Report to Canadian Geoscience Data Model Working Group

Survey of Geoscience Agency Systems and Data Models

Version 1.0

May 26, 2000

This work was done under contract to the Geological Survey of Canada – Calgary. This report is Deliverable 3 of that contract. The report was delivered to the Canadian Geoscience Data Model Working Group.

Executive Summary

This report describes a survey that examined the data holdings and databases of 18 territorial, provincial, and federal geoscience agencies. The survey's goal was to assess:

- The extent of digital geoscience data holdings
- The structure and quality of existing geoscience databases
- The data models and standards used for each type of data.

The survey is part of a larger objective to improve access to geoscience data across Canada.

The survey raised a number of issues that would need to be addressed before more ambitious developments should proceed. Key issues are as follows.

Issue	**Ramifications for improved data access**
No single data model is being used, or is felt to suffice.	A pluralistic approach to data modeling is needed for the near future.
No single approach for metadata or data management is being used.	Many challenges exist for searching data holdings and exchanging data.
A lack of funding makes discretionary spending difficult. Many agencies would like long-term funding in return for participating in federal initiatives.	Some agencies are unable to participate in GSC initiatives to improve or standardize data access because they do not have the funding, or because they are unwilling to commit their own funding without a similar show of commitment by GSC.
Cost-recovery policies exist for geomatics data from Geomatics Canada, as do local user fees for data.	Any initiative to improve access to data will be made more complex by the patchwork of policies governing cost recovery. The ultimate costs recovered are arguably offset by the cost of cost recovery.
Multilingual requirements are unclear.	The patchwork of policies and practices create a challenge, particularly if federal language laws are strictly interpreted for local scientific data. For some aboriginal languages, no translation is possible.
Need for client input to the modeling process.	There is a diverse clientele for geoscience products. Good representation from industry is required if modeling standards are to be established.
In the past, GSC has started, then dropped promising developments.	Many agencies feel unwilling to commit their own resources to developments that can be dropped without consultation.
Leaders in web development lose out on "catch up" funding.	There is a sense that those who spend local monies for technical advancement are penalized when later funding

Issue	Ramifications for improved data access
	is expended to help others catch up. At minimum, the lessons learned by the early innovators should be made available to others as they undertake similar developments.
Giving credit where credit is due, with some balance.	If territorial and provincial agencies accept even minimal GSC participation on a project, they also must accept three GSC logos on the results. The three logos eclipse the single local logo, and make requests for local funding more difficult to justify.

The recommended first step is to develop a meta-information model including a subject classification and other information components that would support a common web portal to distributed Canadian geoscience data. This portal should allow the user to search by geographic area or by subject. Using the search criteria, the portal would direct the user to the appropriate agency's (or agencies') web site. For some agencies, even organizing a subject classification or simplified form of metadata will be a significant undertaking. However, all agencies wished to participate in this process.

The second step would be to implement the portal to test and demonstrate the model and the technology available for web-based information access.

Table of Contents

1 Introduction

This report describes and discusses a survey that examined the data holdings and databases of 18 territorial, provincial, and federal geoscience agencies. The survey's goal was to assess:

- The extent of digital geoscience data holdings
- The structure and quality of existing geoscience databases
- The data models and standards used for each type of data.

The discussion compares the systems used by the 18 agencies surveyed, and includes a description of the advantages and disadvantages of their various data models. Data, data models, and standards are described in each of six broad geoscience categories.

The information presented here was collected by a combination of email survey, telephone follow-up, and in-person visits to agencies to meet with data managers and custodians. The following organizations and individuals participated.

GSC	Provincial Surveys	Territorial Surveys
GSC-Atlantic, Dartmouth GSC-Quebec, Quebec GSC-Ottawa, Ottawa GSC-Calgary, Calgary GSC-Pacific, Vancouver GSC-Pacific, Sydney	BC Geological Survey, Vancouver Alberta Geological Survey, Edmonton Saskatchewan Energy and Mines, Regina Manitoba Geological Survey, Winnipeg Ontario Geological Survey, Sudbury Géologie Québec, Ministère des Ressources Naturelles NS Mines & Energy, Halifax New Brunswick Geological Survey Geological Survey of NF and LB, St. John's	YK INAC, Whitehorse NWT INAC, Yellowknife GNWT RWED Minerals, Oil and Gas Division Nunavut GSC, Iqualuit

1.1 Purpose & Audience

This report summarizes the highly detailed information provided in written surveys that were submitted by agencies using either a long form or a shortened form. Issues and concerns expressed by survey respondents are highlighted as part of the summary. Information provided by the 18 agencies is contrasted in a series of tables.

The full surveys as submitted by each agency will be made available for download and review on an ftp site. Summary tables were compiled from the surveys and are provided in Appendix A.

1.2 Report Structure

Section 2 of this report provides a high-level overview of the survey results and how they are presented in this report. In Section 3, the results are discussed in more detail using six broad geoscience categories.[1] Section 4 presents common issues and concerns raised throughout the survey. Section 5 provides a set of conclusions that characterize existing systems and approaches, define areas of potential cooperation, identify challenges and possible barriers, and establish a set of logical "next steps" toward the goal of improved access to Canadian geoscience data by Canadians.

2 Overview of Survey Results

This survey highlights the diversity of geoscience data holdings and standards across Canada, and the range of data management methods used by different agencies. Data management can be as simple as “a clean sheet of paper” (Nunavut), or as sophisticated as the essentially corporate systems of provincial Quebec (SIGÉOM) and Ontario (ERLIS). One common element across Canada is an enthusiasm for the Internet as a means of disseminating data. Every agency surveyed either has published some of their data sets to the Internet or is looking to do so in the near future.

The survey examined 18 territorial, provincial, and federal geoscience agencies. The results are presented in six (6) broad geoscience categories:

- Geological maps
- Minerals/Hydrocarbon databases
- Geochemistry/Geophysics databases
- Assessment reports and associated databases
- Borehole databases
- Biostratigraphic databases

These categories were defined by the Canadian Geoscience Knowledge Network (CGKN) and the Canadian Data Model Working Group (CGDMWG). The categories are not mutually exclusive (e.g., Hydrocarbons and Boreholes, Biostratigraphy, and

[1] These categories were defined by the Canadian Geoscience Knowledge Network (CGKN) and the Canadian Data Model Working Group (CGDMWG).

Boreholes) which has made it difficult in some cases to know where best to enter an agency's data set. Whenever possible, the designation provided by the agency is used, or the data set (or system) has been placed in the apparent dominant category with references in other categories. In addition, an "Other" category has been added to include those reported data sets or systems that did not fit into any of the six originally defined categories.

Large, integrated systems such as those found in Ontario and Quebec are a special challenge to categorize. These large systems cover a wide variety of geoscience categories and therefore do not have a dominant category. In most cases, these systems have multiple entries in accordance with the reporting mechanism of the surveyed agency.

