VERSITA

Journal of Official Statistics, Vol. 29, No. 1, 2013, pp. 193–201, DOI: 10.2478/jos-2013-0014

# Discussion

*Claude Poirier*[1], *Claude Julien*[1], *and Robert McLellan*[1]

## 1. Introduction

National Statistical Offices (NSOs) must continuously evolve to produce the broad range of information required by law, regulations and numerous users of official statistics. Typically, NSOs have redesigned, enhanced, added or dropped production lines and survey processes that are executed fairly independently from one another. The three articles presented by NSOs in this special issue clearly show that they are facing similar budget constraints, which translate into a set of common drivers for the transformation of their statistical programs and organization. Amongst them, we note the desire to ensure cost efficiency in the processes; to reduce response burden, especially for businesses; to improve coherence; to maintain high quality standards; and to improve the responsiveness of statistical programs.

This discussion reviews the submissions from three NSOs who have moved forward on their path of modernization. Statistics Netherlands writes about its adoption of an "enterprise architecture" approach through process standardization and focus on design and data "steady-states". Statistics New Zealand writes about its Statistics 2020 initiative and the key actions they have taken to transform their statistical production systems, with important changes to their corporate design and database design. Finally, the National Agricultural Statistics Service (NASS) in the United States shares its experiences of moving from a highly distributed approach to centralized infrastructure and data management.

It is clear from reading the three articles that, in order to meet these objectives, redesigning and enhancing current independent production lines or processes is not sufficient. NSOs require much larger scale transformation in how they produce and deliver information and how they organize themselves accordingly. This discussion outlines the main solutions proposed by the articles' authors in Section 2 and how they were implemented in their respective organizations in Section 3.

## 2. Solutions

National statistical organizations are all facing similar issues with respect to efficiency and risk management. At the same time, they have access to technological and methodological advancements. Several good solutions are proposed to react to budget constraints, to

[1] Statistics Canada, 150 Tunney's Pasture Driveway, Ottawa, ON, K1A 0T6, Canada. Emails: Claude.Poirier@statcan.gc.ca, Claude.Julien@statcan.gc.ca, and Robert.McLellan@statcan.gc.ca

improve productivity, to reach unprecedented coherence, to reduce risks of errors, if not simply to adopt common business processes. Most of these solutions address several goals, as described below.

### 2.1. Centralized Databases

Centralized databases provide a number of benefits to a statistical organization. They allow access to a vast amount of survey and administrative data that are at the core of the cost-efficient transformations taking place in the NSOs. Such sharing of data avoids the collection of similar information, if not simply to serve as proxy values for designing surveys or modeling data.

Centralized databases are the result of good information management as long as they are interpretable through appropriate documentation and detailed metadata. In their articles, the three NSOs mention that the various versions of their data – or audit trails – are stored in their centralized data warehouses, in order to allow rerunning or evaluating any process. Other benefits include a less confusing environment for processing purposes, a more efficient use of all information during the survey cycles, a potential increase of NSOs' responsiveness, and an input to corporate memory.

However, centralizing data requires more complex management of data confidentiality within the organization. Indeed, many survey areas collect information that should not be shared with other areas. Therefore, centralization and the related access control bring a risk of mismanaging the accessibility that may result in undesired disclosure.

Users of administrative data often face issues with their timeliness. Data modeling should be considered as a solution to this challenge. Statistics New Zealand plans on using modeling as an opportunity to modify tax data that do not correspond exactly to the variables of interest. This approach can be expanded to introduce trend adjustments to any administrative data when they do not fit the appropriate reference periods, or to enable the exclusion of small units during the regular data collection and allow for adjustments through models.

### 2.2. Standardization

Standardization is a key principle for most NSOs facing financial pressures. Statistics New Zealand targets standardized processes, methods, tools and systems. This will help them to reduce IT maintenance costs and risks. NASS addresses standardization through centralization and consolidation. Statistics Netherlands invests in standardization to enhance the flexibility of statistical processes, to increase their uses and to reduce costs.

