

**UNITED NATIONS STATISTICAL COMMISSION and
ECONOMIC COMMISSION FOR EUROPE**

CONFERENCE OF EUROPEAN STATISTICIANS

Work Session on Statistical Data Editing
(Ljubljana, Slovenia, 9-11 May 2011)

Topic (v): Changing organizational cultures

The Impact of a Changing Business Architecture on Editing

Key Invited Paper

Prepared by Claude Poirier, Statistics Canada

I. INTRODUCTION

1. Statistics Canada has tabled its Integrated Business and Human Resources Plan (Statistics Canada, 2010). This plan focuses on the relevance, quality and accessibility of data, as well as organizational efficiency. The efficiency aspect targets the production of the largest possible volume of high-quality information from the available resources. This principle is not new; indeed Statistics Canada has mounted repeated initiatives to promote efficiency since the 1980s, a suite of generalized systems being one of them. The suite was initiated to offer standardized and harmonized tools to users, with the intention of ensuring data portability in support of all survey steps, from design to dissemination.

2. The Corporate Business Architecture (Statistics Canada, 2009b) also requires harmonized and reusable tools. One of its goals is to make sure generalized systems and processes, wherever they exist, are well-maintained and documented and are used by most programs. Under this goal, no program should be allowed to develop functionality already available through a generalized system unless approved by the System Architecture Review Board.

3. The generalized systems have evolved since their creation, some to a greater extent than others (Kovar, Jeays and Poirier, 1999; Poirier, 2010). Their maintenance and support depend on the complexity of the functions and whether they are menu driven or not (Poirier, 2004). Generalized edit and imputation functions have evolved through two main systems: BANFF and CANCEIS. The first system handles purely numerical data while the second can process a mixture of categorical and numerical data such as the data collected by the Census of population. The migration of other customized systems to one of the two generalized systems is currently being investigated. Subsequently, the pros and cons of a consolidation of the two main systems will be evaluated.

4. Section II presents the recent changes at Statistics Canada, especially the management initiatives that must be considered in identifying future directions for statistical data editing. Section III gives the details of the actual editing tool kit and suggests how it can be reduced. Furthermore, it describes ways to improve the resulting tools and how they may be consolidated. Section IV describes the impact on the research and development program, the human resources and the international collaboration efforts. The paper concludes with some observations in Section V.

II. CHANGING CULTURE

A. Governance and management

5. Statistics Canada's Integrated Business and Human Resources Plan identifies organizational efficiency as a pillar. For a few decades, the planning system has solicited and approved proposals for investments to improve efficiency. As of now, increasing efficiency requires that Statistics Canada innovate its business practices. Processes are being examined from various angles: how information is managed, how systems are built, how programs are governed. In this context, all assumptions are questioned and greater emphasis is put on global (corporate departmental-level) efficiency, as opposed to the historical local (division-level) efficiency. The result of standardization and centralization of processes is that specific programs may lose some flexibility for the best interest of the corporation as a whole. A risk management strategy has been established to make sure that strategic outcomes are still met.

6. Global efficiency will be achieved through various initiatives, including the maintenance of generalized systems and the enhancement of the governance structure that applies to their use. Every survey program that enters into a redesign phase now has to adhere to a formal project management framework. The framework guides managers through the five steps of a typical project: idea generation, project initiation, planning, execution, and close-out. While doing so, managers report to a set of senior management committees, depending on the size of the project. The Corporate Business Architecture Management Committee (CBAMC) is one of them; it developed a check list to make sure that all projects take advantage of the applicable corporate tools and services. In this check list, generalized systems are mentioned as the recommended way to run statistical processes.

7. This governance model has evolved through the decades. When the idea of generalized systems was born in the 1980's, the initial intention was to let managers adopt them as they felt beneficial for their project. A prototyping approach was hidden behind this "if you build it, users will come" principle. Indeed, the requirements of a reduced set of programs were initially targeted by the generalized systems. After several years of fine-tuning and enhancements, more programs could benefit from the extended user-friendly functionality and so systems became recommended. After several more years of development, users were strongly encouraged to investigate generalized systems in lieu of customized system. Later, users were forced to investigate these corporate assets and today they are obliged to use them. Customized systems are now accepted only when an exception case is built and presented to the System Architecture Review Board, and when the board accepts it.