Within each of the six broad geoscience categories, one or more tables provide a direct comparison of the different agencies' reported holdings within that category. The tables present only those data sets (databases) reported in the written surveys submitted by the survey participants, and not all agencies submitted surveys. Thus, the absence of an entry in a particular category does not necessarily mean that an agency does not collect data within that category. Likewise, an entry in a category may not represent all of the data collected by the agency in that category.

The actual submitted surveys that form the foundation of the comparison report are provided on an ftp site. The submitted surveys contain far more detail than is given in the comparison tables, and will be a good resource for anyone wanting to know more about a particular reported data set or system. The surveys are all in Excel 97 format, and are divided by agency.

The territorial agencies of the Yukon, Northwest Territories, and Nunavut have a unique relationship with the federal government agencies of DIAND and/or GSC. In all three cases, the territorial agencies work closely with their federal counterparts, even sharing office space. The close ties between territorial and federal agencies are such that a combined federal and territorial survey was submitted by each of the Yukon, NWT, and Nunavut. Rather than trying to separate systems into federal or territorial initiatives, all survey submissions are assigned to the territorial name in the summary tables presented below and on the ftp site. For details on a particular system and its federal versus territorial component, please refer to the original surveys available on the ftp site.

3 Detailed Survey Results

3.1 Geological maps

3.1.1 Extent of holdings

Provincial and territorial surveys typically are restricted to map coverage within their province or territory. Most of the provincial agencies surveyed were further limited by their agency mandate. In many cases (e.g., Ontario, Manitoba, Saskatchewan, and others), responsibilities for oil and gas have been placed in another branch so the surveyed agency has restricted its coverage of petroleum producing areas. In the case of Nova Scotia, all offshore areas and petroleum-related areas onshore were transferred to the petroleum branch and are no longer within the mandate of the Nova Scotia Natural Resources branch. As a result, the main clients for many of the provincial data sets are mining companies and others interested in mineral holdings. Map coverage within the agency focuses on mineral-rich areas.

The volume and extent of data holdings is highly variable (Appendix A, Table 1). All data volume estimates are rough approximations. The size of a holding does not necessarily indicate the maturity of the mapping system. For example, Newfoundland reports a volume of only 12 MB, but their mapping system, GeoLegend, is mature and innovative.

3.1.2 Standards

Of the agencies that report that they use a geospatial metadata standard, approximately three-quarters use the FGDC[2] standard (Appendix A, Table 1). The Alberta Geological Survey has a well-developed metadata management system based on GILS,[3] which was extended to meet spatial requirements. This method is effective, but Alberta remains open to other alternatives in the future. Newfoundland uses the ISO TC211 metadata standard, and other agencies have an interest in SDTS[4] and OpenGIS.

Most agencies have developed their own data models or structures to manage their data. With some notable exceptions (e.g., Ontario and Quebec), many of the systems are stand-alone and have limited connectivity to other systems within the same agency.

---

[2] Federal Geographic Data Committee, a US committee that has developed a metadata standard.

[3] Government Information Locator Service, see http://www.access.gpo.gov/su_docs/gils/whatgils.html.

[4] Spatial Data Transfer Standard, a file transfer standard with profiles for different data types.

The data model used as a basis for the greatest number of systems surveyed is PPDM.[5] It was reported for three times as many systems as the NADM/CordLink combination (see below). This is hardly surprising, as PPDM has been around longer than NADM (1990 vs. 1997) and has much greater depth and breadth.

Two of the agencies surveyed are using the North American Data Model (NADM v.4.3) or the CordLink model (also known as NADM v.5.2) for their geological maps. Another five agencies are either evaluating or expecting to use the CordLink model in the future. Almost all agencies express some level of interest in a common model which could lead to a "seamless" geological map of Canada. In addition, the ability to adapt the CordLink model specifically to other types of data (e.g., forestry) is seen by some agencies as an advantage in meeting their local mandates to interact with local natural resource agencies.

Despite the interest in NADM and the CordLink model, there are concerns about the lack of capability of these models to handle symbols, and their lack of capability in areas beyond geological mapping (e.g., remote sensing, paleontology, geochemistry). Many agencies indicated that they are waiting for more proof of the effectiveness of NADM and/or CordLink before they would consider migrating their systems to this model.

GSC-Atlantic has loosely used the Ocean Drilling Programme (ODP) model for their Expedition Database (ED), from which geological maps may be produced. Concern was expressed that the use of the ODP model provided questionable interchange with other systems.

NAD27 is the most common spatial reference and UTMs the most common coordinate system. The prairie provinces may capture or store the original data in torrens,[6] which can be converted to geographics (latitude and longitude) or UTM coordinates using standard transformations.

Many of the digital map data were digitized from paper maps, so positional accuracy varies widely. Primary data collected in the field is generally accurate to the limit of precision of the tools and methods used to determine location.

[5] Public Petroleum Data Model, developed by a not-for-profit organizations representing over 100 oil and gas companies, vendors, and regulatory agencies worldwide. The model is a vendor-independent standard that serves the industry but has limited applicability to paleontological and geological elements. See the March 2000 Holonics report "Evaluation of Candidate Models for GeoScientific Data" for more details on PPDM.

[6] A locational indicator based on the township range method of public survey.

3.1.3 Software

The 18 geoscience agencies surveyed use a diverse array of computer software to produce their maps (Appendix A, Table 2). The most popular suite of geographic information system (GIS) tools are those produced by ESRI. Over 80% of agencies use ArcInfo, ArcView, or a MapObjects application at some point to produce or view maps.

AutoCAD and AutoCAD Map by AutoDesk are used by more than half the agencies. These products are often in combination with FieldLog. A common scenario is for field geologists to capture their data in FieldLog and then transfer the data to AutoCAD [Map] to create their maps. These maps may then be transferred to ArcInfo for final adjustments before release.

Both New Brunswick and Quebec are dedicated shops. New Brunswick uses CARIS only, and Quebec uses Microstation only. CARIS is also used by Newfoundland and GSC-Atlantic. Ontario uses Microstation to produce paper copies of its maps. There appear to be regional ties in the type of map production software used.

Data management of completed maps varies enormously. Map data are most commonly stored in the GIS package used. Often as a map is completed, a final copy is produced on CD, and the CD serves as an archive. A common complaint is that once a map is finished, it is no longer part of a searchable database. This is not true for all agencies, however. For example, both Ontario and Quebec have mature and comprehensive systems that manage their map data within a larger relational database management system. Those agencies that do have an underlying relational database management system for their geological maps generally use Oracle, MS Access, or Ingres.

Every agency surveyed expressed enthusiasm about making their maps available via the Internet, but there is no consensus on how best to accomplish this. There is wide recognition of the "Map Place" work of the provincial BC Geological Survey. MapGuide and MapObjects are the most frequently mentioned tools for web publishing, but other tools also are used. Géologie Québec has recently web-enabled their SIGÉOM system using Microstation. GSC-Quebec used ColdFusion in conjunction with MapGuide, and other options are being discussed. GSC-ESS in Ottawa is providing internet access to geological maps using Oracle/CubeSTOR, which supports query from any spatial client that supports either the Open GIS Consortium (OGC Web mapping protocol) or the Information Interoperability Institute (III) Open GeoData Interface (OGDI). This project will also result in a preliminary design for providing access to the spatial warehouse using MapGuide.