Standardization means that common approaches are used to perform business functions. Standardization usually requires the consolidation of some systems to maintain a reduced set of functions. In this context, generalization is carried out to make sure that the reduced set is developed with an acceptable level of flexibility that still satisfies a wide range of applications. This approach targets reusable building blocks for which behavior is controlled with parameters. To be usable in multiple instances, large monolithic solutions need to be broken into smaller components. This concept of breaking large solutions into smaller components is central to the 'plug & play' architecture being pursued among NSOs (HLG-BAS 2012).

Statistics Canada is currently modernizing its Corporate Business Architecture (CBA). Key elements of this involve creating standardized collection processes, tools, and data that feed downstream social, economic, and Census processing. Standard processing platforms for economic and social statistics are also being developed. Since processing tools must be able to interact, it is necessary to standardize interfaces and data. For that reason, Statistics Canada has had in place a central metadata management database, which provides a standardized metadata space covering important aspects of statistical production.

Statistics Canada has adopted the Generic Statistical Business Process Model and has contributed to the Generic Statistical Information Model (GSIM). It is our belief that GSIM will provide a basis for our information standardization activities as part of our CBA modernization.

We realize that standardization should not interfere with extensibility. For instance, while classification is part of standardization, we have found that users frequently extend classifications in use for valid business reasons. In our experience, expecting there to be no refinement of standards is likely unworkable. Establishing clear working mechanisms to allow for controlled extensibility with associated mappings is a reasonable compromise to support business innovation.

### *2.3. Integration*

Integration is the connection of various centralized databases, functions and standard tools to produce an end-to-end production line. NASS uses integration as one of four criteria to determine if a transformation met their objectives. Statistics New Zealand combines a set of processes into a configuration that is applied to a set of data. Statistics Netherlands goes further with an integral approach that optimizes a set of production lines that may share the same source of data.

Statistics Canada's Integrated Business Statistics Program (Ravindra 2012) is somewhere between the latter models. When completed, it will be a large-scale platform that will support the production and analysis of more than 100 business surveys and make extensive use of tax data. It will use generalized systems for sampling, edit and imputation, estimation and time series, and other standard systems for collection, analysis and dissemination. The single platform will require much less IT maintenance than would individual production lines. However, considerable IT resources are required to program the numerous data manipulations to integrate the standard tools into the single platform.

### *2.4. Flexibility*

The NSOs demonstrate a common interest for a flexible environment. As mentioned by MIT-ESD (2001): "Flexibility may indicate the ease of 'programming' the system to achieve a variety of functions. It can also indicate the ease of changing the system's requirements with a relatively small increase in complexity and rework." In other words, flexibility introduces the concepts of (a) adaptability and (b) extensibility.

(a) Concerning adaptability, the idea behind most transformation projects is to standardize processes and achieve systems that can be adapted to most applications. It must be noted that such an effort rarely provides the level of flexibility that the

initial environment offered, since flexibility competes with standardization. This is the price to pay in the quest for more efficient processes. Senior management should make sure that business cases address this "global optimum vs. local optimum" trade-off early in the transformation project.

(b) With respect to extensibility, it is important to set up processes that can be modified in order to satisfy evolving requirements. This is a prerequisite for a responsive architecture. This aspect of flexibility is addressed only implicitly by the NSOs, but Eltinge, Biemer and Holmberg address it explicitly in their proposed framework, through the dynamic features of quality, cost, risk and relevance. They purposely suggest that performance criteria should go beyond traditional metrics like bias and variance.

### *2.5. IT Strategies*

The approaches outlined by the authors vary in the specificity of the underlying information technology and solutions. Central to the approach taken by NASS is the creation of a centralized infrastructure function in line with modern IT practices. The modernization of networks, centralization of servers, use of both desktop and server virtualization, and deployment of standardized desktops are all part of mainstream IT modernization activities in industry today. Both Statistics New Zealand and NASS report on important changes in their approach to providing databases, with Statistics New Zealand highlighting a shift to Online Analytical Processing (OLAP) approaches away from Online Transactional Processing (OLTP). The adoption of more "storage-hungry" solutions is offset by the use of modern storage technology in flexible storage solutions. Similarly, processing power to support a variety of approaches enables them to create a richer data and metadata environment with powerful flexibility. The importance of a comprehensive enterprise architecture which incorporates the IT aspect reflects the need to bring IT and non-IT communities closer together in a collaborative setting.

