Audit Report

Audit of Research Data Centres Program

September 30, 2010

Project Number: 80590-61

Canada

Table of Contents

Executive Summary

The Research Data Centres (RDC) are part of an initiative by Statistics Canada, the Social Sciences and Humanities Research Council (SSHRC) and university consortia to help strengthen Canada's social research capacity and to support the policy research community. RDCs provide researchers with access, in a secure setting, to microdata from population and household surveys. The Centres are governed by a board of directors composed of representatives from Statistics Canada, SSHRC, Canadian Institute for Health Research and the academic directors of each centre. The centres are staffed by Statistics Canada employees.

The objectives of the audit of RDCs are to provide the Chief Statistician and the Departmental Audit Committee with assurance that activities supporting the governance framework are adequate and effective to ensure that services are delivered to researchers; and that access to data follows Statistic Canada policies and procedures. The audit was conducted by Internal Audit Services in accordance with the Government of Canada's *Policy on Internal Audit.*

Key Findings

Activities supporting the RDC network governance are adequate and effective in ensuring continued delivery of services. The corporate management structure is effective, results are measured and reported on; however, performance is not measured against program operational targets and results are not included in the Departmental Performance Report. Once research proposals are approved by Statistics Canada, the scope of the initiative is managed by stakeholders. Since inception, there have been challenges in maintaining program costs within granted funding limits as the number of centres rose from 9 to 24. As a result, increasing deficits have been incurred within the RDC Program. Through a newly increased annual budget, Statistics Canada has stabilized the costs for this program and established a costing structure to ensure RDCs maintain operational costs within the scope of their budgets.

The RDC program and RDC sites do not have formal risk management practices and risks associated to the continuity of the operations of RDCs have not been assessed as part of overall risk management practices for the program.

Researchers conduct their work in restricted physical and technological infrastructures that are secure, which reduces the risk of confidentiality breaches. Information handled in RDCs is subject to governmental confidentiality and security requirements; however, RDCs have not been inspected periodically for compliance to security requirements following initial inspections. After conducting a number of security checks on site, some deviations from security requirements were observed. Security training is provided to new analysts but there was no evidence of continued awareness programs. Finally the disclosure process in place for the program is applied rigorously and consistently, and no unwanted disclosures were observed.

Conclusion

Overall, the activities supporting the Governance Framework surrounding the Research Data Centres Program are adequate and effective, ensuring that services are delivered to researchers. Nevertheless, we noted opportunities to improve the management control framework as it relates to reporting results and assessing risks, as well as opportunities to ensure that departmental security requirements are met.

Introduction

Background

In the late 1990s Statistics Canada and the Social Sciences and Humanities Research Council (SSHRC) commissioned a task force to examine the state of the quantitative social sciences in Canada. Concern was expressed about the future of Canada's capacity to fruitfully exploit the rich sources of quantitative data on households and individuals to inform public policy and public debate. The joint task force published its findings in 1998 in a report titled "The Final Report of the Joint Working group of the Social Sciences and Humanities Research Council and Statistics Canada on the Advancement of Research using Social Statistics" (Statistics Canada 1998). The report outlines a series of recommendations aimed at building social science research capacity in Canada, improving access to Statistics Canada data in order to support research activity and communicating the research findings.

These recommendations led to the establishment of a Canadian Research Data Centres Network (CRDCN) and the RDC Program within Statistics Canada as the organisational unit that represents the agency within the Network.   The first 9 Research Data Centres started offering access to data in secure university-based laboratories across Canada in 2001. A second CFI award (2006), and the joint SSHRC/CIHR 2005-2010 operating grant has supported the Network's rapid expansion, and allowed it to make significant progress. Data access was improved by giving researchers across the country access within Universities to detailed microdata, initially to Statistics Canada longitudinal surveys, and now to a broader range of data sets.

RDCs provide researchers with access, in a secure setting, to microdata from population and household surveys. Statistics Canada employees oversee operations carried out in the centres. They operate under the provisions of the Statistics Act in accordance with all the confidentiality rules and are accessible only by researchers with approved projects who have been sworn in under the Statistics Act as 'deemed employees.' RDCs are located throughout the country, so researchers do not need to travel to Ottawa to access Statistics Canada microdata.