8. The System Architecture Review Board (SARB) is the entity that governs the set of corporate tools and prescribes their use. In more detail, it develops and promotes an enterprise-level system architecture, it specifies and prescribes applications and system development standards and tools to be used, it conducts and guides system project reviews and manages exceptions, and it reviews system risk assessments. While managing exceptions, the board challenges managers for longer term solutions that would better fit the corporate directions, or else recommends enhancements or development of new corporate tools. The CBAMC works in close partnership with the SARB

9. The CBAMC responds in part to the challenges facing Statistics Canada by addressing significant aspects of the quality and cost dimensions. Some of its principles are related to the development and management of tools and methods. They are:

- (a) Metadata: the goal is to make sure metadata is integral to the process and, indeed, drives the process. Metadata should precede the data rather than be constructed at the end.
- (b) Mandatory use of corporate assets: under the authority of several management committees, advice and proposals are managed to ensure the use of corporate services, tools and methods.
- (c) Minimize tool kits: Statistics Canada must minimize the number of software and productivity tools and methods it deploys.
- (d) Proficiency in corporate tools: all employees should be autonomous in the use of generic tools.

- (e) Expand electronic data collection: electronic modes of collection should be offered for most surveys, and used as the sole mode of initial collection wherever feasible.

B. Recent actions and initiatives

10. In response to the evolving governance and management principles described above, a Generalized Systems Planning Committee was recently put in place to investigate how the statistical functions could be better integrated in the generalized systems. Its mandate is to define a methodological vision with respect to the development of generalized systems, especially those related to sampling, editing, imputation and estimation. On that matter, the committee is responsible for defining the scope of the systems and their development priorities given current and future needs of the programs.

11. Resources were put aside to define a development plan that would satisfy the major initiatives, including the Corporate Business Architecture, the Integrated Business Statistics Program, and the Household Survey Strategy. The result will drive budget proposals through the agency long term planning exercise.

12. The committee's working plan includes (a) the documentation of all the statistical tools currently used to support the agency's programs, (b) the identification of programs' needs – current and future – in terms of methods, (c) the solutions resulting from research activities, (d) the prioritization of integration and development activities given gaps, if any, in the current generalized tools, (e) the analysis of opportunities from an international collaboration network, and (f) the definition of a system architecture based on past successful development initiatives.

III. IMPACTS ON EDITING PROCESSES

13. The changing culture has several impacts on statistical processes. This section focuses on the impacts on the edit and imputation processes.

A. Current tool kit

14. The current editing toolkit has several components that present some similar functions. Amongst the tools used at Statistics Canada, we have BANFF, CANCEIS, IMPUDON, AG2000, StEPS and several customized systems.

15. BANFF is the recommended imputation system for business surveys. It was initially architected using Oracle tables, making it heavy to run, before being migrated into a more flexible and user-friendly SAS environment. It performs automatic imputation while respecting specified edits. The system applies linear programming techniques to conduct the localization of a minimum set of fields to be imputed, and then search algorithms are used to perform the imputation. The processing is entirely driven by linear rules defined on numeric variables. Univariate outlier detection is based on quartile distribution. The outlying observations can be flagged for imputation or simply excluded from subsequent calculations. The imputation function offers three imputation methods: deterministic, donor, and estimator. Deterministic imputation identifies cases in which there is only one possible solution given the edit rules. Donor imputation uses data from the nearest neighbour and the estimator function provides a wide set of models using historical or current information.

16. The Canadian Census Edit and Imputation System (CANCEIS) is also a corporately supported imputation system. It was initially developed to treat the categorical variables of the 1996 Census of population. Its functions were expanded to process a mixture of categorical and numerical census variables. This now makes the system suitable for some social surveys. It can perform both deterministic and donor imputation. Its goal is to minimize the number of changes while making sure the imputation actions are plausible. It uses edit rules defined by decision logic tables to identify records that need imputation and records that can be used as donors. For each record to be imputed, the system tries to find the nearest record that can be used as a donor and identifies the optimal imputation action by borrowing as few donor fields as possible to impute the record in error.

17. IMPUDON is a suite of SAS macros that perform donor imputation. It can be used to process numerical or categorical data. It offers several scaling options for distance functions. It also offers some flexibility in donor imputation, for example matching fields can be weighted, it is possible to pick a donor at random from the k nearest donors, and some potential donors can be excluded. IMPUDON does not however offer any editing or error localization tool, it simply offers donor imputation functions.