3.2 Minerals/Hydrocarbon Databases

3.2.1 Minerals

Mineral databases are summarized in Appendix A, Table 3. Most provincial and territorial agencies have a strong focus on mineral occurrences. Clients for many of the products produced by these agencies are mining companies and others interested minerals and base metals. Some agencies (e.g., provincial agencies in Quebec and Ontario) have representatives from the mining industry on their advisory boards. As an example, the Ministère des Ressources Naturelles (MRN) in Quebec has a special advisory board for their SIGÉOM system that includes both ministry and industry members. This board meets at least three times a year to review the direction of development for SIGÉOM.

For metadata, the FGDC standard or a derivation thereof was used by three agencies. Ontario uses GOITS[7] 72.0, which is a province-wide standard for metadata based on FGDC.

For data management, the three minerals database systems that were most commonly mentioned by the surveyed agencies as exemplary are NORMIN.DB (NWT), MINFILE (BC), and MODS (NF). Although these are not the only well-structured minerals database management systems among those surveyed, they do have a high profile across Canada.

3.2.2 Hydrocarbons

This survey provides a restricted view of hydrocarbon databases within Canada (Appendix A, Table 4). As previously noted, responsibilities for oil and gas are beyond the mandate of many of the agencies surveyed. Separate agencies have been created in many provinces to handle hydrocarbons and most of these agencies were not included in the survey list by CGKN or CGDMWG. Many respondents noted that when the new oil and gas agencies were created, the bulk of the budget for geoscience left with them.

For data models, of the eight agencies reporting hydrocarbon databases, half (four) based their systems on the PPDM. This data model is well specified and used extensively by the oil and gas industry. Agencies using PPDM considered the new spatial extensions in PPDM v.3.5 beta to be an encouraging trend for this well-established data model.

[7] Government of Ontario Information Technology Standard

For metadata, only two respondents reportedly used it; one plans to use GILS (Alberta), the other used FGDC (GSC-Atlantic). As there were only two respondents for metadata, it is difficult to suggest any metadata trends for hydrocarbon databases.

For data management, many of the systems reported in this category are quite mature. For example, the BASIN system (GSC-Atlantic) covers offshore eastern Canada and the Arctic and contains approximately 8.9 million records. The WELLS system (GSC-Calgary), MOGWIS (Manitoba), and the Athabasca Oil Sands Database (Alberta) also are quite extensive.

3.3 Geochemistry/Geophysics Databases (raster and vector)

3.3.1 Geochemistry

No cohesive standard ties together the geochemistry holdings of the survey respondents.

For data model, no standard was mentioned for any of the geochemistry databases reported in the survey (Appendix A, Table 5). However, a geochemistry data model is currently under development by the GSC-Ottawa, called the MultiDivisional Database Model.

For metadata, the FGDC standard is used by two agencies (NWT and MB), and Ontario uses a derivation of FGDC in the form of GOITS 72.0.

For data format, Ontario, Quebec, and Newfoundland all maintain their own standards.

For data management software, Oracle, Ingres, and Microsoft Excel are popular options. NS uses ArcInfo to maintain their geochemical data.

3.3.2 Geophysics

The Geophysics category has a greater variety of data set types (Appendix A, Table 6) and file sizes (10 MB to 4 terabytes) than almost any other. As with hydrocarbons, it is likely that this category is under-represented in the survey due to the separation of agencies that deal with mineral occurrences from those that deal with oil and gas. This would be particularly true for seismic data, which often are collected in association with oil and gas data. Four national databases are reported in this category: the National Wave Form Archive, the National Earthquake Database, the National Aeromagnetic Database, and the National Gravity Database. These four systems would be obvious starting points for developing a Canada-wide system for geophysical data.

For data model, PPDM was the only reported standard in use.

For metadata, the FGDC standard and the Ontario derivation were the only reported metadata standards in use

For data format, the Geophysics category is unique in the number of reported international data format standards to which this type of data must adhere (e.g., SEG-Y, SEED, LAS, SINEX, RINEX). In discussions with agencies that deal with geophysical data, all agreed that it is critical to adhere to these international data format standards.

For data management, Oracle, Ingres, and ArcInfo are in use by two or more organizations.

3.4 Assessment Reports and Associated Databases

Many of the assessment reports are managed only via paper-based systems or via an on-line index to paper-based reports (Appendix A, Table 7). However, systems such as BC's ARIS do include the reports themselves and in ARIS's case, reports also are available on the Web.

For data management, Manitoba, Ontario, and New Brunswick manage their reports and/or indexes using Oracle, Ingres, or Access. In some cases (e.g., Ontario), the assessment reports are part of a much larger system.

3.5 Borehole Databases

Borehole data can occur in three of the six categories used by this report: "Borehole," "Minerals/Hydrocarbon," and "Biostratigraphic." The entries in the summary table (Appendix A, Table 8) are a mix of boreholes related to oil and gas wells, and those related to coal and diamond drill holes, among others. In some cases, it was difficult to classify an entry as a borehole database versus a hydrocarbon or biostratigraphic database. Whenever possible, the original designation submitted by the surveyed agency was used. In such cases as the Athabasca Oil Sands and the Manitoba Stratigraphic Database, the main entry is in the hydrocarbon category (Section 3.2.2) or biostratigraphy category (Section 3.6), and a cross-reference is recorded in the Borehole table.

For data model and data format, no use of standards was reported for any of the borehole databases. The Athabasca Oil Sands and the Manitoba Stratigraphic Database (cross-referenced from other categories) use the PPDM data model.

For metadata, Alberta's Index of Groundwater Wells will use GILS in the near future in keeping with the overall GILS metadata standard used by the Alberta Geological Survey. Ontario uses GOITS, which is a derivation of FGDC.

For data management, four databases use MS Access, two use Ingres, two use Oracle, and others use ArcInfo or Excel.

3.6 Biostratigraphic Databases

One stratigraphic and four paleontology entries were reported for this category (Appendix 1, Table 9) although many of the other systems in other categories (especially hydrocarbons and boreholes) likely contain stratigraphic and/or paleontological information.

For data model, both PPDM and Georecords were reported. It is worth noting that the Stratigraphy Work Group within PPDM has made some significant changes reflected in 3.5 beta version of the model. The Georecords data model was developed over 15 years ago at the University of Cambridge in England. This model is used for a variety of geoscience data, including paleontology. GSC-Pacific held their paleontology data in this model for many years and have recently upgraded to a new system that uses a structure based on the original Georecords model.

For metadata, the agencies that reported using a standard for biostratigraphic databases use the FGDC metadata standard.

For data management, Oracle is the software platform most commonly used.

