

Distr.
GENERAL

Working Paper No.18
29 April 2009

**UNITED NATIONS STATISTICAL COMMISSION and
ECONOMIC COMMISSION FOR EUROPE
CONFERENCE OF EUROPEAN STATISTICIANS**

**ENGLISH ONLY
EUROPEAN COMMISSION
STATISTICAL OFFICE OF THE
EUROPEAN COMMUNITIES (EUROSTAT)**

**ORGANISATION FOR ECONOMIC COOPERATION
AND DEVELOPMENT (OECD)
STATISTICS DIRECTORATE**

Meeting on the Management of Statistical Information Systems (MSIS 2009)
(Oslo, 18-20 May 2009)

Topic (ii): Architecture

COLLECTION ARCHITECTURE AT STATISTICS CANADA

Invited Paper

Prepared by Karen Doherty, Statistics Canada, Canada

I. INTRODUCTION

1. Statistics Canada (StatCan) supports a comprehensive statistics program covering a wide range of social and business statistics, the National Accounts and a quinquennial national census. The information used to develop these statistics comes from a combination of survey data and administrative or registry data.
2. Statistics Canada maintains eight regional collection hubs in addition to the Head Office in Ottawa. The collection activity undertaken in the regional sites includes face-to-face interviewing (via CAPI – Computer Assisted Personal Interview), telephone interviewing (via CATI – Computer Assisted Telephone Interview, often supplemented by Random Digit Dialing techniques), reporting via paper questionnaires using Canada Post, electronic data reporting, and, especially in the case of the Census, reporting via a Call Centre.
3. A number of challenges facing the Agency led to a rethinking of how data collection activities should be undertaken. This report describes the approach taken to look at the risks to our programs foreseen over the next few years and identify solutions to mitigate these risks while maintaining the integrity and quality of our statistics.

II. BUSINESS ARCHITECTURE

A. Business Drivers

4. Canada's environment has changed significantly over the past twenty years. The population is concentrated in cities, most of which are located within 100 kilometers of the US border. The use of advanced telecommunications technology is widespread, and, especially among the younger members of the population, the use of, and demand for, services on-line is growing fast. Canada however is a vast country geographically, with people scattered across 10M km² of land with remote areas that are often hostile to human habitation. It is important for the well-being of all Canadians that all communities, not just those located in the southern zones, be represented appropriately in our statistics.
5. The changes in our Canadian reality require that our fundamental assumptions about the services and approaches we use to collect information be challenged, and solutions found to adjust these services to meet

new and changing needs. The following business drivers are forcing us to reconsider how we should conduct our collection activities now and in the future.

- **Cost:** Canadian tax payers are asking governments to deliver services at a lower cost. The government's need to address the fallout from the current economic downturn amplifies this need for costs saving measures.
- **Communications Technology:** Canadians, especially those under 30, rely more and more on cell phones, and in many cases no longer even have a line land. This has significant impact on the success rate of telephone interviewing and random digit dialing. StatCan needs to adjust its telephone interviewing services to meet this new reality.
- **Electronic Reporting:** Canadians, especially those under 30, increasingly communicate via the Internet and expect that services be offered on-line. The success of the 2006 Census Internet Response Channel underlines Canadians' desire to move to internet-based data reporting.
- **Geography:** Not all of Canada's population has access to reliable and cost effective telephone and internet services. Face-to-face interviewing will still be required in our remote areas for some time to come.
- **Admin Data/Registries:** StatCan is increasingly using administrative data sources to replace or supplement survey data. Business surveys now make extensive use of Canadian tax data and on the social side, extensive use is being made of Health and Justice related data. The Census relies on an Address Register that is maintained partially through information from Canada Post, Elections Canada and utility companies. That being said, Canadians are not yet ready to embrace the adoption of a personal identity card or of a comprehensive person registry.
- **Multi-mode Collection:** The need to maintain a range of collection options to satisfy the needs of all Canadians will remain for the foreseeable future.

B. Business Architecture Process

6. To address these business drivers effectively StatCan embarked in 2006 on a business architecture project to develop a new vision and direction of the Agency's collection activities. A company was hired via tender to assist us with this discovery exercise. The project questioned all the assumptions and beliefs in place at the time as well as the existing service model. A large number of stakeholders participated in the process to ensure that a full understanding of the business requirements was achieved.