18. AG2000 was developed for the Agriculture Statistics Program. It offers several editing functions with a high connectivity with the Farm Register. It processes out-of-scope units, changes of farm operators, amalgamation of farms and dissolved partnerships with the appropriate mixture of zero data, missing data and reweighting flags in preparation for the imputation or the estimation phases. Its imputation principle is based solely on a random donor method. Records are first grouped into imputation groups within which missing data are imputed from a random donor. Its strength is its connectivity with the register however this becomes a weakness when one wants to use it with another register.

19. StEPS stands for the Standard Economic Processing System developed by the U.S. Census Bureau. StEPS offers simple edit checks such as ascertaining the presence of data values, range verifications, and verifications of valid categories. It also provides more complex tests such as balance tests which verify the summation of items against selected totals, and survey rules to verify field relationships within observations. With respect to imputation, the system offers deterministic imputation, imputation by auxiliary data items, sum of data items, historical values, means, trends, ratios, and multiple regressions. It is used for the surveys of manufacturers in Canada due to its data point architecture that eases the repetitive processing of commodities.

20. Other customized systems are used within Statistics Canada. One of them was developed 15 years ago to process revenue data for the Survey of Labour and Income Dynamics. It was developed to provide more flexibility than the generalized systems with respect to the definition of donors in the donor imputation. It is used to impute detailed revenue variables in social surveys. Its methodology was recently modified to satisfy other surveys, including the Survey of Household Spending. The Labour Force Survey also has its own customized system, this one very simplistic given the short questionnaire and the tight processing schedule. It is mainly based on validity rules, with no imputation action except reweighting in the case of edit failure or non response.

B. Reducing the tool kit

21. A reduction in the number of software components, tools and methods has already been mentioned in section II above, as a goal of the Corporate Business Architecture. To be reached, this goal requires serious effort and non negligible budgets. Furthermore, its implementation and the migration of applications should not disrupt any statistical programs.

22. To do so, the first step for Statistics Canada is to identify the tools to be targeted by most statistical programs. The Generalized System Planning Committee has already analysed this issue for the edit and imputation processes. Its recommendation was to continue with the current corporately supported generalized systems. These are BANFF, the generalized edit and imputation system, and CANCEIS, the Canadian Census Edit and Imputation System. This does not mean that the two systems satisfy all the requirements of the individual programs. On the contrary, there are many particular aspects to the statistical programs and their generalization would become too costly in terms of development and maintenance effort. Under these conditions, users have to develop their own pre-processors and post-processors to support data reformatting, data connectivity, metadata and integration with other systems. This may eventually become less of an issue with the standardization of the IT architecture. On the editing front, the program managers are now expected to adhere to the set of functions that are already generalized (to reflect the global optimization philosophy) even if it is at the cost of local efficiency. This is possible by making sure a set of robust methods and systems are in place to satisfy most needs, even when data sources differ by their nature and when underlying assumptions change over time. Since there may always be exceptions to this “one set fits all” rule, cost-benefit exercises must be well documented in order to present exceptions to the System Architecture Review Board.

C. Improving the general functionality

23. The quest for robust methods is usually not easy. Statistics Canada started to develop common editing methods for a reduced set of surveys and managed improvements to satisfy more needs thereafter. The resulting corporate methods may not be the most efficient ones for every particular case but they are methodologically sound and they show the robustness that is required to face moving targets. While exceptions to corporate tools are being well monitored, they serve as a source of information that guides more enhancements of generalized systems. In return, such enhancements put the corporation in a better position to reinforce the use of generalized systems. In addition, enhancements constitute the evidence that users expect from a live system, without which a system would become redundant.

24. For data editing, the main complement to BANFF is IMPUDON. This program was developed to offer a donor imputation framework in a SAS environment at the time when GEIS – Generalized Edit and Imputation System, the predecessor to Banff – showed limitations in terms of user-friendliness given its SQL dependency. As a result, several applications were developed in SAS, with IMPUDON or with customized systems. The main issue with these applications was their lack of corporate support. As a risk management initiative, BANFF has since been developed in a SAS environment with a well organized methodology and technical support service. There are still several IMPUDON or customized applications that have survived until now. Their requirements are now steering enhancements of the corporate tool, as follows.