The research data centres provide opportunities to: generate a wide perspective on Canada's social landscape; provide social science research facilities across the country in both larger and smaller population centres; expand the collaboration between Statistics Canada, SSHRC, CIHR, CFI, universities and academic researchers; build on the Data Liberation Initiative and train a new generation of Canadian quantitative social scientists.

The network has grown rapidly since its creation. The number of sites has risen from 9 sites in 2001 to 24 sites in 2010, and the number of projects initiated and completed has risen dramatically. Over 2,600 researchers from a multitude of disciplines and institutions have worked in a growing number of RDCs on over 1500 projects, using an expanding range of micro-data files to examine many health and socioeconomic issues. Statistics

Canada is responsible for the protection of the data, confidentiality vetting and researcher support in centres. There are 53 Statistics Canada analysts and statistical assistants in the Centres, and 5 regional managers dedicated to this program across the country. The program is supported by a Head Office Operations Unit with 3 employees.

Authority

The audit was part of the Multi-Year Risk-Based Audit Plan 2008/09-2010/11 and was approved on March 19, 2008 by the Departmental Audit Committee.

Audit Objectives

The objectives of the audit of RDC are to provide the Chief Statistician and the Departmental Audit Committee with assurance that:

- activities supporting the Governance Framework are adequate and effective to ensure that services are delivered to researchers;

- access to data follows Statistic Canada policies and procedures.

Scope and Approach

The scope of the audit was to assess the effectiveness and adequacy of the activities supporting the current governance structure, as well as the compliance to key Statistics Canada policies and procedures relating to security and confidentiality. Relevant policies are: the Security Practices Manual which relates to the departmental security policy, the IT Security Policy, the Record Linkage Policy, the Security of Sensitive Statistical Information Policy, the Discretionary Disclosure Policy and the Statistics Act. The audit was conducted in conformity with the Treasury Board and the Institute of Internal Auditors standards.

The approach consisted of assessing the processes and procedures of the governance framework in place to control access to data. This was achieved through interviews with key departmental staff and external stakeholders managing the university side of the partnership program, detailed testing of processes and procedures and review of relevant documentation.

The RDC program has 24 centres located across Canada. The examination phase included physical inspections of five RDC sites: the Carleton, Ottawa, Outaouais, Local (COOL) Centre at the University of Ottawa, the Quebec Inter-University Centre for Social Statistics (QICSS) at the University of Montreal, the University of Toronto RDC, the University of British Columbia RDC and the Simon Fraser University RDC in British Columbia.

The examination phase for this audit was conducted from February to May, 2010.

Findings, Recommendations and Management Responses

Governance and Strategic Direction

> Overall, activities supporting the RDC network governance are adequate and effective in ensuring continued delivery of services. The corporate management structure is effective, and results are measured and reported on. Nevertheless, performance is not measured against operational program targets and results are not included in the Departmental Performance Report. Once research proposals are approved by Statistics Canada, the scope of the initiative is managed by stakeholders. Since inception there have been challenges in maintaining program costs within granted funding limits as the number of centres rose from 9 to 24. As a result, increasing deficits have been incurred within the RDC Program. Through a newly increased annual budget, Statistics Canada has stabilized the costs for this Program and established a costing structure to ensure RDCs maintain operational costs within the scope of their budgets

An established governance mechanism providing adequate strategic direction would include a clear committee structure to ensure the effectiveness of relationships and escalation of management issues. The audit found that a clear committee structure is established and is effective. The Canadian Research Data Centres Network (CRDCN) is a partnership consisting of participating universities, Statistics Canada (through the RDC Program funding) and two major granting councils: the Social Sciences and Humanities Research Council (SSHRC) and the Canadian Institutes for Health Research (CIHR).

The CRDCN and Statistics Canada have distinct reporting structures and both of them work in a joint partnership (see Appendix B). The CRDCN committee is the main governing body for the Network. It negotiates grants with the major councils and other funding bodies and sets the policies that determine the membership of the Network, the distribution of grant allocations, the strategic directions for the Network while respecting Statistics Canada criteria for confidentiality, and the dissemination of results of research conducted in the centres and the promotion of the Network nationally and internationally. The CRDCN meets twice a year and is composed of representatives of Statistics Canada, Universities, and partners (SSHRC, CIHR). During these meetings, items such as the funding allocation formula, new IT technology initiatives, training for students, RDC conferences, and member voting rights are discussed. The allocation formula is used by the CRDCN to determine grant allocation to RDCs, based on RDC activity workload, measured by the number of contracted research projects underway or suspended, and status on outputs.