3.7 Other Databases

This category contains data sets or systems that were reported in the submitted surveys but are not readily assigned to one of the six categories defined by the Canadian Geoscience Knowledge Network (CGKN) and the Canadian Data Model Working Group (CGDMWG). Several bibliographic systems are considered to be "other". Also included as "other" are a curation system (SAMPLE), abandoned mine openings, peatlands, and geochronology (Appendix A, Table 10).

The number of items in this category indicates the breadth of geoscience data not covered in the six designated categories. The entries in the summary table are only a small indication of the number of systems that are geoscientific in nature but outside the realm of the listed categories. Because the "other" systems are so diverse, few conclusions can be drawn from the use of standards by these systems.

For data model, PPDM was the only reported standard.

For metadata, FGDC (which includes the Ontario version of FGDC) was the only reported standard.

4 Common Issues and Concerns

A number of common issues and concerns were raised in the interviews (both in person and by telephone) conducted with members from the 18 territorial, provincial, and federal agencies. Many of these issues and concerns were common to all agencies, while some were most important to the provincial or territorial agencies. Many of the concerns were related to the intended use of the common data model. There is a lot of sensitivity about the perceived need to spend significant effort on a model and an associated ***"national system***" that may force participating agencies to modify their internal operational systems or change business practices related to data management.

None of the issues addressed below apply equally to all agencies. However, to preserve confidentiality, only general comments about the proportion or type (territorial, provincial, or federal) of agencies that raised each issue is included.

4.1 Common to all agencies

The issues below are presented in order of how widely and vocally they were expressed.

4.1.1 Lack of financial resources

A lack of financial resources was the most common issue of the agencies surveyed. This issue affects almost every other issue and concern listed below. The lack of financial resources affects an agency's capacity for system development, maintenance, and data administration. Insufficient resources also prevents an agency from buying and maintaining the hardware and software required for their ongoing tasks at hand within their jurisdictions. Many agencies reported that their divisions have undergone large budget cuts, to the extent that they can barely manage the data they currently have. Few agencies could imagine participating in any significant way in an integrated, internet-based geoscience system for Canada under their current budgets.

Some agencies made it a point to state that their top priority is to gather new data. Given limited budgets, agency money is spent on field work and map production based on that field work. Field work and mapping takes precedence over data management or system development. Barring a separate budget that is directed toward data management, many agencies will find it difficult to justify increased expenditures on data management.

4.1.2 Cost-recovery by Geomatics Canada

Many agencies feel that the cost-recovery approach of Geomatics Canada must be revisited before an integrated, internet-based geoscience system for Canada can be considered. Almost all agencies regarded the current policies to be a serious barrier to web-enabling geological and other maps. With rare exceptions, most agencies felt very strongly about this issue, making statements such as “Cost recovery is a disgrace,” and that this issue “...is starting to seriously impact federal and provincial programming.” Many agencies felt that unless they could freely distribute digital geomatics data along with their other data, there was little point in making their geological maps available on the Internet. Some agencies said that they have given up on Geomatics Canada and now get their data from the USGS or other agencies whose geomatics data is freely available.

The strength of feeling about this issue should not be underestimated. It is important that the issue of cost recovery be resolved as part of developing an integrated, internet-based geoscience system for Canada.

4.1.3 Local users fees to acquire data

Different agencies have or are developing different policies for cost recovery of data for which they are custodians. The differing policies may erect serious barriers to the sharing of data assets and the creation of an integrated system.

Most agencies felt strongly that as the Canadian taxpayer has paid for the data, the data should be made freely available over the Internet. Not all agencies agreed, however. In some cases, a cost-recovery policy already exists for non-Internet distribution and the same policy has been or will be extended to Web distribution.

Almost all agencies charge a nominal fee (usually $10.00) for a CD copy of their data, and for reproduction of maps and other hard copy documents. Most do not anticipate translating these reproduction fees to an Internet environment. Many agencies said that they do not want to be in the e-commerce business. In spite of this, all agencies noted that the ultimate decision on whether to charge for data rests with the policy makers.

4.1.4 Lack of industry presence in the modeling process

The surveyed agencies embrace a mandate to serve their clients and recognized that the clientele ranges from the mining or petroleum industry to the Canadian public at large. The question is, how best to determine and meet client requirements? Some agencies have formal client-based advisory boards who advise on system development; others do their best with less formalized contacts.

All agencies felt that industry must be involved in the data modeling process. However, agencies disagreed on the nature of industry involvement. Most felt that representatives from a few of the larger industry clients (particularly those with broad regional interests) should be invited to participate in the data modeling process. It was suggested that a detailed survey of industry clients would not be especially beneficial.

If a few industry clients were chosen, there is a good chance they would be from the mining and petroleum sector. Yet the surveyed agencies emphasized that engineering companies, insurance companies, real estate offices, and environmental protection agencies also are important industry clients. A third tier of clients, equally important, is comprised of other geological surveys, educational institutions (e.g., universities, colleges, high schools), and the Canadian general public. Every agency surveyed commented on the fact that the Canadian taxpayer paid for the data collection so at some level, the Canadian taxpayer is the ultimate client.

Because the clientele is so diverse, a detailed approach to requirements analysis could be all-consuming. As a practical solution, most agencies agreed that targeting the needs of the largest users of the data (generally industry) by involving them in the modeling process should also meet the needs of the wider clientele.

4.1.5 Leaders lose

Some agencies that have already invested time and resources to establish a web presence expressed concerns that they could be penalized for their leadership. This would occur if funding were made available only to agencies without good web sites, so that they can "catch up." None of the agencies expressing this opinion regret the work they have done on their web sites. They noted that all agencies have different priorities, but in times of scarce funding, any little bit extra helps.

To ensure that an integrated, internet-based geoscience system for Canada involves all participants, any available funding should be distributed to *all* agencies. A plan could be established where funding granted to a particular agency would be used to either help that agency establish their Internet presence or expand it. Those agencies that already have extensive web sites could test new options for the benefit of the entire community.

4.2 Provincial and Territorial Agencies

The issues and concerns listed below are ones that were most commonly expressed by provincial or territorial agencies. While some of these issues apply to all agencies (including federal), they tend to have a stronger provincial or territorial component.

4.2.1 Need for a local presence

GSC policy is inadvertently making it difficult for provincial and territorial agencies to answer to local funding authorities, which is critical if they are to meet their mandates. Provincial and territorial agencies must maintain a profile within their community to show local taxpayers and politicians that their monies are well spent. To this end, agencies must produce a variety of products, from hard copy maps to web sites, with a dominant and recognizable provincial or territorial stamp.

For example, it was noted that if the GSC has any involvement whatsoever in a product (e.g., a geological map), all three GSC logos (NRCan, GSC, and the Canadian flag) must be placed on that product. The visibility of the single provincial or territorial logo is vastly reduced, even if the vast majority of funding came from the province or territory. This is a serious issue for agencies whose products justify their funding.

Any integrated, internet-based geoscience system for Canada must recognize the needs of the provincial and territorial agencies to have a local presence. Giving each participant the credit due to them is key to the budgetary process, and to the scientific process. A strategy must be in place to clearly identify data ownership and agency participation. In addition, distributed data holdings should be supported so provincial and territorial agencies can manage their own data assets.