- a) Outlier Detection: more flexibility is required in the Hidioglou-Berthelot method that uses quartiles to define a distance metric for the sake of outlier detection. A distance based on the percentiles would offer more flexibility. A sigma-gap method would also be useful for the business surveys. This method finds a gap between two successive values that is at least as wide as the standard error of the group and flags as an outlier all values that come after this gap. Finally, the top-k method that flags as a potential outlier the k largest values would also be used by several applications. This very simple method may be used to guide the manual verifications or to exclude units from further processes.
- b) Error localization: when the edit and imputation steps are run several times, one after the next, the error localization function should target the previously imputed values as potential errors rather than to target real reported values. This would help the principle of minimal change in the reported data.
- c) Donor imputation: the nearest neighbour method introduces a distance metric. The current one uses the L^∞ metric applied on ranks while several applications would benefit from more choices, for instance L^1 or L^2 based on standardized values. Furthermore, the user should be able to control the number of times a donor is being used.
- d) Diagnostic reports: more diagnostic reports may be introduced to help the user in understanding the processes. However, the user should be warned not to over-analyse every aspects of the edit and imputation otherwise this manual work would become counterproductive and go against the efficiency.

25. Requirements from a few other surveys will be addressed in a second enhancement phase. An Integrated Business Statistics Program (IBSP) is currently in development. This project is an opportunity to bring more statistical programs into the Unified Enterprise Survey model. While the model is being rethought, its goal is to streamline the statistical processing, including the edit and imputation steps. On the social side, a project was initiated to develop common tools for the Social Survey Processing Environment (SSPE). Like IBSP, the SSPE project already identified some gaps in the editing tools. The two projects will develop new editing functions that satisfy the global efficiency flavour. The functions that follow may eventually be integrated into the generalized framework.

- (a) Selective editing: selective editing is a procedure that targets only some of the micro data items or records for review or follow-up activities by prioritizing the manual work and establishing appropriate and efficient process and edit boundaries. It is usually driven by the size of the units

or by characteristics through a score function or a set of quality indicators. Selekt – a Swedish system – will be considered to meet selective editing requirements.

- (b) Connectivity: connectivity refers to the interaction with a central database in order to avoid managing independent files that would otherwise be created by the individual processes. Typically, this means the edit process would first access the database to get the raw data and the metadata, then run the edits and store results back in the database. Next, the imputation process would read the database to get the list of variables requiring correction, then perform corrections and store imputed data back in the database.
- (c) Metadata: as explained in section II, metadata should be integral to the process and, indeed, drive the process. Furthermore, metadata should be integral to the connectivity aspect in order to precede the data rather than be constructed at the end.
- (d) Editing of e-collection: the benefit of early data editing is already well known. E-collection offers such an opportunity. While functions for self administered e-editing constitute a subset of regular batch editing functions, they differ by the environment in which they are applied. Making the batch functions accessible from an e-collection environment would make the statistical data flow smoother.

26. With respect to CANCEIS, its development plan is still guided by the quinquennial Census of population. The 2011 census is being prepared and all the editing resources are being dedicated to the CANCEIS applications for the census short-form and the new National Household Survey long-form. A review of the CANCEIS functionality may take place after the census data processing in order to facilitate more social surveys.

D. Consolidating existing functions

27. CANCEIS is intrinsically different from BANFF since it finds the donor before identifying the minimum imputation while BANFF does the opposite. It is believed that the two systems deserve a place in the generalized suite because one would not perform well in a purely numerical context and the other cannot process categorical variables.

28. However, from a maintenance and support perspective, there may be benefits of bringing BANFF and CANCEIS into a unique environment. The success of the BANFF development project and the flexibility of the resulting functions based on SAS components served as a model for the redevelopment of the corporate system for disclosure control, CONFID2. Now, the same model is being used for the generalized sampling system G-SAM and it is being planned for the generalized estimation system G-EST.

29. The SAS model can be easily justified for editing functions. First, SAS has been adopted by the corporation as a standard for statistical data analysis and processing. Hence, the adoption of SAS for the editing systems is in line with this strategic direction and does not involve any extra costs on the part of the user community. The second reason is even more important: SAS is a highly suitable and effective technology that provides a rich set of functions and modules for statistical processing and analysis. An editing system simply enriches this functionality. Also, SAS implements a runtime environment highly scalable and suitable for processing a high volume of data in batch mode – a typical scenario in data editing processes. Since pre-built SAS procedures can be called independently, the user is given the opportunity to insert his own SAS processing instructions between procedure calls.

30. These facts, along with the evolving culture in favour of a common IT strategy, suggest that SAS may support a consolidated editing environment for BANFF and CANCEIS. On the other hand, such a long-term vision is competing with the aim to have CANCEIS meet the evolving census requirements. A study of strengths, weaknesses, opportunities and threats of such a SAS model for the editing of both numerical and categorical data would be required before bringing the idea into the scope of the corporate 10-year investment plan.

