

Connecting quality and planning systems in the Italian National Statistical Office

Simona Pace, Riccardo Carbini, Concetta Pellegrini, Maria Cara

Abstract The quality system and the planning systems have been implemented in Istat almost in an independent way, connected ex ante only by the sharing of basic information on statistical processes carried out by Istat. The products of the two types of system have been used in a synchronized way to respond to specific internal and external needs. An activity of integration and re-use of information between systems ex post has been carried out, and efforts were made to meet planned and unplanned needs. The recent increase of these needs, due to internal and external factors, pushes for a reorganization of the role and connections among these systems. Hence the design of a well connected meta-system - including the quality system, the planning systems and the process map system - can provide opportunities to create a virtuous circle in order to realise high quality and effective statistical products.

Key words: Data quality, planning system, statistical process map

1. Introduction

Statistical quality of European statistics is a fundamental requirement, as stated in article 12 of European Community Regulation No 223/2009 on European statistics. In addition, the above mentioned regulation lists a series of statistical principles, such as professional independence, impartiality, objectivity, reliability, statistical confidentiality and cost effectiveness, that should govern the development, the production and the dissemination of European statistics.

¹

Simona Pace, Istat, Via C. Balbo 16, 00183, Rome, sipace@istat.it

Riccardo Carbini, Istat, Via C. Balbo 16, 00183, Rome, carbini@istat.it

Concetta Pellegrini, Istat, Via C. Balbo 16, 00183, Rome, pellegrini@istat.it

Maria Cara, Istat, Via C. Balbo 16, 00183, Rome, cara@istat.it

Hence the necessity of evaluating the way to comply efficiently with these interconnected requirements, such as quality, reliability, cost effectiveness, etc

2. The Istat Quality System

Istat has contributed to the definition of the conceptual framework of quality of the European statistical system in the last 10 years. The European Code of Practice (Eurostat, 2005) and the Recommendations of the LEG on Quality (Lyberg L., 2001) are the pillars of the European policy for quality.

Istat has developed a systematic activity regarding the quality of statistical processes according to the European framework. The statistical processes are divided in two groups, the new and the current ones. A procedure has been defined to assure that the project of the new process was consistent with related sound methodologies, defined relevance, valid and tested collection mode and adequate methods to prevent sampling and non-sampling errors.

The evaluation of the quality of the current processes is supported by a double activity: First, the processes are documented by collecting process metadata in a central system (Information System for Survey Documentation – SIDI). This activity is carried out by a network of quality pilots, ad hoc trained. Process metadata are metadata on process content, its operational characteristics and the quality considered both in terms of activities of prevention, monitoring and evaluation of errors (quality actions) This information is available to external users through a section of Istat website, siquat.istat.it, to guarantee the transparency of Istat statistical production.

Secondly, a set of standard quality indicators, oriented to monitor a process, such as timeliness and punctuality, coverage, non response, edit and imputation, revision, consistency between preliminary and final result indicators, are collected (Brancato et al, 2004). This set is used to monitor broadly the quality of statistical process. Analysis of trend of timeliness, coverage, non response according to specific metadata such as type of process and periodicity have been described before (Brancato et al, 2005), as well as analyses of editing and imputation procedures tailored to internal and external users (Brancato et al, 2009).

3. The Istat Planning Systems

Istat runs a planning system that consists of two parts: the National Statistical Programme (PSN) that includes the statistical production of public agencies (including Istat) and the Istat internal planning system.

The legislative decree n. 322 of 6 September 1989 "Provisions on the National Statistical System and re-organisation of Istat" commissions Istat to draw up the PSN, in order to coordinate the statistical production of all the public agencies. The PSN identifies the surveys, the secondary and preliminary studies that the offices and the bodies within the National Statistical System (Sistan) have decided to carry out during a three-year period. Istat statistical production is included in PSN.

The internal planning system includes the Strategic triennial Planning (PST) and the Annual Plan of Activities (PAA), logically connected.

PST includes all the planned objectives, statistical and non statistical, and represents the basis for the resources acquisition plan (personnel, services, IT).

PAA includes the objectives and the assigned resources in a reference year. All the activities in PAA are monitored step by step, from their beginning to their complete realization.