4.2.2 Issues of trust

One of the most difficult issues raised by several provincial and territorial agencies is that of trust. Why should agencies involve themselves in any large-scale initiatives with the GSC if the GSC can pull out without warning or consultation as has occurred in the past? Many agencies are concerned about how an integrated system would be established, and all agencies agreed that at the very least, a common portal to their own local web sites would be a good first step. This is seen as a way of maintaining local presence while re-establishing trust. Most agencies felt that small steps rather than a large leap to an integrated system were required.

All agencies stated that they have good relationships with the federal Geological Survey of Canada at a personal level; but numerous concerns were raised about dealing with the GSC as an entity. The comment was made that the only reason provincial and territorial agencies have been included in this process to develop an integrated, internet-based geoscience system for Canada is that the GSC was required to include them to obtain the funding. More than one agency felt that the GSC would end up dominating the process, and that those agencies with minimal budgets (and thus unable to attend meetings) would be excluded.

Two GSC developments were frequently mentioned as past examples of this type of behaviour, where the GSC involved provincial and territorial agencies only to drop them later without consultation. The first was GeoScan. Many agencies are still bitter that they invested time and data in GeoScan only to have it dropped. The second development commonly mentioned is FieldLog. Agencies noted that the GSC promoted FieldLog as a useful tool for the field, and a mechanism to exchange data with the GSC. The GSC has now announced that it will no longer support FieldLog. This is seen by some agencies as yet another example of the GSC involving them only to withdraw support when no longer convenient.

There may have been good reasons for both developments to be dropped by the GSC; still, a significant number of provincial and territorial agencies feel that there was a breach of trust. This is an extremely important issue to address in the development of an integrated system for Canada.

4.2.3 Long term funding

Most agencies would like to see funding put in place to support a long-term strategy (minimum of five years). This strategy should involve *all* agencies and ensure adequate funding to allow their participation.

The concern is that the development of an integrated, internet-based geoscience system for Canada would otherwise be just an ongoing series of ad hoc meetings without any real development or commitment. Some agencies thought it a waste of their time to participate unless long-term funding was guaranteed. They do not believe that good will and the occasional injection of funds is enough to make this happen. This sentiment was widely expressed by provincial, territorial, and various GSC agencies.

4.2.4 Multilingual solutions

To ensure that all agencies can participate equally in an integrated Canadian system, it will be necessary to establish clear language policies on the Internet publication of data. Most agencies stated that nothing seems to be written down and that everyone is looking to everyone else for a policy in this area. If very strict interpretations of the language laws are to be enforced, then funding should be made available to help with translations.

It is assumed that a national, internet-based geoscience system for Canada must support multiple languages. The issues is how agencies will meet this requirement when making data available over the Internet. Agencies appear to have different policies regarding the language laws and how they apply to primary data. In some agencies, internal regulations require the translation of all data to both official

languages; in other agencies, primary scientific data may remain in the original language of the scientist.

The situation becomes more complex when dealing with aboriginal languages. Many aboriginal languages have no geoscientific vocabulary, which precludes effective translation.

5 Conclusions

5.1 Characteristics of existing systems and approaches

Geological maps and hydrocarbons are the only two categories where there is a consistent use of a public data model. Within the geological map category, there is significant interest in the North American Data Model (NADM v.4.3) and the CordLink model. For hydrocarbons and related data, the Public Petroleum Data Model (PPDM) is dominant.

The only other standard that makes a common appearance among the surveyed agencies is the Federal Geographic Data Committee (FGDC) standard for digital geospatial metadata. No consistent data format standards transcend a particular data type. Geophysical data are somewhat unique, in that international data format standards do exist for some types of geophysical data (e.g., SEED for natural source earthquake data and SEG-Y for seismic data).

Agencies producing digital geological maps commonly use a Geographic Information System (GIS) tool to visualize the results. In most cases, all the map data are stored in the GIS tool rather than in a separate relational database management system. Almost three-quarters of agencies use ArcInfo, ArcView, or MapObjects in some aspect of interaction with their geological maps.

For the remaining categories, it is more difficult to identify a common approach. The data sets are disparate and the solutions are often as disparate as the data sets. However, some general trends may be identified.

Most agencies use either Oracle or Ingres for their high-end database management tools. At least some of the agencies using Ingres are looking to move to an Oracle environment. Thus, Oracle is likely to achieve dominance among the surveyed agencies. For PC-based database management, MS Access was used most frequently. Ease of use and general availability were often cited as reasons for using Access. All of these dominant databases have an SQL underpinning, as do most of the other database management systems reported. In addition, almost all database systems

were relational rather than object-relational or fully object-oriented. However, there is some interest in an object-oriented approach.

All agencies indicated an enthusiasm for the Internet. The ideal is seen as a map of Canada where the user can drill down to the actual, primary data. Most agencies consider this completely unrealistic at this time but find the general idea appealing.

Every agency surveyed either has published some of their data sets to the Internet or is looking to do so in the near future. No single approach has been established for publishing to the Internet, but MapGuide and MapObjects are often mentioned for geological maps. For other types of data, data sets are commonly provided as downloadable files from a simple HTML or FTP interface. Some agencies already link a variety of data sets to their maps, and almost all other agencies are looking to do the same.

The overwhelming majority of agencies do not want to see user fees charged for geoscience data made available over the Internet. Many feel that the data belong to the public already and they are not interested in maintaining and e-commerce site. Nonetheless, there are agencies with cost-recovery policies presently in place.

5.2 Areas of potential cooperation

Throughout the interviews (in person and by phone), a common question was “What is everyone else doing?” There is widespread interest in learning how other agencies have approached a particular type of data: what worked, what didn’t work, and why. Also, many agencies are grappling with decisions on how best to publish data to the Internet; those agencies could benefit greatly from the experience of others.

Each of the six categories defined by the Canadian Geoscience Knowledge Network (CGKN) and the Canadian Data Model Working Group (CGDMWG) could benefit from a technical working group to analyze requirements in that category, keeping local, regional, and industrial/client needs in focus. In addition, a technical working group on Internet strategies would be useful. This group would perform and present evaluations on the technical options for web-enabling various types of data sets; this service would save money across Canada by permitting each new development to build on the knowledge of previous efforts. In addition, each agency could choose a method whose characteristics (e.g., cost, hardware needs, software needs, complexity) were in keeping with that agency's goals.

5.3 Challenges and Potential Barriers

Numerous challenges must be faced in defining a common Canadian GeoScience data model for an integrated, internet-based geoscience system for Canada. Some high-level challenges are as follows.

- Developing a long-term strategy for future development and keeping funding in place to support that development.
- Addressing the cost recovery issue surrounding geomatics data.
- Involving industry in the modeling process.
- Defining a strategy for the sharing of data assets given differing policies among agencies related to user fees.
- Ensuring that *all* participating agencies are supported.
- Recognizing that provincial and territorial agencies must maintain a local presence.
- Building trust among agencies.
- Drafting policy on linguistic responsibilities of agencies publishing to the Internet.

