

Integrating Data Collection: wins and challenges

Paulo Saraiva (Statistics Portugal)

paulo.saraiva@ine.pt

Abstract

Statistics Portugal adopted in 2005 a centralized Data Collection Department and it has an integrated survey management system since 2007, covering both businesses and household surveys. The Integrated Survey Management System (SIGINQ), based on statistical units and metainformation components, is broadly identified as an innovative practice that could benefit the European Statistical System.

Having an organization and an infrastructure in which data collection is integrated is clearly not a trivial matter. There are more management aspects than technical challenges that must be known in advance and dealt with. It results that the main critical success factors are the top management sponsorship, the involvement of a multidisciplinary team, an extensive requirements design, and the internal change management.

This experience provided wide and significant insights to Statistics Portugal. We faced various barriers and successes during the processes of deciding, designing, developing and implementing a huge transformation.

On the other hand, there are very challenges that are now placed and deserve to be discussed. Having a pipeline data collection organization and infrastructure does not assure that a Statistical Office has a complete integrated and flexible workflow of the production. For example, this approach may unexpectedly perpetuate the dependence of statistical production in surveys and improve resistance to new alternative data sources.

Consequently, the author believes that it is essential that Statistics Offices should undertake a complete transformation of the production system, in particular within its data collection component, both from the perspective of design and in its operational aspects. It is understood that this transformation of the collection should be based on the integration of all potential data sources for statistical purposes: those originated from surveys, from administrative sources, from Simplified Business Registers and from other non-traditional sources (such as Big Data), also providing an easy and frequent feedback to data providers.

This challenge requires other processes and infrastructures, because simply replacing surveys by alternative sources rarely work without changing the systems and existing capacities.

This presentation provides an overview of wins and challenges in integrating data collection for official statistical production and is intended to stimulate a fruitful discussion of this matter.

Keywords: Statistical data collection, integrated systems, statistical production.