

Innovate or perish – Italy's *Stat2015* modernisation programme

Emanuele Baldacci

Italian National Institute of Statistics - Istat, Rome, ITALY

e-mail: baldacci@istat.it

Abstract

In the last decade, official statistics has undergone a dramatic shift, both in the production model and in its output. The traditional chain, based on the vertical integration of different survey-specific tasks carried out to collect, process, analyse and disseminate statistical data, has become out-of-date. In this framework, a concrete instance is represented by *Stat2015*, the flagship innovation programme that Istat has been carrying out since 2010: it is designed to allow the transition to a business environment based on a Service-Oriented Architecture supported by plug-and-play technology, consistent with the industrialisation and standardisation of production processes.

Key Words: Industrialisation, innovation programme, plug-and-play, Service-Oriented Architecture, standardisation.

1. Introduction

The innovation process aims at enhancing efficiency and better serving increasingly diversified user requests. This reflects the need to meet the inability to exploit the economies of scale by industrialised production systems. Moving from a stovepipe model to a new structure based on a Service-Oriented Architecture (SOA) enhances productivity, supports better quality control and allows quicker innovations through a plug-and-play technology. In order to do this, the adoption of an Enterprise Architecture (EA) and of a Business Architecture (BA) is required to reformulate statistical processes and to implement generalised services, methods and standards.

However, this change involves a major redesign of the production model. In addition, the demand for statistical information is growing exponentially. Data have become the soil out of which knowledge can bloom and mining data to discover information has become an essential task in almost all human activities. As new IT and communication tools make it easier for most individuals, companies and institutions to access data, the demand for tools that allow to transform data into higher value-added knowledge to take decisions has grown rapidly. This, in turn, puts pressure on data providers to adjust their products and create tools to increase data accessibility, allow data management and mining by users, and support result dissemination.

These changes are occurring in a context where the private sector provision of statistics has also improved. Both supply-side factors, such as innovation in technologies to collect and analyse data and demand-driven forces, have increased the scope for unofficial statistical productions. Thus the traditional monopoly of statistical agencies has changed, although the need for high quality statistical data based on international standards has not been replaced. However, in this new environment public and private producers of statistical data can compete and cooperate in a market that has grown in size and complexity.

In the light of the above, this presentation aims at illustrating Istat commitments towards a change and a full-fledged modernisation route, inspired by a common vision shared with the European and international statistical communities. For this purpose, since 2010 Istat has been carrying out a flagship innovation programme called

Stat2015 with the target of allowing the transition to a business environment based on SOA, consistent with the industrialisation and standardisation of production processes.

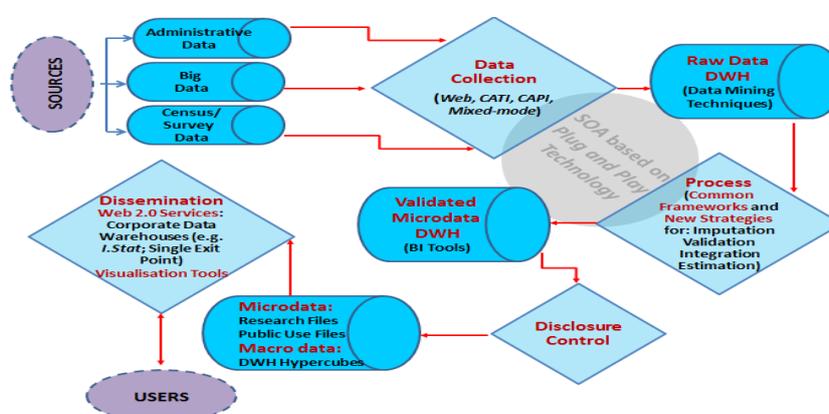
2. *Stat2015* in a nutshell

This programme is both ambitious and challenging, requiring structural changes in the working environment, in its technological and methodological foundations.

Stat2015 governance is provided by a Steering Committee and a Working Group, that are in charge of its design, coordination, monitoring and assessment. The programme is executed through different projects and the main results are documented also to internal topical committees (e.g., Innovation and Research, Methodologies, Quality, IT) and networks (Methodologists, Socio-economic Research).

The main objective of the programme is to develop soft and hard infrastructures to enable moving to an integrated metadata-driven architecture for the whole statistical production process (Figure 1).

