Comstat (2018). Atto di indirizzo (Guiding Act) n.3.
https://www.sistan.it/fileadmin/Repository/Home/NORME_E_PROCEDURE/ORGANIZZAZIONE_E_FUNZIONAMENTO/UFFICI_DI_STATISTICA/Comstat_Atto%20di%20indirizzo%203_2018.pdf

Eurostat (2023). Italy Peer Review Report on compliance with the European Statistics Code of Practice and further improvement and development of the national statistical system https://www.istat.it/it/files//2021/10/peer-review-report-2023.pdf