In order to increase the efficiency of the decision making process at meetings, the CRDCN went through a re-organisation and implemented a sub-committee structure. The following sub-committees have been created: the Executive Committee, Implementation of CFI Award, Coordinate Information Gathering and Dissemination, Thresholds and Measures for Allocation. The CRDCN Executive Committee was created

to address strategic decisions involving the university network. This group meets twice a year prior to the CRDCN Committee bi-annual meetings. Representatives of Statistics Canada attend these meetings. The CRDCN sub-committees were created as a supporting role to the CRDCN Committee. These sub-committees provide subject-matter expertise on research related questions for consideration and decision making at the CRDCN Committee meetings. Statistics Canada representatives often attend these meetings and provide workload information used to set the funding allocation formula.

The RDC Program structure within Statistics Canada consists of the lead program manager and five regional RDC managers. Meetings are held on a weekly basis to discuss day-to-day operations of the RDC Program and ensure that processes and procedures are applied consistently across the program. Members also meet as a working group to discuss and resolve operational issues that may arise. The operations of the centres are managed by Statistics Canada employees in partnership with academic directors from the universities. Universities provide IT support. One representative of the research community in each university holds the function of Academic Director. The Statistics Canada analysts report to the regional RDC managers, who are located in Statistics Canada regional offices. Regional RDC managers are responsible for the management of several RDCs and report to the Assistant Director, Microdata Access Division. Because managers are not on site, management of the centres is done through e-mails, phone conversations and twice-yearly visits to centres.

To access the microdata housed in the Research Data Centres (RDCs), researchers submit a project proposal to the Social Sciences and Humanities Research Council (SSHRC) for peer review and internal review by Statistics Canada. SSHRC invites applications from individual researchers or from research teams led by a principal applicant. The principal applicant is responsible for submitting application forms on behalf of the team. Each proposal is evaluated by two academic peers and a Statistics Canada analyst. The SSHRC facilitates the peer review process and the head office operations unit facilitates the internal Statistics Canada review. The proposals are assessed based on: scientific merit and viability of the proposed research; relevance of the methods to be applied; demonstrated need for access to detailed microdata; and expertise and ability of the researchers to carry out the proposed research as illustrated in the resumes and list of contributions. Statistics Canada is not involved in the prioritization of projects undertaken in RDCs; however, it is the responsibility of each RDC to ensure that resources and funding is available prior to taking on a new research project.

Management Control Framework

An adequate management control framework is required to ensure effective planning, organizing, controlling, directing, communicating, as well as compliance with the Treasury Board Management Accountability Framework and the TBS Risk Management Policy. Clear program objectives should support strategic direction, operational plans and priorities, and should provide clear direction on how resources should be allocated to achieve these plans. Planning for the program involves the production of annual operational plans, budgets, staffing plans, and an annual RDC activity report, which are

produced on a timely basis.

To ensure appropriate delegation of authorities, roles and responsibilities need to be documented and communicated. It was found that roles and responsibilities were clearly documented and understood. The organisational chart of the overall structure, including the relationships between the CRDCN and Statistics Canada, are up to date and indicate the linkages between the different stakeholders.

A communication process should be in place to ensure consistency of program activities and compliance to policies and regulations. The CRDCN, Statistics Canada and each RDC have their own websites, and communication for the program is done through the use of web technology. Communication between the RDC Regional Managers and the Statistics Canada analysts is informal. Items discussed are mainly related to RDC management and statistical operational processes. Operational issues not resolved at the regional level are escalated to the RDC Program Manager.

Values and ethics are promoted and communicated on an on-going basis within the researcher community through formal documentation, training, and policy frameworks. Microdata research contracts signed by researchers include several clauses regarding values and ethics. Also, researchers receive an orientation session where values and ethics are introduced and presented. Researchers acknowledge the "Researcher Guide", which includes a section on values and ethics. Values and ethics awareness is also reinforced through the Oath required under the Statistic Act.