7. The following high level business requirements were identified:

- Use common services, processes and tools for surveys and the Census.
- Provide respondents with a menu of response options so that they can select the option that best meets their needs.
- Allow for greater flexibility so that managers can react quickly to changing circumstances and cost factors.

8. The main recommendations from the study were:

- Adjust the organizational structure to better align with how business should be conducted:
 - Create a centralized accountability structure for collection activities and long-term planning.
 - Revisit capacity, workforce, workflows and costs.
- Streamline the collection process:
 - Separate responsibility for operations from development.
 - Eliminate duplication and overlaps in functions.
- Invest in new technology to create a more robust and cost efficient platform for all surveys:

- Develop a system to control and track all collection activity and allow a case to move seamlessly from one mode to another (Survey Master Control System).
- Discontinue the existing electronic data reporting service and replace it with an Internet based solution (e-Questionnaire Service).
- Implement wireless support for field interviews.
- Move to thin client desktops for collection staff (Virtual Desktop Infrastructure).

9. In 2007, the Operations Field proceeded with the implementation of measures designed to meet the recommendations of the Business Architecture initiative. A reorganization of the Field was completed in 2007 which included eliminating overlapping responsibilities, separating development from operations, and moving the field collection activities to the Central Regional Office. In January 2009, responsibility for Census operations was moved to the Operations Branch. This move will strengthen the links between regular survey collection activities and the Census operations activities and ensure that common systems are developed whenever possible.

10. Several IT projects were also launched in 2006 to address the recommendations related to new technology and streamlined applications. The remaining sections of this paper will cover the IT architecture of these new systems and the approach adopted to implement them successfully.