## 3. Good Practices

The transformation of a statistical production system can only be successfully implemented using sound practices. This section outlines good practices that are reused by NSOs. While these objectives address current risks, they may introduce new risks that should not be ignored. Special attention is paid to these new risks below.

### *3.1. Clear Objectives*

Transformation means changing the way we do things with the objective of improving the process and reducing vulnerabilities. Transformation projects always start with a business case to demonstrate benefits of proposed solutions and list clear objectives. On that matter, Statistics New Zealand, Statistics Netherlands and NASS initiated their overall projects with well-defined objectives. In-scope and out-of-scope processes are clearly listed, not only to drive the governance but also to avoid misinterpretation by affected employees.

Such projects generally try to streamline processes within a new architecture where tasks and activities are centralized and standardized. The several processes involved increase the impact that any failures would have on the overall success. This translates into

a new risk: the interdependency of transformed processes. We would then recommend keeping transformation tasks as small as possible and spreading these into realistic timelines to avoid overloaded schedules.

The determination of priorities comes after the objectives. It was noted that most NSOs adopted a reactive approach in motivating priorities based on obsolescence of processes and systems. We believe that a more proactive strategy should be considered. A long-term development planning (say over ten years) could be initiated, and then updated on an annual basis. This would ensure a continuous maintenance of the statistical programs. It would also avoid surprises in the budgeting process, and would serve as input for a human resource plan. Statistics New Zealand provides a good example with such a long horizon.

Employee involvement or buy-in is essential to the success of the transformations. The articles offer different approaches, that is, top-down or bottom-up. From our experiences, miscommunication may emerge when employees do not understand the pressure the organization is facing. Reluctance to change may also be observed while ideas may contradict each other. We would then recommend planning for an appropriate combination of top-down and bottom-up approaches. Senior management would describe the indisputable pressures facing the organization, consult end users, identify areas to be improved, set targets and action plans, invest in a communication strategy, and then listen to employees' concerns while looking for innovative ideas on their part on how to implement and optimize a transformation.

### *3.2. Governance*

Governance is probably the most important aspect of large projects. Without governance, the project scope, schedule and budget would not be monitored closely enough to ensure the success of the project. Statistics Netherlands describes a well-structured governance model that involves a board of directors, a portfolio board, an architecture board, and steering committees dedicated to each project. Statistics New Zealand implemented a very similar structure with a transformation team, a standards review board, portfolio committees, and steering committees. While NASS also has a similar structure, they go beyond the Statistics Netherlands and Statistics New Zealand model by setting a change control board to govern the transition to the new architecture – an exemplary approach.

Statistics Canada has put in place a similar governance structure to those demonstrated by the NSOs. The governance starts with a corporate management framework, that is, a vision that lays out the mandate and objectives of the organization and the related vulnerabilities. It brings together management principles to enable financial and operational success as well as to monitor the interdependency of projects.

Statistical organizations would benefit from a project management office, on site, to assist senior management and project managers within the governance structure. Such an office would offer advice and tools to guide project managers in their activities and assist them in documenting their projects throughout their phases.

### *3.3. Cost-Benefit Analysis*

All articles in this special issue describe cost-benefit or business case analyses; some more explicitly than others. Statistics Netherlands states it very well: "The goal of

cost-efficiency is only reached if cost-benefit considerations are correctly taken into account when taking decisions, of course. We look at cost-benefit considerations first for the architecture, then for deciding on standardization and finally for project decisions on individual (re)design projects in the context of portfolio management."

Cost should not only be measured financially but also in terms of workload, expertise required, and capacity to manage large and complex projects. In the articles, NSOs do not formally address transition costs from the former to the new architecture. This requires considerable effort, especially in migrating applications, training staff with new systems, and putting in place support units. We would recommend neither to underestimate transition costs nor ongoing maintenance costs when planning transformations.

The transformation projects described by the NSOs are all driven by financial pressures, but to varying degrees. Pressures to do more with less have been around for many years, but particularly so recently; it seems likely to remain this way in the years to come. However, financial pressures should not be the only driver for transformation. It should be initiated because and when NSOs can, not when NSOs must. Statistics Netherlands' strategic decision to transform processes is a good example of this. That said, a portion of any savings harvested from a transformation should be reinvested in research and development or other transformation initiatives to further improve the efficiency or the quality and quantity of information delivered by the NSOs.