An effective performance management system should be in place to measure and report on performance. Relevant performance targets should be identified, and information on results should be gathered and used to make departmental decisions. Program outcomes should also be reported as part of the Departmental Performance Report (DPR). The RDC Program Manager, who reports to the Director General of Census Subject Matter, Social and Demographic Statistics, is accountable for reporting on performance. Different reviews of the RDC initiative have been conducted, such as: the mid-term review produced by SSHRC and CIHR, the Quadrennial Program Review of the program produced by Statistics Canada for 2000-2008, the Program Manager Report, and Satisfaction Survey. Although results are measured and reported on for internal administrative purposes, performance is not measured against comprehensive program operational objectives, nor is it reported in Statistics Canada's DPR.

Investment Management/Funding

The overall initiative is funded primarily by Statistics Canada (through the Research Data Centre Program), the SSHRC, CIHR and the universities. SSHRC and CIHR provide grants to the CRDCN, which then re-distributes funds to the RDCs. An allocation formula is used as a basis to distribute funds, taking into consideration factors determining RDC activity levels such as the number of researchers, branches, and projects. Statistics Canada provides in-kind contributions, by not only providing access to data, but also by assuming the operational costs to manage the program. In the event

that costs associated with administrating additional research projects exceed the limits set by the program, RDCs enter into an agreement with Statistics Canada for additional resources on a cost-recovery basis.

Every RDC has its own budget, which is managed by the Academic Director. The Academic Director is a university researcher and does not have authority over the Statistics Canada analysts. A large portion of the RDC budget consists of the cost associated with Statistics Canada analysts and for daily operations. Universities provide IT services and support to RDCs. Enhancements or development of new initiatives in centres are undertaken only if additional sources of funding are received.

The RDC program budget is intended to fund Statistics Canada's head office support for the RDC Network, including: the maintenance of the Management Information System on research activity conducted by deemed employees for Statistics Canada, the administration of contracts, maintenance of the RDC Web site, provision of the data to the RDCs, preparation and maintenance of the documentation required by researchers and RDC project staff, LAN support for head office, and the provision of methodological support to the RDCs.

The demand for Statistics Canada data access has grown over the last ten years, and increasing annual deficits have been incurred within the Program as the number of centres grew from 9 to 24. Over the past three years, deficits incurred by the RDC program were partially covered by funds available in other field programs.

RDC Program Expenditures		Actuals			Going forward Budget/Target '(2)	
		2007-2008	2008-2009	2009-2010	2010-2011	2011-2012
Base (PE 1884)						
ITSD	(3)	138,689	112,592	182,320	-	-
Salary		364,181	697,726	595,040	711,640	711,640
Non-Salary		92,238	102,158	100,761	90,555	90,555
Prog. Admin Costs		595,108	912,476	878,121	802,195	802,195
Recoverable Expenses (PE 6587)						
ITSD		2,374	33,053	66,102		
Salary		1,920,355	1,943,262	2,126,833		
Non-Salary		458,614	330,798	220,105		
Total Recoverable Costs	(1)	2,381,343	2,307,113	2,413,040		
Total Base & Recoverable	(1)	2,976,451	3,219,589	3,291,161		
Minus: costs recovered		- 2,381,343	- 2,307,113	- 2,413,040		
Net Program Costs		595,108	912,476	878,121		
Budget/Target		527,000	527,000	527,000		
Surplus (Deficit)		**- 68,108**	**- 385,476**	**- 351,121**		
Notes:						
(1) Source: Financial system aggregate data.						
(2) Going forward, base budget for the Program has been increased by an additional $400,000.						
(3) Going forward, ITSD costs will be covered under the recoverable revenues.						

In April 2009, the RDC program was transferred to the Microdata Access Division. For the fiscal year 2009/10, the deficit incurred for the research data centre initiative reached $351,121. In order to remedy this situation, a long-term proposal has been submitted to ensure on-going funding to cover head office costs. Going-forward, the annual funding for the program has been set at $800,000 annually.

From this time forward, RDCs are expected to maintain operational costs within the scope of their budgets.

Recommendation No. 1

It is recommended that the Assistant Chief Statistician (ACS) of Social, Health, and Labour Statistics ensure that performance indicators are measured against operational

targets set for the program and that program results are included in the DPR for future target references.