E. Migration plan

31. The plan to migrate editing applications to the generalized systems is relatively simple. It is based on the following principle: don't fix something that is not broken. This means the current applications will be in scope for migration when they show serious risk of failure or have issues related to their maintenance, or when a change in the program needs requires major modifications to the editing applications. At that point, applications become candidates for moving to generalized systems.

IV. IMPACTS ON OTHER DIMENSIONS

A. Research initiatives

32. The Statistical Research and Innovation Division (SRID) is relatively young. It was created in 2006 for researching, developing, promoting, monitoring, and guiding the adoption of new and innovative techniques in statistical methodology in support of statistical programs. Its mandate includes the provision of technical leadership, advice and guidance to employees elsewhere in the Methodology Branch or in the department. This assistance is in the form of advice on methodological problems that arise in existing projects or during the development of new projects.

33. SRID also monitors research activities initiated in the Methodology Branch and sponsored by the Methodology Research and Development Program (MRDP). While SRID is young, the MRDP was created about 25 years ago. The program has been very successful for carrying out advanced research in methodology, and implementing their results. The main development initiatives included the development of unified methods in estimation, editing and imputation, data analysis for complex surveys, confidentiality, time series and record linkage. These methods resulted in a set of generalized software that included GES (estimation), BANFF and CANCEIS (edit and imputation), SEVANI (variance estimation due to imputation), and CONFID2 (disclosure control). Recent research activities have included the development of small area estimation techniques and associated flexible software to estimate below the sampling level, fine-tuning variance estimation due to imputation, outlier detection techniques, and the analysis of score functions in a selective editing context.

34. The corporate desire to improve data collection in terms of efficiency and quality has had the following impacts on the methodology research program:

- (a) Nonresponse follow-up: the objective of this research is to optimize follow-up rates within groups of units to maximize the overall quality while minimizing the editing costs. Impact on potential bias is the main concern of this work.
- (b) Mode effect: it is clear that the trend for collection strategies relies on a mixture of collection modes. For example, electronic data collection is increasingly being used. This type of self-reporting introduces a risk of misinterpretation on the part of respondents. However, we do not have measures on how it impacts on data quality and comparability. Furthermore, since respondents can provide different responses among the different modes, it becomes essential for the editing process to ensure that responses across modes are comparable, and if they are not that there is a mean to standardize them.
- (c) Quality indicators: these should summarize the state of the survey data by reflecting the outcomes of the editing process. Response rates, coverage errors, measurement errors, error rates, imputation rates and resulting bias are the existing indicators that summarize the quality of the data. There are also emerging indicators (R indicators) in the literature that reduce non-response bias in the estimation process. The properties and usefulness of these indicators need to be better understood in the data collection and editing context.
- (d) Selective editing: the challenge related to the editing of a reduced set of data points consists of identifying a robust approach that will perform well for most variables and most domains. Since selective editing may compete against bias, the goal is to find the right balance between cost reduction and induced bias.

35. Unfortunately, these areas cannot all be examined given the limited research budget. Therefore, it is important to ensure that the research program be managed so as to prioritize activities likely to provide results that enhance relevance, quality, accessibility and efficiency.

B. Efficiency with respect to human resources

36. The impacts of the evolving business architecture on editing principles have already been mentioned with the reduction of recommended tools, their mandatory use, the phased-in migration, the extension of functions within the editing tools, and the coordinated research activities. All these impacts are related to the technical aspects. The current section now addresses human resources aspects.

37. When the review of the corporate architecture was initiated, employees had the impression that it would become an obstacle to their creativity. From some angles they were right, but from many others we realize they weren't. In the world of editing, initiative and creativity should be demonstrated through well thought out applications. For instance: useful edit checks, well parameterized socio-economic laws, a good understanding of strengths and weaknesses of proxy and historical values for the purpose of editing, and the choice of sound actions for correcting failed-edit records. Therefore, one should dedicate one's valuable time to making a good use of a standardized editing system, rather than redeveloping customized systems that happen to be similar to each other after all. In that context, a BANFF user for instance would not simply run a sequence of preset functions, he would rather try to understand the data and identify the best parameters for his editing application. More importantly, he should interpret editing outcomes in order to fine-tune the design for the next cycles of the statistical program. On that topic, Granquist, Nordbotten and Kovar (2006) say that the role of editing is enlarged to include primarily the identification of error sources and, only secondly, identification and correction of important errors. Providing intelligence related to survey design is now one of the goals of editing.