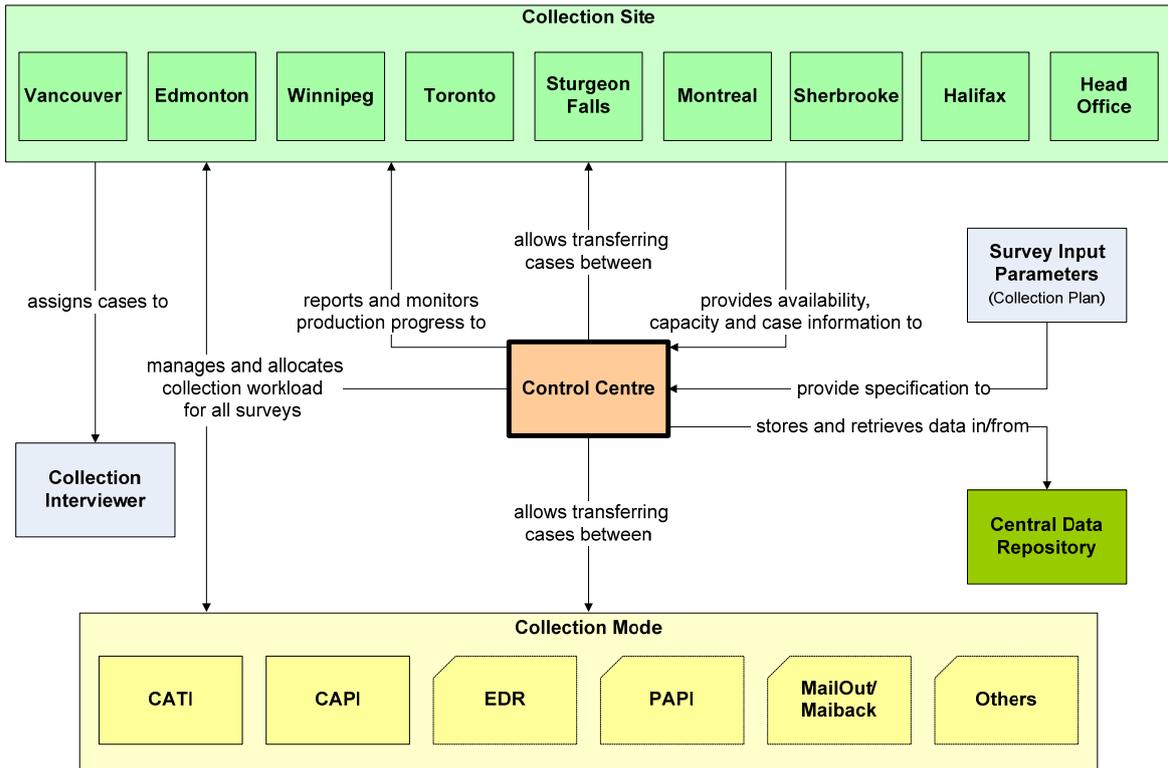
III. SURVEY MASTER CONTROL SYSTEM

A. Architecture

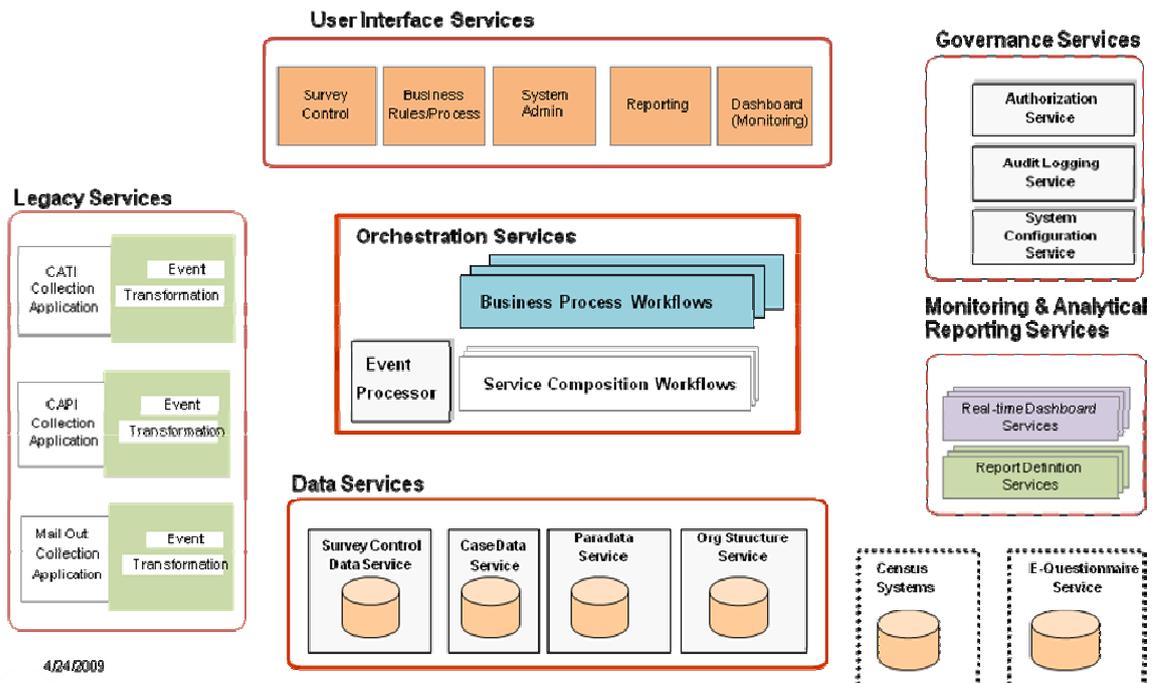
11. The Survey Master Control System (SMCS) is designed to be the central hub through which all collection activity is managed. All other systems perform functions in the collection tool kit but are fed and/or feed the SMCS. All data is stored in a Central Data Repository.

12. The SMCS implements the vision for collections at StatCan as illustrated in the diagram below. Survey cases are stored in the Central Data Repository. Survey managers choose the collection modes to be supported by a particular survey and the Control Centre (the box in the centre of the diagram) manages the transfer of cases between these modes as required. Interviewer case loads are fed to the Collection Sites. Collection managers can view, and download when appropriate, information on all cases regardless of the collection mode used by the respondent. Managers can also monitor the case load within each region as the survey collection occurs over time and move cases around to make the most effective use of interviewer time.

13. Once all the redeveloped collection instruments are operational, all case information will be held in only one place, the Central Data Repository. In the meantime the SMCS will manage the transfer of case load information to and from the tools used to capture survey information. For example, the SMCS will manage the CAPI case load sent to individual interview laptops for interviewing using Blaise until such time as CAPI collection can be performed directly from the Central Data Repository.



14. The application is developed using Service Oriented Architecture principles. The following diagram illustrates the logical components that make up the architecture.