Management Response

Agreed. Performance indicator measures and results are now an integral part of the Performance Program Review with detail in the program logic model of outcomes and how they are measured. Microdata and Access Divison will thus include them in its upcoming Performance Program Review. Performance of the RDC program has been positively reviewed through an external international panel and through client satisfaction feedback.

Deliverables:
- Performance indicators and measures identified in the Performance Program Review and Departmental Performance Report – by March 31, 2011; and
- Program results identified in the Performance Program Review and Report on Plans and Priorities – by March 31, 2011.

Risk Management

> The RDC program and RDC sites do not have formal risk management practices, and risks associated to the continuity of the operations of RDCs have not been assessed as part of overall risk management practices for the program.

In a well-controlled program environment, management should have a solid and up-to-date understanding of the internal and external factors that may expose their strategic and operational objectives to risk. Resources and strategic risks should be monitored proactively to assist in decision-making and Statistics Canada's Long Term Planning process. Formal risk management practices enable program managers to identify, assess, monitor and report on risks that may result in threat or opportunity. Although some risks for the RDC program are discussed informally in meetings, there is no evidence of a risk management framework in place or a link to Statistics Canada's Risk Management Model. In the absence of a formal risk management framework in place, management's ability to identify and influence risk throughout the program lifecycle is weakened.

The policy on government Security states: "Continuity of government operations and services is maintained in the presence of security incidents, disruptions or emergencies". The departmental Business Continuity Plan policy also states that "it is the responsibility of the RDC Manager to develop and implement a Business Continuity Plan for each of the Research Data Centres and integrates the plan with those developed by the institution or university in which the centre is located". Although the RDC program is not a mission-critical project, risks associated to the continuity of the operations of RDCs should be identified and assessed as part of the overall sound risk management practices for the program.

Recommendation No. 2

The Assistant Chief Statistician (ACS) of Social, Health and Labour Statistics should ensure that the RDC program be included in Statistics Canada's Risk Management Model and that risks associated to the continuity of the operations of RDCs are identified and assessed as part of the overall risk management practices for the program.

Management Response

Agreed. The RDC program has identified program risks for Statistics Canada's Risk Management document in July 2010. Risks associated with RDC operations will be an integral part of the Performance Program Review, including completion of the Risk Register. The program manager will continue work on a business continuity plan for the program.

Deliverables:
* Inclusion in Statistics Canada's Risk Management document, a Business Continuity Plan, and a risk assessment in the Performance Program Review. – by March 31, 2011.

Access to data

The audit found that researchers conduct their work in restricted physical and technological infrastructures that are secure, which reduces the risk of confidentiality breaches. Information handled in RDCs is subject to governmental confidentiality and security requirements. Nevertheless, RDCs have not been inspected periodically for compliance to security requirements. After conducting a number of security checks on site, some deviations from security requirements were observed. Security training is provided to new analysts, but there was no evidence of continued awareness programs. Finally, the disclosure process in place for the program is applied consistently, rigorously and consistently, and no unwanted disclosures were observed.

It is expected that controls over access to Statistics Canada data in RDCs are in compliance with Government and Statistics Canada policies and procedures on security and confidentiality. Applications are submitted through the Social Sciences and Humanities Research Council, and assessed within Statistics Canada through the SBSD, SM Area, DACS, and the RDC Program divisions. Once access has been granted, Universities provide RDC space and Statistics Canada ensures the site complies with departmental physical security and confidentiality policies. The audit found that researchers conduct their work in restricted physical and technological infrastructures that are secure, which reduces the risk of confidentiality breaches, and no unwanted disclosures were observed. Information handled in RDCs is subject to governmental confidentiality and security requirements.

Infrastructure

Prior to approving the provision of access to Statistics Canada data to RDCs, physical and technological infrastructure of RDCs must meet the security requirements of the department. The universities provide physical sites as well as IT support, Statistics Canada ensures the sites meet the departmental security requirements which are based on the TBS Policy on Government Security. To review the security of the infrastructure, a security inspection checklist has been developed. Inspections include physical security infrastructure items such as site access, locks, and keys; and electronic security infrastructure items such as server access, and passwords. This security inspection checklist is used to ensure departmental security requirements are applied when a new RDC/branch is opened. In summer 2009, the security inspection checklist was reviewed by the Data Access and Control Services Division to ensure that the requirements of the new Policy on Government Security continued to be met. A review of a sample of completed security inspection checklists revealed several forms that were only partially completed; consequently, there was insufficient evidence that complete security inspections were conducted.

