

MANAGING AN INTEGRATED RESPONDENT COMMUNICATION: STATISTICS PORTUGAL EXPERIENCE

Paulo Saraiva¹, Almiro Moreira²,

¹ *Statistics Portugal, Lisboa, Portugal, paulo.saraiva@ine.pt*

² *Statistics Portugal, Lisbon, Portugal, almiro.moreira@ine.pt*

Abstract

Statistics Portugal started ten years ago a process of modernisation of its production system. In this process, focusing firstly on business surveys, three key elements worth to be highlighted:

i. The implementation of the Simplified Business Information (an administrative internet based source) that conveys each year almost census accounting data on the corporation sector, covering the information requested by four public institutions;

ii. The full implementation of an Integrated Survey Management System, which consists on components that support the basic statistical production sub-processes collect, process, analyse and disseminate;

iii. The promotion of a better relationship between data providers and Statistics Portugal notably by improving the dialogue with the data providers and by designing a service of customized feedback information to them.

This paper provides an overview on the content of these three elements.

Keywords: Official statistics production; process integration; respondent management; data collection; survey management.

1. Introduction

Likewise the majority of official statistics producers, Statistics Portugal faces increasing challenges, as the decreasing funding and of other new phenomena that changed the offer of new kinds of information, like alternative sources of data that do not fit into the concept of official statistics, but which are valued by the society. In fact, while society wants diverse and faster results, other sources are becoming available, based on the large amounts of data that are automatically collected, as Big Data.

Similarly to other National Statistical Offices (NSO), Statistics Portugal is still very survey-oriented. As the perception of the utility of statistics for the companies and individuals is constantly jeopardized, as well of pressures to cut production costs, data providers resist collaborating with Statistics Portugal, and respondents are also demanding for advanced and simpler modes to provide the required data. This new context requires that NSO must be prepared to change its processes, infrastructure, and skills.

Statistics Portugal continually seeks to introduce innovative solutions to scientific and technical level to the deepening of quality and minimizing the costs of data collection.

Integrating information systems is one major concern among companies and institutions, being widely accepted as a critical success factor of any organization.

After an extensive internal discussion prior to 2005, Statistics Portugal has decided to re-engineer its production architecture from a traditional stovepipe approach to an integrated one. Business statistics were the first area to benefit from this new system, followed by social statistics.

This paper offers an historical view of this journey and lists some results achieved. For the purpose of this Conference, we will focus on business surveys, and two key elements will to be highlighted: (1) the full implementation of an Integrated Survey Management System, and (2) the promotion of a better relationship with data providers by designing a service of customised feedback information to the companies.

2. Statistics Portugal

Statistics Portugal is the National central authority for the production of statistics. Its main task is to develop and supervise the national statistical system.

Data collection (using traditional surveys) is a core function of Statistics Portugal, consuming around 40% of its annual budget and 30% of its human resources. A Data Collection department assures mainly the operation of statistical production phases of collection, processing and analysis of collected microdata, covering all business and social surveys. Data collection staff is spread all over the country (mainland and islands), especially in Lisbon, Oporto, Coimbra, Évora, and Faro, but under centralized system. The Autonomous regions of Madeira and Azores have their own authorities for the production of regional specific

statistics, while being the data collection centres for those areas for Statistics Portugal, and for national purposes, under common technical requirements and infrastructure.

Statistics Portugal has ten years of experience in the modernisation of its statistical production and services, improving the efficiency and flexibility. The strategy that was implemented implied the replacement of the existent stovepipe model with an integrated one.

Business surveys involve around 85.000 different companies, 99% of them are considered as small and medium enterprises. In 2015, 635.000 self-completed questionnaires were collected, 95.5% of them electronically via WebInq, which is an online service dedicated to information providers.

The current production organization of Statistics Portugal (simplified) is shown in Figure 1.