Any one of the challenges listed above could develop into a significant barrier for activities leading to an integrated system and so must be addressed as part of any proposal for a Canada-wide system.

5.4 Next Steps

Given the numerous challenges, the need for a firm policy basis, and serious issues regarding levels of trust, the first steps should be small ones. Small incremental steps are a virtual necessity to ensure cooperation. All agencies agreed that a good first step would be to build a common web portal to Canadian geoscience data. This portal should allow the user to search by geographic area or by subject. Using the search criteria, the portal would direct the user to the appropriate agency's (or agencies') web site. For some agencies, even organizing a subject classification (or simplified form of metainformation) will be a significant undertaking. However, all agencies wished to participate in this process. The subject classification and framework that support it will represent the beginning of a common information framework.

A common geoscience data model for Canada is more problematic. There are no simple answers. This issue and recommendations for a common web portal are addressed in the follow-up draft report entitled “Recommendations for a Geoscience model to support an integrated, Internet-based Geoscience Network for Canada.”

Appendix A: Summary tables

In all the tables that follow, the absence of an entry reflects only the absence of an entry on the respondent's written survey. It does not necessarily indicate that an agency does not collect a data type, or use a type of software, model, or standard.

Table 1. Geological map standards and extent

	Name or Description of data set	Metadata standards	Data model standards	Spatial Reference	Coordinate System	Internet	Coverage	Accuracy of Geographic Locations	Approx. size of holdings
Yukon	Yukon Digital Geology			GRS 80	Albers, UTM, LL	f	Yukon wide	>100m	650 MB
	Yukon Geoprocess File			GRS 80	UTM, LL, Trans. Mercator		South & Central YK	>100m	350 MB
NWT		FGDC		NAD 27 NAD 83	UTM	f	Parts of NWT & Nunavut (10%)	30-200m	900 MB GIS 1100 MB CAD
Nunavut									
BC	Mineral Potential Maps			NAD 83	Geogr.	c	BC	1:250,000	100 MB
	Digital Terrain & Soils Map Library			NAD 83	Geogr.	c	BC	Variable 1:250K to 1:15K	> 10 GB
Alta	Radarsat	GILS	Possible for future: NADM	NAD 83	UTM		Alta – north of 55 degrees N	+- 25 m	37 GB
Sask	Geological Map of Sask.			NAD 27	UTM	f	All of Sask.	1:1 million	2 MB
	Summary of investigations			NAD 27	UTM	f	Northern Sask.	1:20K	300 MB
Man	Geological Maps	FGDC		NAD 27 NAD 83	UTM	f	Manitoba	5-50m	15 GB
Ont	ERLIS	GOITS 72.0 (rev'd FGDC)	Future: Combine NADM and LIO	NAD 27 NAD 83	UTM LL Lambert	b	Ontario	Variable 1:4M to precise	220 GB

	Name or Description of data set	Metadata standards	Data model standards	Spatial Reference	Coordinate System	Internet	Coverage	Accuracy of Geographic Locations	Approx. size of holdings
Que	SIGÉOM			NAD 83	UTM	c		precise	1000 sheets at 1:50K
NB	Bedrock Geology			ATS 77	NB Stereo LL		New Brunswick	relative	10 GB
	Surficial Geology			ATS 77	NB Stereo LL		New Brunswick	relative	10 GB
NS	Onshore NS Geological Maps			NAD 27	UTM	p	NS - Onshore	Variable. Digitized to <1mm at map scale.	120 sheets, 200 MB
Nfld	GeoLegend	FGDC, ISO TC 211. Future: OpenGIS	Expected: NADM	NAD 27	UTM	f	Nfld & Labrador	+- 1mm at map scale	12 MB
GSC-Sidney		Expect to follow: SDTS	Expected: NADM	NAD 27	UTM				
GSC-Pacific	Marine & general geology		Expected: NADM				W. coast & Cordilleran	Variable	
	Bedrock geology		NADM by fall 2000		UTM		Southeast BC		
GSC-Calgary	Central Forelands NATMAP		NADM	NAD 83	UTM		94N, 95C, 94G	50 m	
GSC-Ottawa									
GSC-Quebec	Surficial geology, geology			clrk 66, WGS 80	UTM, LL, lambert				
	Hydrogeological maps		NADM	WGS 80	LL				5 MB
GSC-Atlantic	ED (Expedition DB)	FGDC	ODP in part	NAD 27	LL	c	Marine & coastal areas of Canada	1 m	10.7 M records
	PAD (Physical Attribute DB)	FGDC		NAD 27	LL	c	Marine & coastal areas of Canada	1 m	36K records

	Name or Description of data set	Metadata standards	Data model standards	Spatial Reference	Coordinate System	Internet	Coverage	Accuracy of Geographic Locations	Approx. size of holdings
	MetaDB	FGDC		NAD 27 NAD 83	LL	c	Canada but mostly marine areas	1 m	640.7K records

Legend: b = in beta f = Future
c = currently used p = partial

Table 2. Geological mapping software

	Adobe Illustrator	ArcInfo	ArcView	AutoCAD	AutoCAD Map	CARIS	MapGuide	MapInfo	MapObjects	Microstation	Other	Associated DB Software
Yk		c	c	c							SurView	
NWT		c	c	c[1]	f				c			
Nunavut		c	c									
BC		c	c				c			c	QuikMap	MS Access
Alta		c	c				f				ER Mapper	
Sask	c[2]	f	c		c[1]							MS Access
Man		c			c[1]			c				
Ont		c	c	c	c[1]				c	c[3]		Oracle, MS Access
Que										c		Oracle
NB						c						
NS		c	c	c[1]					c			MS Acess, FoxPro
Nfld			c	c		c		c			Geosoft Oasis Montaj	MS Access
GSC-Sidney		c	c		c		c[4]					
GSC-Pacific												
GSC-Calgary		c	c							c		Oracle, MS Access, Excel
GSC-Ottawa		c	c	c[1]	c[1]			c	c[4]			MS Access
GSC-Quebec		c	c	c[1]	c[1]		c[5]	c	c[4]	c	Idrisis 32,Geomatter, ColdFusion	MS Access, Delphi
GSC-Atlantic		c				c					Z-Map Plus	Oracle

Legend: b = in beta f = future
c = currently used p = partial

[1] Used by field geologists with FieldLog to create maps.
[2] Used with ArcView to publish maps.
[3] Used to produce paper maps
[4] Used to publish to the web
[5] Used with ColdFusion to web-enable

Table 3. Mineral databases standards, software, and extent.