D. Development Approach

15. The project adopted a structured approach, with oversight committees, clear governance and the use of project management techniques to keep the project on track. The steps that were followed included:

- User (business) requirements were identified and documented in Use cases.
- The application architecture was developed.
- The infrastructure architecture was developed.
- The development was outsourced to a private sector firm that was asked to follow the StatCan application development approach. In this case, since the developer chose to use the Microsoft .net technology, they used StatCan's Microsoft .net Development Framework.

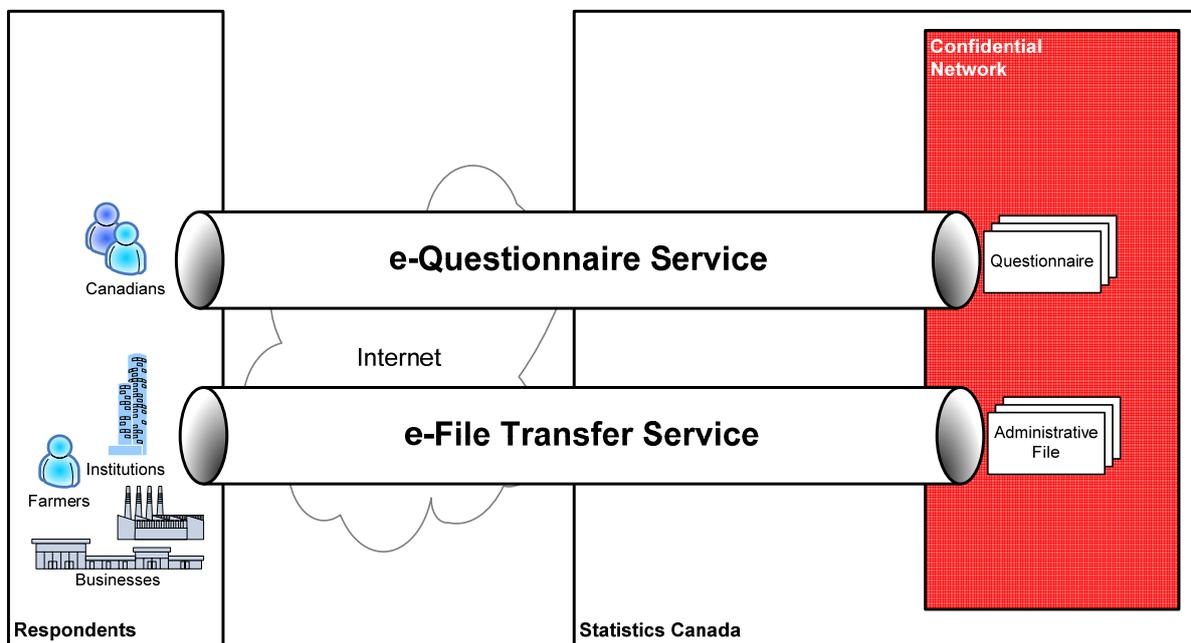
IV. COLLECTION TOOLS

16. In parallel with the development of the Survey Master Control System, Statcan embarked on an ambitious project to replace its aging collection tools. The first phase covered:

- Electronic data reporting (e-Questionnaire service)
- Managed file transfer (e-File Transfer service)
- Tracking, specifically of paper questionnaires, etc. (Generalized Tracking System)

17. The second phase of the project starts in 2009 and is aimed at developing a strategy to modernize the existing interviewer tools including the Case Management System and Blaise questionnaires.

18. The e-Questionnaire and e-File Transfer services have similar data transport requirements as shown in the diagram below.



A. e-Questionnaire Service

19. Following the success of the Internet Response Channel developed for the 2006 Census, it was clear that a growing number of Canadians were ready to move to electronic reporting. The Agency already had an aging electronic data reporting service for business questionnaires, but this service was costly to maintain and had not seen the take-up originally anticipated when the service was developed.

20. The e-Questionnaire Service was designed to provide respondents with an easy-to-use internet based questionnaire that was compliant with Canada's language and accessibility guidelines. The service was not meant to cover every single type of survey that StatCan supports but needed to handle most common types of surveys as well as the Census.

21. The decision was made to acquire a Commercial-Off-the-Shelf (COTS) product through a competitive process and a comprehensive set of requirements were developed in collaboration with a wide range of stakeholders from across the Agency. A contract was signed in March 2009 for IBM Lotus Forms Server and the product is being used this year to support a proof of concept to test whether electronic reporting is a viable response option for the Labour Force Survey.