Figure 1: Production organization of Statistics Portugal

3. Integrated Survey Management System

Prior to the year 2004 Statistics Portugal produced statistics through a non-integrated organizational architecture, based on numerous parallel processes, domain by domain, place by place. Data collection activities were conducted by individual survey departments and regional directorates. A survey organizational unit would design their survey, but the development and operations were usually not harmonized. Centralized activity was mainly restricted to some software development and to a weak coordination of the household interviewer workforce.

After a reflection and a reorganization process in 2004, a project to re-engineer the production architecture was undertaken based on an integrated and process driven approach aiming at improving its efficiency and flexibility.

This effort resulted in an Integrated Survey Management System (SIGINQ), which covered firstly the business surveys, and later on the social surveys.

From 2008 until now, the number of features supported by SIGINQ grew up continuously. Four years after its first version, SIGINQ covers all business surveys, which are fully supported by this system.

The SIGINQ aims at offering an integrated infrastructure to better support the statistical production and development in an efficient way, covering all the statistical operations (business and social). It unifies the main components into a comprehensive and interdependent system based on the architecture illustrated in Figure 2.

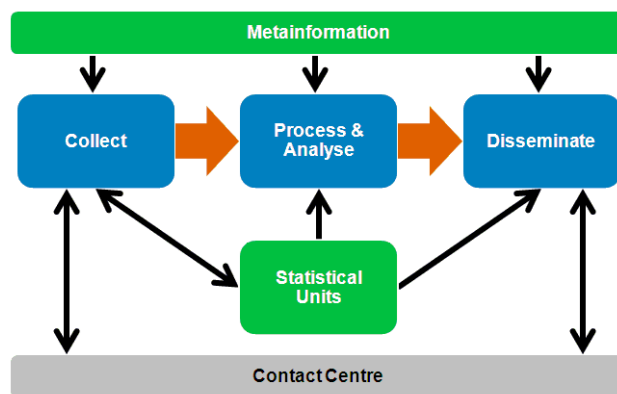


Figure 2: Integrated Survey Management System architecture (Level 1)

The system follows the basic production sub-processes collect, process, analyse and disseminate. Statistical units registers and metadata support the flow of the processes. A contact centre system offers the infrastructure to telephone interviews (for social surveys), and the support to data providers. The next points describe these subsystems, especially the processes and systems that support the production of business statistics.

The main components of SIGINQ are described below.

Statistical Unit Registers System

The registers aggregate all of statistical units, which provide the basis for the selection of the populations and samples for all surveys.

These registers include variables of identification, localization, and characterization of their statistical units. The register is updated to assure quality to the samples and surveys themselves are one of the most important sources for this purpose.

Population and Samples System (SIGUA)

This component aims at creating and maintaining a repository to store all of the reference population and the samples selected to the surveys. This is a key element to the Collect System.

Online processes to update the repository, managed centrally, are available. One internal collection agent can submit proposals to update some information about the statistical units. These proposals are analysed centrally by the Methods team (Methodology Unit ?), and information is updated, if considered reliable. Finally, the proponent receives the feedback of the proposal, i.e., if it was accepted or rejected, and why.

Collection System

The Collection System aims at feeding the Statistical Production Chain with microdata, and it is composed by three components: (1) Process Management; (2) Questionnaires and Data Capture, and (3) Respondent Management. The diagram of Figure 4 shows these three components, and its relationships.

- **Process Management**

This component is responsible for the management and control of all data collection processes, including information about respondents and paradata. These processes are fully supported by the Metadata System.

- **Questionnaires and Data Capture**

This system offers the data collection supports and IT solutions, each of them dedicated to a particular mode of collection: electronic, paper, telephone, face-to-face interview, etc. For business surveys, questionnaires and data capture have two components: WebInq and WebReg.

WebInq: Surveys in the Web

WebInq is the area of Statistics Portugal Portal where data providers can respond to surveys' questionnaires via Internet webforms. Respondents can find information about the surveys and

further information about other surveys where the statistical unit is included, the status of response, and also extensive information about the respondent and its response behaviour.

WebReg is a clone of WebInq, aiming at supporting the internal data entry of the remaining paper questionnaires, been also a data editing tool.

