

Prelude to the Special Issue on Systems and Architectures for High-Quality Statistics Production

Many survey professionals have become aware of an increasing dependence among statistical methods, statistical practice as implemented in tools and procedures, and information technology (IT). This dependence is reflected in a shift towards standardized processes and tools, integrated systems for production of statistics, statistical modernization programmes, transitions from stove-pipe based to systems-based statistical production, and similar activities. These developments have arisen within statistical organizations and through cross-organizational collaboration.

In developing this special issue, we identified two trends – one in methodology and one in IT systems – that were influential in these developments. The methodological trend leads towards a balanced evaluation and treatment of the components of aggregate error associated with a statistical product. The result, sometimes called Total Survey Design or, more generally, Adaptive Total Design (ATD), is also well integrated with the quality perspectives on statistics production. The IT system trend centers on developments in “enterprise architecture,” which very broadly is a formal, structured description of an enterprise¹. These developments are also influenced by increasing financial pressure experienced by large-scale statistical organizations.

Twenty years ago, introducing a special issue of JOS (1993, Vol.9:1) on “Current Research and Policy Issues at Statistical Agencies and Organizations,” the then editor-in-chief Lars Lyberg noted that statistical agencies are undergoing “a perennial flow of the flux and change”, and followed with a list of areas affected by new developments. The list included, among other items, user needs and the external survey environment, development of new programs, interagency harmonization, new technologies, and confidentiality protection. All of these were said to be influenced by increasing pressure for productivity and quality gains. Most of these issues remain with us today in one form or another.

Within the context defined by these evergreen topics, there has been substantial progress in both of the abovementioned trends. Embracing the enterprise architecture approach, recent years have seen an active international collaboration regarding the foundations of production of statistics. Common conceptual frameworks, like the Generic Statistical Business Process Model (GSBPM) and Generic Statistical Information Model (GSIM), are the results of such international initiatives. In addition, there is a wide range of

¹ The terms that we are using in this introduction are well explicated in the contributions to this special issue, which is the reason we, for brevity, refrain from providing references here.

international cooperation efforts, including a high-level group, a statistical network, a thriving wiki organised under the UNECE auspices², and a set of projects that Eurostat is coordinating among European Union (EU) member states towards implementing an EU Vision, among other forms of cooperation.

On the methodology side, progress has been slower and international collaborations have yet to form to the extent that they have in the former area. Nevertheless, substantial progress has been made within many statistical organisations that view product quality as an ongoing process of error assessment and reduction. Applying the ATD paradigm to statistical production systems, improved product quality is achieved by assessing the error risks associated with multiple error sources and applying error reduction strategies that mitigate the major risks. The new quality improvement strategies rely less on indirect indicators of data quality such as response rates and agreement rates, and more on direct evidence of bias and error variance obtained through nonsampling error analyses.

Also, from a broader perspective, formalised methodological approaches (based on the ATD paradigm and generally relying on some architectural features) are important in ensuring that statistical agencies have robust structures through which to address the abovementioned “flux and change.” These robust structures can support both development of additional statistical products that meet stakeholders’ current needs, and improvement in the balance among quality, cost and risk in statistical production processes.

This special issue of JOS is intended to support current and future developments regarding systems and architectures for high-quality statistics production in two ways. First, the special issue presents the “state of the art” in this area. How far have organizations at the forefront of the initiated changes come, what have they achieved thus far, what “lessons learned” are they drawing from their experiences, what are they currently working on, and what issues have they identified for discussion by the wider community? We hope that answers to these questions will help the reader compare and contrast the approaches to extract useful insights for their own organisation, or to identify points of departure for further work.

Second, we hope to enhance and accelerate further integration of methodology and IT for statistical production systems. For instance, from an ATD perspective, there is an increasing need for paradata and other forms of process data, administrative data, and metadata. Allocating costs to detailed system components or to the steps in production of statistics – which is needed for improvement in the balance of cost vs. quality – appears at the moment to be an almost impossible task in many organisations. It is our impression that only advanced IT solutions can help provide these data in a manageable way. Therefore, enterprise architecture is needed.

However, a formal system that describes the production of statistics is somewhat distinct from a system for production of consumer goods or services. Thus, we assume there will be some important differences in enterprise architecture for statistical production, relative to the architecture developed for other commercial or governmental work.

² www1.unece.org/stat/platform/display/msis/

Development of the GSBPM and GSIM are important steps in the exploration of these differences, and related insights have been presented at symposia organized independently by Statistics Sweden (“Modernisation of Statistics Production” in 2009) and Statistics Canada (the International Methodology Symposium 2011: “Strategies for Standardization of Methods and Tools – How to Get There?”). We hope that the content of this special issue will make additional contribution in this area.

Pursuing the two abovementioned goals, in May 2011 we initiated work on a JOS special issue on systems and architectures for high-quality statistics production. The call for main papers resulted in nine abstract submissions, of which five resulted in articles included in this special issue. Our foremost aim with the articles is to provide a solid review and synthesis of the current state of knowledge and practice. We were fortunate to be able to work with representatives from two of the intended contexts, official statistics and private sector, while we regret we were not able to include any academic contributions.

To broaden the perspectives of the articles, and to expand consideration of methodological issues, we invited prominent survey methodologists and other experts to provide formal comments on the articles. The aim with the discussions is to facilitate creation of a broad agenda that identifies several areas in methodological research that will be crucial to subsequent development of systems and standardization that are consistent with rigorous, high-quality methodology. These constitute the second part of this special issue.

Departing from the usual format for a journal with articles followed by invited comments, we do not provide rejoinders by the articles’ authors in this issue. The lack of a “coda” (in the musical sense) in this discussion is intentional in two ways. First, we want to open the floor to our readers (including the authors and other interested stakeholders) to submit further papers on the topic, or share their thoughts on the subject through the Letters to the Editor section. Second, we do not believe it is currently possible to give any authoritative final judgment on the matter – the lack of a coda hopefully emphasises this view.

Instead of a coda, the sixth article, by the three guest editors for the special issue – John Eltinge, Paul Biemer and Anders Holmberg – seeks to place work with statistical production systems and methodology in a broader context defined by multiple dimensions of quality, cost, and risk.

This special issue represents the results of the activities outlined above. We hope that it contributes to the work of building better systems for production of high-quality statistics, and better integration with the methodological accomplishments. Together, these two components ought to lead to statistical products that are of high quality and broadly used. We sincerely thank all the contributors to the special issue for their dedicated and persistent work towards achieving this goal.

Having received an exceptional contribution by Jack Nealon and Elvera Gleaton, we were saddened to learn about the illness and subsequent passing away of Jack Nealon in December 2012. Jack spent his entire 37-year career with the National Agricultural Statistics Service (NASS), serving as the Director of the Information Technology Division and the Chief Information Officer for NASS, and at the time of his retirement was the Senior Project Manager, Office of the Administrator. Jack was known as a great

communicator who was passionate and had a deep desire to excel. His dedication to completing the NASS contribution for this special issue despite his failing health was an inspiration to us all and will long be remembered.

Boris Lorenc
Ingegerd Jansson

Paul P. Biemer
John L. Eltinge
Anders Holmberg