38. From Statistics Canada's point of view, professional knowledge and employee job satisfaction are issues of concern. Putting emphasis on "learning from the data" is an effective way to improve employees' abilities and should be considered as a source of pride by the employees. The use of a standardized editing approach will ease the employees' learning curve in the long term, and make the transition from project to project more affordable. This portability of staff will be beneficial to the employees by enriching their job opportunities, as well as to the corporation by introducing efficiencies in productivity and creativity.

C. International collaboration network

39. The idea of an international collaboration network started about a year ago. Under the leadership of Australia, a meeting with senior managers representing the interested countries was organized in Paris last June. The participating countries, Australia, New Zealand, Sweden, Norway, UK and Canada, agreed to share knowledge on the following six opportunities: Innovation, Confidentiality, Metadata, Interoperability, Editing, and E forms (ref.: Value Creation Group, 2010). The fifth opportunity – the industrialization of editing – has already started to take shape. The partners shared their vision and a business case will be presented as part of agenda topic (vi) of the 2011 UNECE data editing work session.

40. This international opportunity is similar to the Statistics Canada generalized systems project that was launched in the 1980's, but on a larger scale. Its goal is to remove the barrier between National Statistical Institutes in order to bring together sound and well tested editing functions. On that matter, BANFF and CANCEIS are already in a mature state. These flagship systems are already used around the world and their functionality is highly recognized. The collaboration network may help Statistics Canada to determine the benefits of maintaining the two systems, building new functions on top of existing ones, and maybe consolidating the two systems into one. Statistics Canada will also benefit from other NSI's tools. This may fulfil the need for additional functions as listed in section III above, selective editing being one of them.

V. CONCLUDING REMARKS

41. Statistics Canada's priorities include accessibility, relevance, quality of data, and organizational efficiency. In order to address the efficiency aspect, the corporate business architecture is being reviewed. This initiative has an impact on all processes, including data editing. Among the goals we find the use of common systems, the reduction of system maintenance, the quest for robust methods, the search for global rather than local efficiencies, the shorter learning curve from a user perspective, the portability of staff, and more. Editing systems will be reduced in order to avoid duplication of methods, thus increasing the use of corporately supported systems (BANFF and CANCEIS). On the other hand, the functionality of generalized systems will be improved given needs and priorities, with the result that some applications may lose in order to achieve a global gain. While production systems are being centralized, research activities are still going on with close coordination. Behind all this, opportunities will most likely be created in order to get a better understanding of the error sources and to initiate changes to the design that would address them.

VI. ACKNOWLEDGEMENTS

Acknowledgements to Michael Hidioglou, Claude Julien and Jacqueline Mayda of Statistics Canada for their contribution.

VII. REFERENCES

- Granquist, L., J. Kovar and Nordbotten, S. (2006). Improving Surveys – Where Does Editing Fit In? *Statistical Data Editing. Vol. 3: Impact on Data Quality, Chapter 4: Looking forward.* Conference of European Statisticians. United Nations Statistical Commission and United Nations Economic Commission for Europe. <http://www.unece.org/stats/publications/editing/SDE3.htm>
- Kovar, J., Jeays, M. and Poirier, C. (1999). Generalized Systems: Where are we at and where are we going. Statistics Canada internal document presented to the Advisory Committee on Statistical Methods, meeting #28, April 1999.
- Poirier, C. (2004). The Processing Environment Behind a Statistical Program. *Proceedings of the Survey Methods Section*, 2004 Joint Statistical Meeting of the American Statistical Association, August 2004.
- Poirier, C. (2006). A Functional Evaluation of Edit and Imputation Tools. *Statistical Data Editing. Vol. 3: Impact on Data Quality, Chapter 2: Quality Measures.* Conference of European Statisticians. United Nations Statistical Commission and United Nations Economic Commission for Europe, pp. 146-156. <http://www.unece.org/stats/publications/editing/SDE3.htm>
- Poirier, C. (2010). Planning for our Generalized Systems. Statistics Canada internal document presented to the Advisory Committee on Statistical Methods, meeting #51, October 2010.
- Statistics Canada (2009). Report of the Task Force on Corporate Business Architecture. Statistics Canada internal document, April 2009.
- Statistics Canada (2010). Integrated Business and Human Resources Plan. Statistics Canada internal document. <http://www.statcan.gc.ca/about-apercu/plan2010-2013/pdf/plan2010-2013-eng.pdf>
- Value Creation Group (2010). Informal CSTAT Workgroup on Stronger Collaboration on Statistical Information Management Systems – Draft Output Report. International Conference, Paris, June 2010.