4. Statistical Process Mapping - MaPros

The statistical process map system - called MaPros - is the result of a specific request coming from ISTAT top management. The request is based on the necessity to explore all statistical processes in a full integrate way, highlighting the links with the others existing processes and taking into account temporal and institutional constraints.

The system is also aimed at evaluating and assessing priorities on the statistical production depending on the interconnections between processes, as well as the legal and financial constraints.

MaPros aimed primarily at supporting the planning operations: for this reason, integrated information is timely collected and available. These characteristics qualify MaPros as something more than a classical system of metadata.

Part of the information in the statistical process map system comes from planning and statistical information systems already existing in Istat: in addition, new information is ad hoc collected in order to deep and highlight links and interconnections between statistical processes. Thus MaPros gives the opportunity to consider the production process not only in an isolated way, but within an integrated system of information exchange.. This is the principal element of novelty introduced by MaPros in Istat planning system. In this way it is possible to analyze the contribution, and also the importance, of each statistical process in the broader context of interconnections between different processes.

The implementation of MaPros is based on maximizing the level of integration among existing systems and minimizing the respondent burden on the production structures.

MaPros does not constitute a new informative system, but rather a system that collects and combines information already available from statistical and administrative archives, returning information in addition to those provided by individual systems. From this point of view, MaPros represents a connecting system, with its own characteristics, from two integrated components: one part consists of systems characterized by a planning character (National Statistical Programme - PSN and the Planning system - PAA); the other is composed of quality and documentation systems.

MaPros is powered by statistical production process, in particular Survey (SDI), Register Statistics (SDA) and Secondary studies (SDE).

For each statistical production process in a given year we can consider four groups of information:

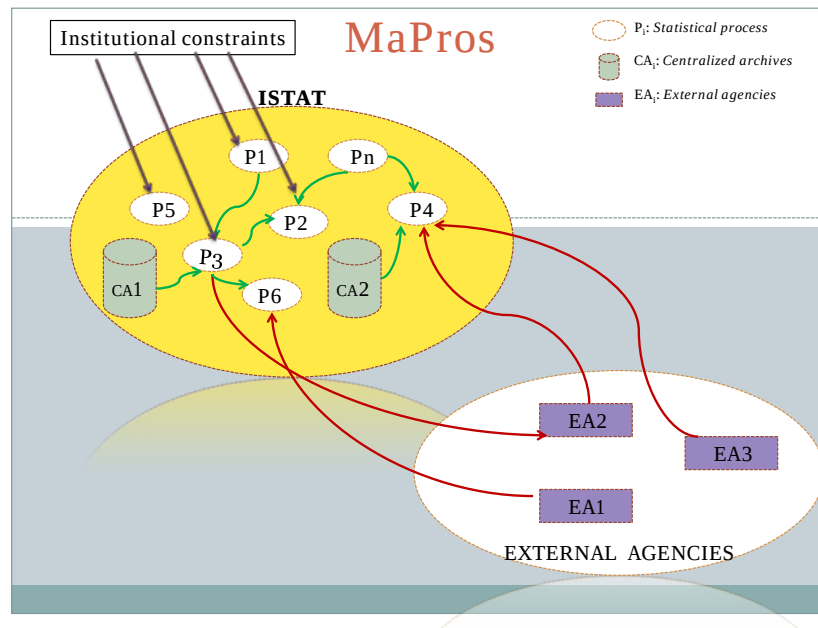
- 1 Structural data (code PSN, code PAA, type of process¹, etc.);
- 2 Institutional constraints (European and National regulations, relationships with other institutions regulated by formal agreements, etc.);
- 3 Internal and/or external input (with description of these inputs, methods and timing of use);
- 4 External output (archives or products other than scheduled dissemination, for regular external entities or products for subjects of extreme relevance).

The first section derives from existing information systems. The second is a study of data from other informative systems. The third and fourth parts contain new information.

The following Figure 1 shows a graphic representation of interconnections between statistical processes and institutional constraints inside MaProsirst section is derived from existing information systems. The second is a study of data from other informative systems. The third and fourth parts contain new information.