- **Respondent Management**

This component aims to maximize the relationship with the data provider and the respondent. This is achieved through a repository of all respondents, including information about the identification, localization, contacts, relationships and their collection behaviour, (history of the collection activity, quality of the data provided, response timing, etc.). This tool is very important when the processes are repeated regularly.

Process and Analyse System

This system prepares the cleaning of data records, preparing for data analysis, the stage where statistics are produced, examined in detail and made ready for dissemination. As the "Process" and "Analyse" phases are iterative and parallel, the system assures the principle of “interdependence but no interference” between both phases, which is represented in the Figure

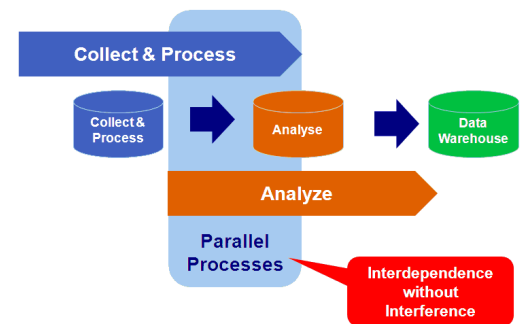


Figure 3: Parallel Processing

3. Process and Analysis System supports the following GSBPM phases and sub-processes: (5.) Process: (5.4) Impute; (5.5) Derive new variables and statistical units; (5.6) Calculate weights; (5.7) Calculate aggregates; (5.8) Finalize data files; (6.) Analyse: (6.1) Prepare data outputs; (6.2) Validate outputs; (6.3) Scrutinize and explain; (6.4) Apply disclosure control; (6.5) Finalize outputs.

Metainformation System

This system is composed by several components, such as: Terminology and concepts; Statistical Classifications; Repository of questionnaires; Methods Documents; Variables and Questions. Metainformation is closely interrelated with Collection, Process and Analyse, and Dissemination systems.

Contact Centre System (SICC)

As mentioned before, this system offers the infrastructure to telephone interviews (households' surveys), telephone reminders and the support to data providers. It also facilitates the access to context information about data providers, respondents and surveys. From March 2012, SICC supports another component of Questionnaires and Capture System: Telephone Data Entry (TDE), which is a solution by which respondents can return their data using the keypad on their telephone.

4. Customised feedback to data providers

Having an integrated production system is a key element to improve the relationship with respondents. As long as the data collection processes occurs, paradata is automatically collected, offering an extensive knowledge of the data providers and respondents.

Nowadays, it is easy to answer some questions like: What are the surveys that a company or a respondent have to reply to? What is their response behaviour? Those questions were nearly impossible to answer with the former approach.

Harmonized skills are used to provide an improved capability for respondent communication. A standard reminder format (?) and follow-up strategies are in place, having an active capability to modify the implemented approaches accordingly to the development of a certain data collection campaign.

Written questions, complaints and suggestions from respondents are handled centrally to ensure a consistent reaction based on standard templates.

Statistics Portugal efforts to improve the consistency of respondent communication are illustrated by one example. In 2013 Statistics Portugal carried out this ad hoc survey to data providers, which allowed obtaining a picture of how the obligation to reporting statistics is viewed by the respondents. With the results was possible to have further information about the Perceived Response Burden and the quality of the data reported, and obtain critical suggestions for the improvement of procedures in data collection. After the survey, one of the major results was the following: "Companies consider that the statistical information disseminated by Statistics Portugal have significant utility to the society, but they are more reluctant to admit their own interest on it." These results are presented in the two figures indicated as followed.