Figure 1. Stylised *Stat2015* framework



Source: our processing, 2012

Stat2015 infrastructure is compliant with the SOA framework and aims at:

- collecting survey data, accessing administrative and other big data sources using digital technologies (e.g., from web portals, to computer-assisted interview collection, to administrative files, single entry point and web scraping tools);
- organising data in a raw data repository to be processed;
- carrying out data integration, validation and estimation tasks with generalised tools (e.g., methodological software, etc.);
- storing validated micro data in a data warehouse to be used for the following dissemination stage;
- disseminating micro and macro data through different tools reflecting user needs (e.g., from Micro data Files for Researchers through Research Data Centres and anonymous Public Use Files to the corporate macro data warehouse I.stat and related visualisation and communication tools).

Development projects included in the *Stat2015* programme are led by teams that involve human resources spanning across various Istat organisation units and focusing on specific stages of the GSBPM-based¹ production process (Figure 2). Specific resources in the budget are allocated to the programme projects and investment planning (e.g., IT and methodological support) reflects the high priority assigned to the programme. More in detail, for 2013 about one fifth of the whole budget has been allocated to *Stat2015*, corresponding to 107 projects with a relevant impact on the programme, shared in *masters* (54%) and *supporters* (46%).

¹ Generic Statistical Business Process Model (GSBPM).

Figure 2. Phases and classifications of Istat production process

Phases of Istat production process	Classification
Metadata	Referential metadata
	Structural metadata
	Strategic planning metadata
Design/Build	Determine needs for information
	Check data availability
Data Collection	Design production system and rules
	Survey data (Total/Sample)
	Data from Administrative Archives
Process/Analyse	Big Data
	Integration, editing and imputation
	Process and estimation
Dissemination	Analysis
	Validated microdata
	Corporate Data Warehouse
Procedures/Methods	Visualisation and dissemination tools
	Standardisation
	Quality
Infrastructures	Support to processes
	Common Reference Environment
	Archives
Coordination	Supporting actions

Source: our processing from GSBPM, 2012.

Each phase of Istat production process refers to specific activities carried out by Stat2015 projects, as illustrated in Figure 3.

Figure 3. Phase classifications and related activities

Metadata

CLASSIFICATIONS	MAIN ACTIVITIES
Referential metadata	Integrated information system on quality of statistical production processes
Structural metadata	Metadata modelling through standard formats
Strategic planning metadata	Information system for the National Statistical System (Sistan)

Design/Build

CLASSIFICATIONS	MAIN ACTIVITIES
Determine needs for information	Analysis of administrative sources for the continuous population Census
Check data availability	Input data in the <i>Archimede</i> System
Design production system and rules	Methodological and organisational design of the continuous population Census

Data Collection

CLASSIFICATIONS	MAIN ACTIVITIES
Survey data (Total/Sample)	Setting up a <i>Business Statistical Portal</i>
	Designing an integrated management system for all phases of Surveys on households and institutions
	Generalising and re-using total survey management systems
	Developing, innovating and strengthening data collection tools and methods
Data from Administrative Archives	Implementing a statistical virtual population register based on micro-demographic accounting
Big Data	Using information from the Internet

Process/Analysis

CLASSIFICATIONS	MAIN ACTIVITIES
Integration, editing and imputation	Logical/physical integration of administrative and statistical microdata
Process and estimation	Testing and developing generalised tools for statistical production
	Implementation of the <i>Archimede</i> System
Analysts	Developing a general government portal

Dissemination

CLASSIFICATIONS	MAIN ACTIVITIES
Validated Microdata	Enhancing Research Data Centre
	<i>I Stat</i> management system
Corporate Data Warehouse	Using <i>Single Exit Point</i> for Census Hub
	Implementing the geographic portal prototype (<i>GISTAT</i>)
	Primary Data Warehouse for population Census
	<i>Sistan</i> portal
	Updating <i>Single Exit Point</i> for Open Data formats
Visualisation and dissemination tools	Data Archive
	Developing applications for mobile devices
	Implementation of dynamic and interactive visualisation tools

Procedures/Methods

CLASSIFICATIONS	MAIN ACTIVITIES
Standardisation	Adoption of the <i>Business Architecture</i> model
	Using administrative data for economic surveys
	Implementing <i>Memobust</i> project
Quality	Methodologies for statistical quality of administrative sources
Support to process	New software technologies for classification, clustering and estimation
	Administrative processes to plan the acquisition of goods and services
	<i>Risk Management</i> model

Infrastructures

CLASSIFICATIONS	MAIN ACTIVITIES
Common Reference Environment	Systems and software supporting standards for data exchange
	IT architecture for continuous Censuses
	Developing and managing common reference environment
	Integration of management information systems
Archives	Development of territorial bases
	Updating the street and street number archive
	Designing integrated archives of social and economic microdata
Supporting actions	Istat data storage, organisation and preservation
	Developing technological infrastructures for Censuses (<i>DIDAC</i>)
	Developing internal communication
	Methodologies to support the dissemination of statistical culture (<i>e-learning</i>)

Coordination

CLASSIFICATIONS	MAIN ACTIVITIES
Coordination	<i>Stat2015</i> technical-scientific coordination
	Implementing <i>Enterprise Architecture</i>
	Implementing tools of <i>Change Management</i>
	Developing Istat human capital

Source: our processing from GSBPM, 2012.