The following Figure 1 shows a graphic representation of interconnections between statistical processes and institutional constraints inside MaPros.

Figure 1: MaPros



1 Reference year; Project; Type of process; Identification code for each process (ex-code PST); Noun of process; PSN code; Type of process; Reference operative structure (Department/Division/Office); Sector; Sub-Sector; Institutional constraints ; Available human resources by level and skill; Additional human resources by level and skill.

In this figure, P_i represents a statistical process, CA_i centralized registers of Istat and EA_i public and private external institutions. The red arrows illustrate inputs and outputs from/to external bodies; the green arrows represent exchange inside Istat; the purple arrows are the institutional constraints that insist on statistical processes. For example, P_3 is a statistical process subjected to institutional constraints (at European or national level), it uses as inputs the centralized register CA_1 and the process P_1 from Istat, the register EA_2 from external institutions; moreover P_3 is used as input by processes P_2 and P_6 .

The overall representation of production processes in terms of internal and external actors and exchanged information flows is conceptualized from a theoretical point of view in the input/output model.

Table 1 illustrates MaPros in terms of input/output model referred to statistical production processes scheduled in 2010.

Table 1: Scheduled statistical production process in 2010

Suppliers	Users												
	DIV1	DIV2	DIV3	DIV4	DIV5	DIV6	DIV7	DIV8	DIV9	Istat	Private agencies	Public agencies	External agencies
DIV1	1		15		1		1	1		19	1	14	15
DIV2								1		1			1
DIV3			31		1		2	3	14	51		88	88
DIV4			18	36	3		2			59	2	41	43
DIV5					1					1		5	5
DIV6				10	9	19			1	39		38	38
DIV7	1		27		3		34		1	66	10	61	71
DIV8	1		76		2	1	20	25	22	147	11	59	70
DIV9			1		1					2		10	10
Istat	3		168	46	21	20	59	30	38	385	24	316	340
Private agencies	8		27	6	1	11	12	13	4	82			
Public agencies	13	4	52	8	15	40	23	23	7	185			
External agencies	21	4	79	14	16	51	35	36	11	267			
ALL	24	4	247	60	37	71	94	66	49	652			

In the rows of the table internal suppliers (DIV1-DIV9 represent the Divisions of production) and the external suppliers (private and public institutions) are shown. Internal (DIV1-DIV9) and external users (DIV1-DIV9) are in the columns of the table. The exchanges at home and abroad in terms of informative flows are represented inside

the table. For example, the DIV3 receives 247 different informative flows: of these, 168 come from sectors of Istat (in particular 76 from DIV8), 79 from outside (52 from public institutions). If we consider the side of users, DIV3 supplies 139 flows of output of which 88 outside: of the 51 that it supplies to fields of Istat, 31 are flows of information exchanged to the division itself.

The process map allows to perform analysis at different levels: divisions, sectors of analysis, single process, explaining in detail the characteristics of each considered process.

MaPros consists of three subsystems: *navigation*, dedicated to the representation of individual processes; *reporting*, to analyse data in more aggregate reports; *management*, dedicated to collect the information in the system.

As previously said, MaPros feeds itself, at least partly, with information coming from existing data bases in Istat. Particularly, structural information originate from planning systems.

The management of the information of MaPros is consistent with the standards adopted by SIDI. From the contents point of view, the adopted definitions and the thesauri assure the coherence among both the systems. Besides, a general comparison has been realised with the information collected by Sidi too, allowing the elimination of possible duplications or overlaps of information in a view of integration of the collected information.

The added value of MaPros consists not only in providing new information on the interconnections between statistical production processes, but also in building a "place" of comparison and of treatment of the information collected from the different systems, and the point of confluence and restitution to the other systems of information "standardized, normalized and integrated."

5. Connections and opportunities

The design of a well connected meta-system including the quality system (SIDI), the planning systems (PEC, PSN) and the statistical process map system (MaPros) gives opportunities to create a virtuous circle in order to support the production of high quality and effective statistical products. The main efforts are obviously oriented to the resource planning and activity scheduling. But today the frequent necessity of changes in planned activities, calls for a development of the connection between systems. Indeed, these changes require analyses of their consequences, in terms of the impact on other processes, (in case of a modified or suspended process previously used as a statistical source), and specific analyses of the overall effect on the quality of the statistical results.