### *3.4. Statistical Tool Box*

The development of a tool box that offers key statistical functions is also a good practice suggested by the NSOs. Whether the tools are standardized, generalized, centralized, consolidated, or common, these tools are to be used by most statistical processes within organizations. A tool box is a way to reduce duplication of functions, to help guarantee their sustainability, to ease the transition of staff across surveys, to reduce implementation errors, to focus research efforts, and to ensure sound methods.

Any tool box must be maintained. The supported functions must stay in line with the survey requirements. Therefore, resources must be set aside to ensure the relevance of tool box components. This means research activities must be funded to keep up with state-of-the-art methods, as alluded to in Section 3.7. Tools must be classified as emerging, recommended, phasing-out or obsolete, and their renewal should be included in the organization's long-term plans. Furthermore, governance must prescribe mandatory tools and have exceptions considered, approved and monitored by corporate review boards.

Given all its positive aspects, a tool box is not free of issues. For the employees, a tool box may be perceived as an obstacle to their creativity. In the envisioned environment, statisticians should dedicate their valuable time to making good and innovative uses of standardized systems and conducting research into improving these systems, rather than redeveloping customized systems that happen to be fairly similar to each other. For other employees, there is a risk of the press-button syndrome where statisticians simply run systems with a set of parameters without proper theoretical knowledge on how the system works and critical look at what the system produces. Such behavior may generate costly errors, which could include revising estimates and could contribute to a reduction of end

user trust. This should be avoided by ensuring that staff have the appropriate training and knowledge to run these systems effectively.

### *3.5. Data and Metadata Management*

Data and metadata need to be consistently and carefully managed, and appropriate decisions must be made upon the known quality of a particular dataset. Automation of processes requires metadata (parameters) to control the flow of a process. Reuse of data products requires clear metadata to support its use within quality and privacy constraints. All three articles have some form of data and metadata management platform as a key component of their vision. These data service centers or common metadata databases ensure a consistent, reliable management of data and metadata with high quality. Statistics New Zealand highlights a key set of attributes for their platforms which enable the creation of statistical processes through specifications with little to no custom IT development required. This puts the power of process design in the hands of methodologists and statisticians who can create new capabilities in a flexible and timely manner.

Statistics Netherlands highlights the importance of managing "steady-states", and starting the process of design from the analysis of outputs, then designing the data, and finally the process. Chain management is highlighted to provide cross-process coordination. Data Service Centre concepts are introduced to provide careful management of these "steady states". Statistics New Zealand similarly notes the importance of data and metadata management, and highlights innovative approaches they have taken to associate powerful metadata with their data records. In addition, they highlight a shift from processing-centered data to analysis-centered data with the adoption of OLAP-based techniques optimized for analysis as opposed to more conventional OLTP-based techniques, which favor transaction throughput. They adopt powerful architectural principles (e.g., "Administrative data first"), reflecting their transformational shift away from traditional collection. NASS places the creation of centralized data and metadata management at the heart of their vision, remediating the traditional decentralized and *ad hoc* approaches used in their distributed field office environment. They leverage their infrastructure centralization and virtualization to create a core set of data and metadata databases, and similarly reflect on differences between OLTP and OLAP implementations.

Collectively, the focus on data and metadata management and use implies the need for a central information architecture, which provides the reference framework for governance, standardization, and specification to support these activities.

### *3.6. Processing*

Several aspects of the survey processing are being improved by the three NSOs. Selective editing is one of these. This technique contributes to the reduction of the amount of manual work and follow-up activities devoted to observations that have negligible impact on statistical products. In the same vein, an iterative process – also referred to as responsive design – is proposed by Statistics New Zealand to derive estimates with a reduced set of sampled units. This approach uses quality indicators to determine whether collection activities should be modified (maybe interrupted) or not, in order to dedicate the limited resources available to where they will have a higher positive impact. Statistics Canada

works on a similar approach based on rolling estimates (Ravindra 2012). Under this scenario, estimates will be tabulated regularly during the collection phase until they reach a desired level of quality, as measured by some indicators, at which point collection would stop. Along with a selective editing strategy, the scenario will contribute to reduce the number of manual interventions by analysts. Furthermore, it will highlight a connection between the collection process and the quality of the estimates.