	Name or Description of data set	Metadata Standard	Data Model Standards	Data Format Standard	Spatial Refernce	Coordinate System	MS Access	Ingres	Oracle	SQL Server	ArcInfo	Excel	Other	Internet	Coverage	Approx. size of holdings
Yk	MINFILE				GRS 80	UTM LL	c								Yukon wide	35 MB
NWT	NORMIN.DB				WGS 84	LL			c					c	NWT & Nunavut	4000 MB incl all db objects
Nunavut																
BC	MINFILE Mineral Occurrence			ASCII .dbf	NAD 27 NAD 83	UTM LL							FoxPro, R&R, MINFILE/pc, Cold Fusion, HTML	c	BC complete	70 MB
Alta	Coal/CBM database					ATS LL	c	c								> 1 GB
	MIS database					various	f	c	f		c		ArcView		All of Alberta	
Sask	Mineral Deposits Index				NAD 27	UTM	c							c	Northern Sask.	10 MB
Man	Mineral database	FGDC			NAD 27	UTM				c				f	Manitoba	50 MB
Ont	Mineral Deposit Inventory Database (MDI2)	GOITS72.0 Rev'd. FGDC		ERLIS	NAD 27	UTM	c		c					b	Ontario	20,000 mineral occur.
Que	SIGÉOM Mineral deposits			Que.	NAD 83	UTM			c					c	Quebec	
NB	Mineral deposits				ATS 77	NB Stereo LL	c	c & f							New Brunswick	20 MB
	Mineral Rights				ATS 77	NB Stereo LL		c							New Brunswick	3 GB

	Name or Description of data set	Metadata Standard	Data Model Standards	Data Format Standard	Spatial Refernce	Coordinate System	MS Access	Ingres	Oracle	SQL Server	ArcInfo	Excel	Other	Internet	Coverage	Approx. size of holdings
NS	Mineral occurrences database			NS	NAD 27	UTM					c		ArcView, FoxPro	c	Nova Scotia onshore	5 MB
	Mineral Rights				NAD 27	UTM					c		Origin Advanced Revelations		Nova Scotia onshore	5-10 MB
Nfld	MODS	in CEONet using FGDC		NFLD	NAD 27	UTM	c							c	Nfld & Labrador	25 MB
GSC-Sidney																
GSC-Pacific																
GSC-Calgary																
GSC-Ottawa																
GSC-Quebec																
GSC-Atlantic																

Legend: b = in beta f = future
c = currently used p = partial

Note: The absence of an entry does not necessarily indicate that the agency does not collect minerals data. Also, the given entry may not represent all of the minerals data collected by the agency. This table only represents those data sets (databases) reported in the survey.

Table 4. Hydrocarbon databases standards, software, and extent.

	Name or Description of data set	Metadata Standard	Data Model Standards	Data Format Standard	Spatial Refernce	Coordinate System	MS Access	Ingres	Oracle	SQL Server	ArcInfo	Excel	Other	Internet	Coverage	Approx. size of holdings
Yk																
NWT																
Nunavut																
BC																
Alta	CIFEDB Drill stem tests	Plan for GILS				LL		c	f						Alta, BC, Sask	153 MB
	Athabasca Oil Sands Database		PPDM	LAS, Synth. LAS									PRISM			40 MB (geophys.logs) 60 MB (log analysis) 2 MB (strat. picks)
Sask	Well information system					LL							CICS		Southern Sask.	200 MB (67000 wells)
Man	MOGWIS		PPDM		NAD 83	UTM LSD			c					p	Manitoba	1 GB
Ont																
Que																
NB	Energy Rights				ATS 77	NB Stereo LL		c							New Brunswick	1 GB
NS	Onshore Wells Database				NAD 27	UTM					p			f	NS- Onshore	100 records
Nfld																
GSC-Sidney																
GSC-Pacific																

	Name or Description of data set	Metadata Standard	Data Model Standards	Data Format Standard	Spatial Refernce	Coordinate System	MS Access	Ingres	Oracle	SQL Server	ArcInfo	Excel	Other	Internet	Coverage	Approx. size of holdings
GSC-Calgary	WELLS		PPDM		NAD 27	LL			c						BC, Alta, Sask., Yk, NWT, Nunavut, Arctic Ocean, some east- and west-coast offshore	2.5GB
GSC-Ottawa																
GSC-Quebec																
GSC-Atlantic	BASIN	FGDC in Metadb	PPDM (in part)		NAD 27 & NAD 83	LL			c					c	Offshore eastern Canada + Arctic	8.9 M records

Legend: b = in beta f = future
c = currently used p = partial

Note: The absence of an entry does not necessarily indicate that the agency does not collect hydrocarbon data. Also, the given entry may not represent all of the hydrocarbon data collected by the agency. This table only represents those data sets (databases) reported in the survey.

Table 5. Geochemical databases standards, software, and extent.

	Name or Description of data set	Metadata Standard	Data Model Standards	Data Format Standard	Spatial Refernce	Coordinate System	MS Access	Ingres	Oracle	SQL Server	ArcInfo	Excel	Other	Internet	Coverage	Approx. size of holdings
Yk																
NWT	KIDD (Kimberlite Indicator & Diamond DB)				WGS 84, NAD 27	UTM LL						c	dBase, Arc Explorer		Parts of NWT & Nunavut	20 MB xls 10 MB dbf
	Economic Geology Series	FGDC			NAD 27	UTM							Various		Parts of NWT & Nunavut (sporadic)	300 MB
Nunavut																
BC	Geochemisty											c	Shape		B.C.	
Alta																
Sask	Litho geochemistry				NAD 27	UTM						c		f	All of N. Sask.	10 MB
Man	Multimedia geochemical surveys	FGDC			NAD 83	UTM								f	Manitoba local	2 GB
Ont	Lithogeochemical Database (LGC)	GOITS 72.0		ERLIS	NAD 27	UTM	c		c					b	Ontario	30,403 rock sample analyses
Que	SIGÉOM Sediments geochemistry			Que.	NAD 83	UTM			c					c	Quebec	550K samples 14 M results
NB	Till, Silt, Rock				NAD 27 NAD 83	UTM LL		c							New Brunswick	10 GB
NS	Geochemistry database				NAD 27	UTM					c		Origin in dBase	c	NS Onshore	2 MB

	Name or Description of data set	Metadata Standard	Data Model Standards	Data Format Standard	Spatial Refernce	Coordinate System	MS Access	Ingres	Oracle	SQL Server	ArcInfo	Excel	Other	Internet	Coverage	Approx. size of holdings
Nfld	Lake Sediment Geochemistry			NFLD	NAD 27	UTM							dBase Point Shape files, Contour line shape files, Georef tif imagaes		All of Nfld and Labrador	260 GB
	Geochemical Atlases				NAD 27	UTM							Shapefile mif		All of Newfoundland & Labrador	20MB
GSC-Sidney																
GSC-Pacific																
GSC-Calgary																
GSC-Ottawa																
GSC-Quebec	Till geochemistry, Water temperatures				clrk 66, wgs 80	UTM LL lambert	c						GIMS			3000 entries
GSC-Atlantic																

Legend: b = in beta f = future
c = currently used p = partial

Note: The absence of an entry does not necessarily indicate that the agency does not collect geochemical data. Also, the given entry may not represent all of the geochemical data collected by the agency. This table only represents those data sets (databases) reported in the survey.

Table 6. Geophysical databases standards, software, and extent.