Company size	Very useful	Useful	Not so useful	Useless	Do not know
Big	25,5%	57,2%	9,2%	2,8%	5,3%
Medium	20,5%	58,3%	12,4%	1,9%	6,8%
Small	17,2%	55,3%	16,7%	3,7%	7,1%
Total	18,5%	56,1%	15,2%	3,2%	6,9%

Figure 4: Perception of the usefulness of statistical information for the society according to the company size

Company size	Very useful	Useful	Not so useful	Useless	Do not know
Big	7,1%	44,8%	35,9%	4,4%	7,8%
Medium	4,6%	42,4%	36,3%	9,2%	7,5%
Small	3,6%	40,4%	35,3%	11,7%	9,1%
Total	4,1%	41,1%	35,5%	10,6%	8,6%

Figure 5: Perception of the usefulness of statistical information for the company according to the company size

In order to deal with these contradictory perceptions and also offering a sign of recognition for the effort of the information providers, Statistics Portugal recently started to make available a new feature of the WebInq: the Customised feedback to data providers.

This feedback corresponds to the periodic provision of three types of reports:

- Customised Reports, which bring together, in a synthetic and targeted way by themes, the information collected, including indicators of the relative position of the company given the results of investigations in which it participates and other information of a specific nature, provided they do not compromise the principle of statistical confidentiality;
- National macroeconomic framework, quarterly updated;
- A link to the electronic brochure of economic activity more updated.?

For confidentiality reasons, this feature is available to specific WebInq users with the proper authorization from the companies to access this kind of sensitive information.

After authentication by username and password, the respondent can access customized reports for all companies who have authorized his or her access to WebInq, as shown in the Figure 6.

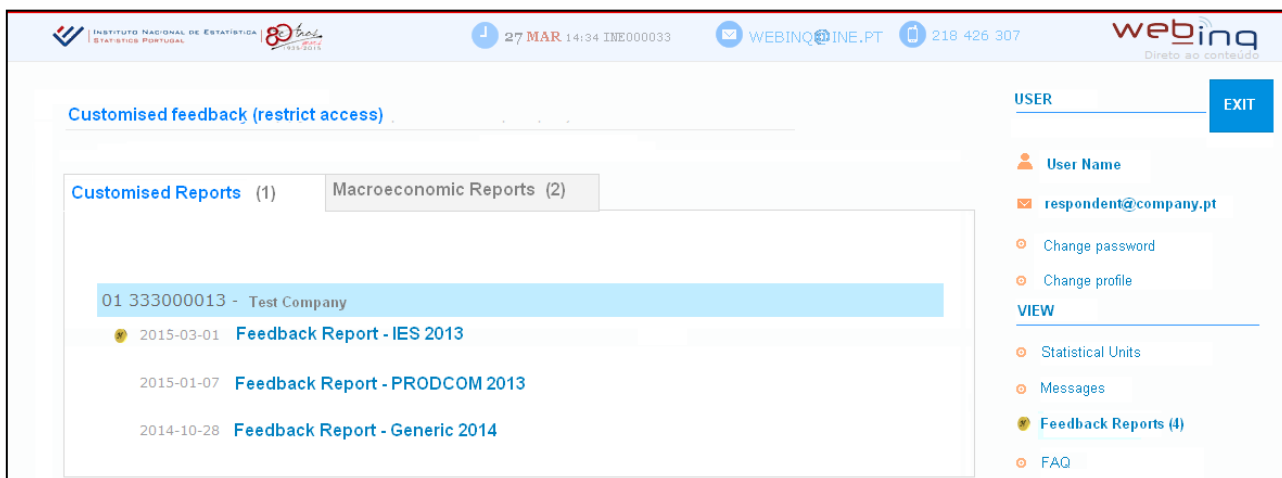


Figure 6: WebInq menu to access the customised feedback reports.

The reports have been designed in a very concise manner, using graphics and images, as in the example shown in Figure 7.

5. Future developments

Statistics Portugal will continue to extend the process of feedback of information to respondents in the field of business surveys, promoting good value and enhance the collaboration with information providers.

Firstly, Statistics Portugal will create other types of custom reports, according to the characteristics of each company as well as their involvement in the various surveys.

The next step will be to implement a new feature that allows respondents to choose the most appropriate indicators for its management and produce other customized reports available in more formats and offering a dynamic way for automatically updated versions of this type of specialized information.