The TBS Policy on Government Security referring to the Directive on Departmental Security Management recommends regular security inspections of working sites in order to identify potential security risks. In reality, RDCs have not been re-inspected since their opening, in some cases for the last ten years. After conducting a number of security checks on site, some deviations from security requirements were observed. As an example, the Statistics Canada Security Practices Manual states that "to protect the integrity and data availability of the Agency's information assets, all information should be on servers and network drives that receive daily automatic backup". The audit found that, in some instances, backup storage procedures were not consistently applied as described in the Manual.

The TBS Policy on Government Security states that "A departmental security awareness program covering all aspects of departmental and government security must be established, managed, delivered and maintained to ensure that individuals are informed and regularly reminded of security issues and concerns and of their security responsibilities." Departmental requirement for security training has been enforced through the KLICK security course offered to RDC analysts; however, there was no evidence of mechanisms in place to ensure continued awareness among Statistics Canada analysts since the opening of the centres.

The demand for greater access to data continues to increase. In response, Statistics Canada's program management is examining other means of data access such as remote access and synthetic files as a means to provide improved data access.

Confidentiality

One of Statistics Canada's key values is the preservation of the confidentiality of its data and the respondents. Data handled by RDCs must also meet departmental confidentiality and security requirements. Data requested by researchers is extracted in Headquarters, encrypted on disk and transmitted to RDCs by courier, using confidentiality rules for protected B information. Data is decrypted by the Statistics Canada analyst and installed on the RDC server, which is not connected to a network. Researchers obtain access to Statistics Canada data samples through individual user accounts for each of their research projects. Disclosure guidelines are implemented within the department to avoid publication of protected information to the public. To ensure Statistics Canada information is protected within RDCs, a disclosure analysis process conducted by the Statistics Canada analysts is performed prior to the publishing of results. Researchers submit their information to the analysts, who then review and remove information that could lead to a breach of confidentiality. The audit found that the disclosure process in place for the program is applied rigorously and consistently.

Recommendation No. 3

The Assistant Chief Statistician (ACS) of Social, Health and Labour Statistics should work with Corporate Services to ensure that regular inspections are instituted to ensure infrastructure surrounding the access to data continue to meet requirements of the departmental security policy

Management Response

Agreed. In addition to completeinitial site inspections, the RDC program will conduct regular inspections of all RDCs. The Microdata Access Division has in place a number of procedures to safeguard the physical and electronic security of the centres and their data.

Deliverables:
• Inspector from corporate services will report on the security of RDCs by submitting the completed security checklist - The schedule will start with the first re-inspections of the RDCs in 2011-2012 Fiscal Year. All current RDCs will be re-inspected by 2015-2016..

Appendices

Appendix A: Audit Objectives and their Criteria

Objective 1: The activities supporting the Governance Framework are adequate and effective to ensure that services are delivered to researchers.

Governance and Strategic Direction: Review the adequacy of the management control framework (MCF) of the program including processes and practices related to planning, organizing, controlling, directing and communicating.
- Relationship exists between the strategic plan and the objectives of the programs.
- Roles/responsibilities are defined and communicated through the program.
- A clear organizational structure is established, documented and reporting relationships are effective.
- The values and ethics are promoted among stakeholders.
- There is a communication process in place to communicate program activities.
- There is an effective performance management and accountability framework in place to measure and report on performance.

Risk Management:
- A risk management mechanism should exist to identify, assess, monitor and report on risks.

Investment Management/Funding
- The current funding planning/budgeting processes are effective. Expected costs are defined and assessed based on expected benefits.

Objective 2: Access to data follows Government and Statistics Canada policies and procedures.

Infrastructure:
- Physical and technological infrastructures are secure.

Information Security and Confidentiality:
- Electronic data access is controlled efficiently.
- Confidentiality is maintained, data are adequately protected.
- Unwanted disclosure is avoided.