B. e-File Transfer Service

22. StatCan has the requirement to move data to and from business respondents, in particular for monthly and quarterly business surveys, and to and from organization and government departments that sponsor surveys, or require and/or provide data to the Agency on a regular basis.

23. The e-File Transfer Service was designed to provide users with an easy-to-use internet based mechanism to move data and files to and from StatCan in a secure manner. The Agency acquired a COTS product in 2008 through a competitive process and the solution has been implemented.

24. The new service allows any type of file, of any size, to be transferred safely and in a timely fashion to its destination. The product meets Canada's language and accessibility guidelines, has a very simple user interface, and maintains the confidentiality of the information through end-to-end encryption. The service has been adopted as a standard for all file transfer activity between the Agency and its stakeholders.

C. Generalized Tracking System

25. A significant amount of effort is needed to track the movement of collection related items between StatCan sites. The Census has always had a tracking system to ensure that Census related material, including paper questionnaires, arrive safely at their destination.

26. As there is very little difference conceptually between tracking questionnaires, blood samples, offices supplies, etc. the Agency decided to develop a generalized system that could be used to track any type of commodity rather than a customized solution of use only for questionnaires and survey material.

27. The system is being built in-house using the Agency's standard Microsoft .net Development Framework and is in the final stages of development and testing. Once operational, the system will add another generalized solution to the growing number of such systems in use at StatCan.

V. ADHERENCE TO ENTERPRISE ARCHITECTURE

28. All the Phase I applications have been designed to be fully compliant with the principles of the StatCan Enterprise Architecture. The developers used technology choices endorsed by the Agency's Architecture Review Board, in particular:

- (i) All the applications, including the COTS products, comply with the technical architecture in place at StatCan. They are designed to run on our enterprise infrastructure services, including our enterprise server, storage and backup services.
- (ii) The home grown applications are designed using Service Oriented Architecture techniques and are developed using the Agency's standard development framework for Microsoft .net applications.
- (iii) Requirements were collected and documented using the IBM Rational Suite, in particular the Rational Unified Process, which emphasized the development of Use cases and Test cases.

- (iv) Automated testing tools were used wherever appropriate.

VI. LESSONS LEARNED

29. No project is without its problems and challenges and the first phase of the modernization of our collection processes and tools was no exception. Some of the conclusions that have been drawn so far include:

- (i) It is worth the time to review the business architecture before embarking on a re-engineering initiative. The business architecture review undertaken by the Collections area resulted in realignment of activities, streamlining of processes and a better understanding of what the final objective was. All of these activities increased the likelihood of success of the system re-engineering work.
- (ii) Strong governance and decision-making processes are required throughout the process.
- (iii) To streamline systems and control costs it is important to standardize to some extent the business processes and requirements of the subject matter areas. This is a substantial cultural shift for StatCan and we still have a long journey ahead of us to achieve effective results.
- (iv) Buying a COTS product can be very expensive especially if the requirements have not been trimmed down. Although the products we purchased met our needs and are seen to be a good investment, the actual cost of both of the COTS products we acquired significantly surpassed the original budget set aside for these acquisitions.
- (v) A minimum of nine months should be allocated to the procurement process for a COTS acquisition since the procurement rules in place within the Government of Canada make tendering processes long and arduous. Extreme care must be taken to ensure that the process delivers a result that meets the needs.
- (vi) Using standard tool sets and development frameworks saves money and decreases the risk of failure, however, the use of these tools must be governed closely and projects must have access to at least some resources with knowledge of these tools and practices. It is also important to properly account to the learning curve for staff with little or no experience with the standard tools and methods.

VII. CONCLUSION

30. Over the past two years, a significant amount of work has gone into the reengineering of our collections operations, systems and tools. The road has often been bumpy, with many challenges to overcome. Over the next few years the systems developed during the first phase of the re-engineering project will move into full production mode and the early indicators suggest positive outcomes. By the time the dust settles at the end of 2009, we fully expect that the new products will be up and running and meeting the requirements they were designed to address. We will have learned a lot about how to apply sound architectural principles to our systems projects and will have a larger pool of resources trained in these techniques.

31. There is still a lot of work to be done to fully realize the potential of these systems and of the standard development methods, tools and frameworks. However with each success comes a deeper understanding of how to maximize the benefit of these approaches and we are confident that the initiatives will both improve applications available to StatCan employees and deliver savings through increased automation and reduced duplication of processes and systems.