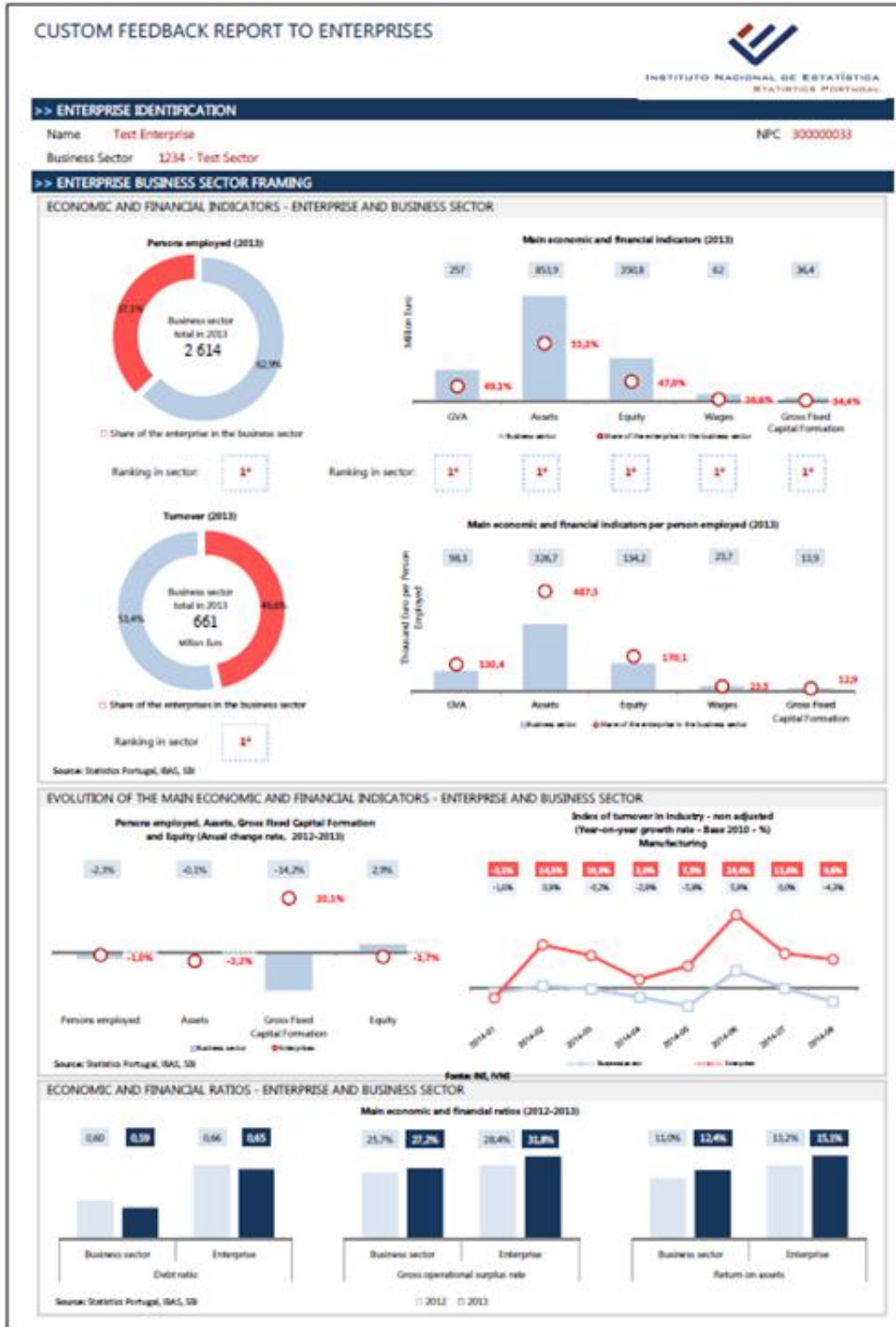


Figure 7: Example of a personalised report to a data provider (fictitious).