3. Focus on the programme key components

In the short run, *Stat2015* development is planned with an emphasis on the following key components of the programme:

- design and implementation of the shared Business Architecture (BA);
- design and implementation of the Corporate Metadata System (CMS);
- tools to enhance user need assessment capacity;
- data collection and management web portals;
- generalised methods and tools for data processing and analysis;
- new dissemination and communication tools for micro and macro data to enhance user interactions.

With regard to BA, this is central for defining, managing and monitoring the necessary changes for *Stat2015* development: this approach can guide the operational changes in the production chain, highlighting the required infrastructure that should be put in place under this framework².

CMS planned design and implementation are expected to consolidate existing systems (e.g., the quality metadata information system). The medium-term objective of the process is to adopt an Istat metadata architecture that guides users through all stages of the production chain.

Assessing user needs represents the very first stage for tailoring products and services, developing new tools and partnerships that allow Istat to assess what users really need. Concerning this, the following two topics feature prominently in the agenda of *Stat2015*:

- some web-based tools to gather customers feedbacks are being developed to allow maximum user interactions with Istat. The key portal to collect these feedbacks and suggestions is the corporate web site *www.istat.it*, but also social platforms and networks are increasingly exploited to gather feedbacks in a dynamic way;
- the web portal of the Italian National Statistical System (NSS)³ to provide adequate information and specific tools, helping to disseminate the results of the innovation programme, its infrastructure and services among this network of public and private producers of statistics.

Concerning data collection tools, *Stat2015* aims at achieving significant progress moving to fully digital data acquisition systems. In this case, the strategy is differentiated by type of data to be collected and within survey data it acknowledges the differences between business and institution statistics on the one hand and household surveys on the other hand. Although tools developed are different, the strategy for data collection is built around common concepts. Data collection tools are being developed with the purpose of minimising respondent burden, reusing information that has been already collected on respondents through a variety of sources and building platforms that allow seamless exchange and data integration. The strategy to develop platforms for generalised data acquisition systems is built around the following main pillars:

- Business statistical Portal, for collecting data effectively via web through a one-stop-shop, using all available information consistent with the e-government strategy; back-office operations, moving away from a vertically integrated stovepipe model; structured interactions with respondents, including returning key tailored information to them through the web portal itself.

² For a deep analysis of this topic see Barcaroli and others (2013).

³ Istat is responsible for the coordination and the technical support of this portal.

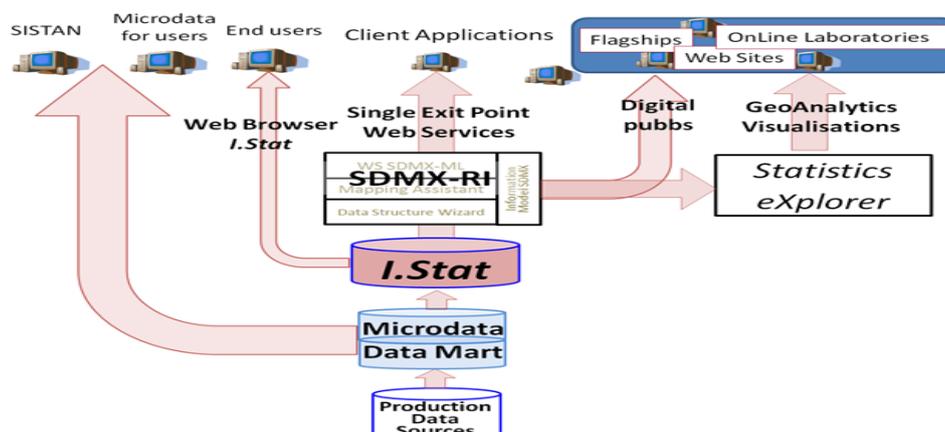
- Computer-assisted techniques (CATI/CAPI/CAWI), for all household surveys, building the related generalised tools.
- Infrastructure for a Single Entry Point, for all administrative data used in the statistical production (consolidating the technologies of the different access systems in place), providing on line documentation to enhance data integration.
- Big Data mining engines and web-scraping infrastructures (also in partnership with research centres), for collecting auxiliary or substitutive information for statistical production, using new sources available on the web and other sources.