The quality and planning systems have been developed for different purposes and with different technologies. Today the technological infrastructure is homogeneous with each other, thanks to the implementation of technical standards adopted by the Institute: for example, all quality and planning systems are using an Oracle relational database, and this facilitates the integration of the data among them.

The integration of the technological infrastructure is a fundamental starting point; the "philosophy" that inspires the implementation of MaPros is another step.

Today there are good opportunities on the floor.

The short-term objective could be the development of an higher degree of integration among quality and planning systems, allowing the reuse of information.

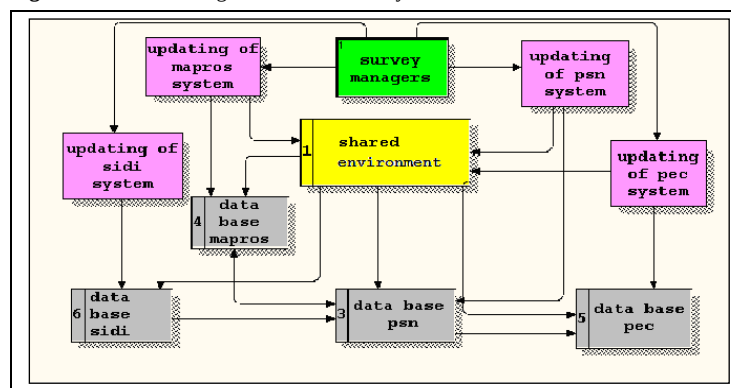
The first step, of course, is to continue the analysis of the information that are collected in the different different systems in order to compare and evaluate the opportunity of harmonization and to explain and manage the differences.

The workflow between the systems, in new design, should be the following:

- The PSN collects the provisional information of the processes.
- The PAA system gets the meta information collected in the PSN e gives the data of the provisional management costs of the processes.
- The MAPROS system shares the PAA and PSN information and it manages the input and output flows of the processes, helping the Istat top management to evaluate the planned processes.
- The SIDI system gets all the information of interest collected in the other systems and tests it out for confirmation, validation and integration purposes.
- The information collected in SIDI returns to the other systems (i.e. validated process metadata) as input for the planning systems in the following years.

The main problems are the complexity of each system that manages a great number of information and their constrains in supporting different needs. Another problem is the assessment and management of the validity period of every single information.

Figure 2: Dataflow of agents and data base system



A long-term objective could be the implementation of a centralised environment containing all the shared data according to the thesaura and the classification variables, accessible from the different systems to assure an updated and univocal codification of information on processes. The dataflow of agents and data base system is reported in figure 2.

The integrated management among the different systems might guarantee coherence of information, efficiency and above all reduced burden for the survey managers.

References

- Brancato G., Carbinì R., Simeoni G., Metadata and quality indicators to report on editing and imputation to different users, UNECE, Work Session on Statistical Data Editing, Neuchâtel, Switzerland (2009)
- Regulation (EC) No 223/2009 of the European parliament and of the council (2009) <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:087:0164:0173:en:PDF>
- Brancato G., Carbinì R., Pellegrini C., Signore M., and Simeoni G., Assessing Quality through the Collection and Analysis of Standard Quality Indicators: The ISTAT Experience", Proceedings of the European Conference on Quality in Survey Statistics, Cardiff, UK (2006)
- Eurostat: European Statistics Code of Practice, http://epp.eurostat.ec.europa.eu/portal/page/portal/quality/documents/code_practice.pdf (2005)
- Eurostat: Standard Quality Indicators, Doc/ESTAT/02/Quality/9/EN (2005)
- Brancato G., Pellegrini C., Signore M., Simeoni G., Standardising, Evaluating and Documenting Quality: the Implementation of Istat Information System for Survey Documentation – SIDI", Proceedings of the European Conference on Quality and Methodology in Official Statistics, Mainz, Germany (2004)
- Lyberg L., Summary Report from the Leadership Group (LEG) on Quality, Eurostat (2001)