### *3.7. Research Activities*

The standard methods and systems that are essential to addressing the current challenges faced by NSOs need to continuously evolve in order to further improve efficiency or respond to future needs. To effectively do so requires research and development, as well as an architecture that can easily add new functionalities or their extensions into the integrated production lines. This enables an evolutionary state of the architecture after it goes through its transformational state. Regardless of any transformation initiative, a minimum support of research and development activities is essential, or the statistical organization will not survive the changes observed in the society it serves.

The role of methodologists is to steer statistical research and development priorities. Typically, the development cycle should go through four main phases: (1) idea generation, (2) proof of concept, (3) prototyping, and (4) integration. The methodologist leads the first three phases in order to contribute to specifications for the fourth phase.

On the other end, the IT analysts steer technical priorities. They are responsible for identifying technologies that can bring benefits to the production of statistical products. This research aspect enables IT development based on methodological specifications. A good symbiosis must exist between methodologists and IT analysts because technological and theoretical challenges are interdependent.

The strategy of the High-level Group for Strategic Developments in Business Architecture in Statistics (HLG-BAS 2012) is to promote international cooperation in which NSOs would share tools in order to reduce development efforts. Sharing can apply at multiple levels. Sharing of architectures and interface specifications can lead to more rapid development and enable future convergence. Co-development of components can lead to an open source approach allowing the community of interest to enhance elements in a controlled fashion. Sharing of executable modules is also effective at saving effort, although it is limited due to the fact that inevitably the owner is the only one who may evolve the module. In the end, the relative maturity of participating organizations and the catalogue of existing components each brings to the table will determine the right mix of approaches.

### *3.8. Evaluating the Transformation*

Although most transformation projects are still ongoing, NSOs should already have thought about the way they plan to evaluate them. Lessons learned are documented but the success of the projects is not measured against initial objectives. An evaluation framework would highlight both the positive and negative outcomes of the transformation, such as: identifying key factors (methodological, technological or environmental), elements or decisions that contributed to successes where other similar initiatives had failed or been

less successful in the past; better understanding of weaknesses in the planning or implementation phases; or, recognizing limitations in the transformation itself. This information is essential to improving the governance for future initiatives.

## 4. Concluding Remark

The three articles from NSOs describe transformations of survey functions, production lines, and management practices that are taking place or being considered in many NSOs. These transformations are essential to ensure that NSOs continue to deliver the information needed by the society that they serve for the years to come. These transformations are ambitious and some will take years to be fully implemented. In order to be sustained, they will have to deliver benefits along the way, even very early on. To achieve this, difficult decisions will likely have to be taken and some may even consist in prioritizing objectives. The key element is to set this transformation on an architecture that will allow the more standardized tools, methods and processes to grow and evolve through research and innovation in a well-governed manner. It is important to communicate from the onset that the new steady state will be globally better, that is to produce as much, if not more, relevant information of better quality at lower costs. It is also important to recognize that this information will be produced differently with fewer locally optimal features. Finally, some may resist standardization, integration and streamlining of processes claiming that it curbs research and innovation. On the contrary, such a transformation is innovation by itself and research is essential to achieving the objectives of the transformation and ensuring that the new steady state meets future challenges.

## 5. References

HLG-BAS (2012). Strategy to Implement the Vision of the High-level Group for Strategic Developments in Business Architecture in Statistics. Document presented at the Sixtieth plenary session of the Conference of European Statisticians, UN-ECE, June 6–8, 2012. Available at: http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.58/2012/10_-_HLGBAS_Strategy.pdf (accessed March, 2013)

MIT-ESD Symposium Committee (2001). ESD Terms and Definitions – Version 12. Working paper ESD-WP-2002-01, Massachusetts Institute of Technology, Engineering Systems Division, October 19, 2001. Available at: http://esd.mit.edu/wps/esd-wp-2002-01.pdf (accessed March, 2013)

Ravindra, D. (2012). The Integrated Business Statistics Program: An Overview. Proceedings of the Fourth International Conference on Establishment Surveys. Montreal, June 11–14, 2012. Available at: http://amstat.org/meetings/ices/2012/proceedings/KESIV_TOC.pdf (accessed March, 2013)