Concerning the processing and analysis components of *Stat2015*, they are focused on the creation of generalised tools to carry out data integration, validation and estimation functions. There are several activities that are conducive to establishing a set of standard methods and tools that would constitute the library (a real *Istat Store*) to be used by analysts for their tasks, addressed to the following main issues:

- development of SOA-compliant standard software to perform data processing functions and to achieve a gradual replacing of existing instruments, so as to move away from the stovepipe model;
- improvement of new methods to support the integration of survey data and administrative sources in the statistical production process, enhancing estimate quality;
- construction of a generalised conceptual methodology to use administrative data for sampling, validation and estimation;
- production of enhanced methodological documentation and manuals, on line and user-friendly software repository tools, standard certification guidelines and procedures.

Regarding data dissemination components of *Stat2015*, they represent some of the most advanced elements of the programme. The corporate dissemination infrastructure is already centred around the macro data warehouse called *I.stat* and its satellite thematic systems (e.g., social cohesion). New visualisation tools have also been released recently using Statistics eXplorer technology. Progress has been made on developing and implementing a Single Exit Point using SDMX for machine-to-machine data communication. Micro data are available through a network of research data centres and anonymous research files (Figure 4).

Figure 4. Istat data dissemination framework



Source: our processing, 2012.

The strategy followed is to build on existing infrastructures for enhancing their capacity and strengthening the links between production and dissemination data warehousing tools, following these key priorities:

- consolidating the overall dissemination infrastructure design to enhance its efficiency and flexibility;
- expanding aggregate data dissemination tools linking the corporate data warehouse to the geo-analytics visualisation tools (GIS-based and eXplorer-based platforms) with a view to strengthening authoring tools and interactive access to all data by users;
- strengthening machine-to-machine data communication tools and hubs to support information exchange and enhance access to data produced by NSS network;
- centring micro data dissemination on the availability of Public Use Files and Files for Research, enhancing the access to Istat Research Data Centres and developing the Italian Data Archive in partnership with other public and private institutions;
- continuing to link communication tools to dissemination platforms, ensuring consistency and constant updating of data on the website, on mobile apps and e-books.

4. Concluding remarks

In the last decade official statisticians have revisited their business model to remain relevant in a fast-changing world where advanced technologies and communication impose challenging structural reforms also in the statistical production chain. Statistical offices around the world are at work to implement these changes, guided by a common vision. Achieving modernisation is a task that combines different skills and the general lesson learned is consistent with the main findings in the economic reform literature. Decision-makers need both to translate their vision into a clear political mandate for changes and to generate political consensus for reforms. This requires:

- clear political commitment from top management;
- buy-in on vision from middle-management;
- consensus on modernisation targets from staff.

Good management, a change-driven culture, risk taking and open minds ensure that results are delivered effectively. All these factors are related to people: human capital is the core ingredient in innovation and creating a strong and sustainable support for the ambitious transformation agenda needs to invest in people. As data are the soil that can grow beautiful flowers, people (producers and users) are the hands that mould the soil into something new and better. A changing process relies on people; consequently, a good people training is one of the most relevant factor to put in place modernisation programmes and should build together human resources and strategy managers with a view to generating an environment prone to innovative out-of-the box thinking.

References

- Baldacci, E., McHugh, J. and Petrova, I. (2011). Measuring Fiscal Vulnerability and Fiscal Stress: A Proposed Set of Indicators. IMF Working Paper 11/94 (Washington: International Monetary Fund).
- Barcaroli, G., Falorsi, P.D., Fasano, A. and Mignolli, N. (2013). Business Architecture Principles to Foster Industrialisation and Standardisation at the Italian National Institute of Statistics. 59th World Statistics Congress, Hong Kong, 25-30 August.
- Bredero, R., Dekker, W., Huigen, R. and Renssen, R. (2009). Statistics Netherlands Architecture; Context of change. Discussion paper (09017), CBS Statistics Netherlands, The Hague Heerlen.
- Giovannini, E. (2010). Statistics 2.0: The Next Level. National Conference of Statistics, Istat, Rome.
- McCandless, D. (2010). Data is the new soil. TED talk on visualising data, Columbia Journalism Review.
- Statistics Netherland (2012). Strategy to Implement the Vision of the High-level Group for Strategic Developments in Business Architecture in Statistics. Conference of European Statisticians. Sixtieth plenary session, Paris, 6-8 June.