Business Continuity Planning:
- Business Continuity should be defined to assure continuity of the operations.

Appendix B: Overview of the RDC Network

RDC NETWORK
Universities, Centres & Branches
15 Data Centres
9 Branches
26 Partners
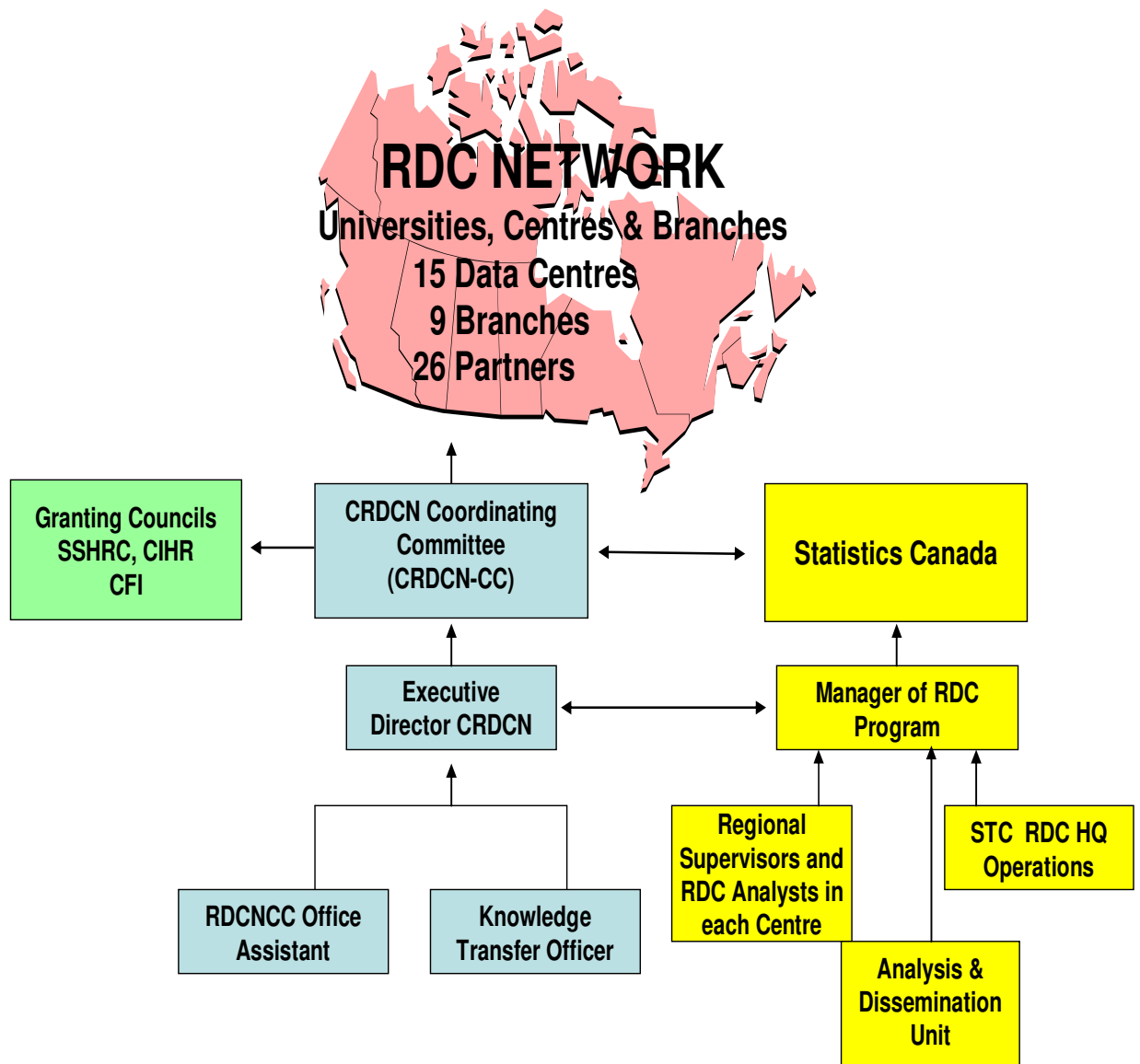
Reference: Quadrennial Program Review: 2000 – 2008, page 19.