	Name or Description of data set	Metadata Standard	Data Model Standards	Data Format Standard	Spatial Refernce	Coordinate System	MS Access	Ingres	Oracle	SQL Server	ArcInfo	Excel	Other	Internet	Coverage	Approx. size of holdings
Yk	Geophysics				GRS 80	Albers UTM					c				Yukon wide	57 MB
NWT	Airborne mag & EM			ASCII xyz; raster & vector	WGS 84	UTM							Geosoft Geopak		Parts of NWT & Nunavut (1%)	60,000 line-km
Nunavut																
BC																
Alta	Geophysical logs		PPDM	LAS		LL & ATS							PRISM			40 MB
	Athabasca Oil Sands (see Table 3.2.2.1)															
Sask																
Man																
Ont	ERLIS	GOITS 72.0		Raster Vector ASCII Ontario Master Grid Format	NAD 27	UTM LL Lambert			c		c		Geosoft Grid Tiff	b	Ontario	32 detailed surveys. Province wide grids.
Que	SIGÉOM Input anomalies			Que.	NAD 83	UTM			c					c	Quebec	

	Name or Description of data set	Metadata Standard	Data Model Standards	Data Format Standard	Spatial Refernce	Coordinate System	MS Access	Ingres	Oracle	SQL Server	ArcInfo	Excel	Other	Internet	Coverage	Approx. size of holdings
NB	Geophysics				NAD 83	LL		c							New Brunswick	40 GB
NS	Geophysics image data set. Aeromag & GammaRay				NAD 27	UTM							ArcView	p	NS – Onshore	500 MB in images
Nfld	Aeromagnetics				NAD 27	UTM							dBase Point Shape files, Contour line shape files, Georef tif images, Geosoft grids		Various areas of Nfld & Labrador as submitted by exploration companies & regional govt surveys	40-60 MB
GSC-Sidney	Geodetic data			For GPS: SINEX& RINEX												
	Aeromag&Gravity														Cordilleran	
	Digital Elevation (DEM)			Future SDTS											Cordilleran	
	Geomatics														Patchy	
	Natural Source Earthquake			SEED											Western Canada	
GSC-Pacific																
GSC-Calgary	Reflection Seismic field data			SEG-Y	WGS 84	UTM LL							ProMax		Canadian Arctic Islands	169 GB
	Reflection seismic stacked data			SEG-Y	WGS 84	UTM LL							ProMax Geoquest		Canadian Arctic Islands	50 GB
	Shotpoint locations			ASCII	WGS 84	UTM LL							Zycor		Canadian Arctic Islands	45 GB

	Name or Description of data set	Metadata Standard	Data Model Standards	Data Format Standard	Spatial Refernce	Coordinate System	MS Access	Ingres	Oracle	SQL Server	ArcInfo	Excel	Other	Internet	Coverage	Approx. size of holdings
	Gravity data coincident with arctic seismic			ASCII	WGS 84	UTM LL							Zycor		Canadian Arctic Islands	10 MB
GSC-Ottawa	National Wave Form Archive			SEED									Veritas		All of Canada	4 TB (plus 0.5 TB/yr)
	National Earthquake DB					LL		c	f						All of Canada	2.5 GB
	National Aeromagnetic DB	FGDC		.gxf, grid files	NAD 27	LL			c				Cube Werx	p	Canada offshore, some foreign	8-10 GB
	National Gravity DB	FGDC			NAD 83	LL			c					p	Canada & offshore	3 GB
GSC-Quebec																
GSC-Atlantic	BASIN (see Table 3.2.2.1) include multichannel seismic reflection															

Legend: b = in beta f = future
c = currently used p = partial

Note: The absence of an entry does not necessarily indicate that the agency does not collect geophysical data. Also, the given entry may not represent all of the geophysical data collected by the agency. This table only represents those data sets (databases) reported in the survey.

Table 7. Assessment reports databases standards, software, and extent.

	Name or Description of data set	Metadata Standard	Data Model Standards	Data Format Standard	Spatial Refernce	Coordinate System	MS Access	Ingres	Oracle	SQL Server	ArcInfo	Excel	Other	Internet	Coverage	Approx. size of holdings
Yk																
NWT	Exploration industry assessment reports				various	various									Most of NWT & Nunavut, esp. mainland	Paper based 700 line-ft
	Normin References				WGS 84	LL			c					c	Much of NWT & Nunavut, esp. mainland	8400 MB incl all db objects
	MINERS (mineral claims, leases, permits)				GRS80	LL			c				shp (arcview)		NWT & Nunavut	
Nunavut																
BC	ARIS			.pdf	NAD 27 NAD 83								FoxPro Cold Fusion	c	B.C.	311 MB 26,000 reports
Alta																
Sask	Metadata file of survey and assessment reports				NAD 27	UTM	f						InFocus ArcView (f)	f	All of Sask.	200 MB
Man	Mining Recording Information System				NAD 83	UTM			c					p	Manitoba	
Ont	Assessment File Research Imaging (AFRI)	GOITS 72.0		Tiff JPG					c					b	Ontario	60,000 files 150,000 maps 2.5M pages
Que																
NB	Assessments				ATS 77	NB Stereo LL	c	c							New Brunswick	5 GB

	Name or Description of data set	Metadata Standard	Data Model Standards	Data Format Standard	Spatial Refernce	Coordinate System	MS Access	Ingres	Oracle	SQL Server	ArcInfo	Excel	Other	Internet	Coverage	Approx. size of holdings
NS	Assessments				Mapsheet claim reference area									p	NS – Onshore	Paper based
Nfld	GEOFILES – Biblio. Assessment			NFLD	NAD 27	UTM							STAR	f	All of Nfld and Labrador	Approx. 20K records
GSC-Sidney																
GSC-Pacific																
GSC-Calgary																
GSC-Ottawa																
GSC-Quebec																
GSC-Atlantic																

Legend: b = in beta f = future
c = currently used p = partial

Note: The absence of an entry does not necessarily indicate that the agency does not collect assessment reports. Also, the given entry may not represent all of the assessment reports collected by the agency. This table only represents those data sets (databases) reported in the survey.

Table 8. Borehole databases standards, software, and extent.

	Name or Description of data set	Metadata Standard	Data Model Standards	Data Format Standard	Spatial Refernce	Coordinate System	MS Access	Ingres	Oracle	SQL Server	ArcInfo	Excel	Other	Internet	Coverage	Approx. size of holdings
Yk																
NWT																
Nunavut																
BC	Coal Borehole					UTM									B.C.	
Alta	AGSWDB					LL		c	f						AB plus sparse coverage BC & SK	
	Index of groundwater wells	Plan GILS				LL		c	f						Most of Alta	88 MB
	Athabasca Oil Sands (see Table 3.2.2.1)															
Sask	Mineralized core collections				NAD 27	UTM	c							f	Northern Sask.	1 MB
Man	Manitoba Stratigraphic Database (see Table 3.6.1.1)															
Ont	ODHDB	GOITS 72.0		ERLIS	NAD 27	UTM	c		c				ASCII	b	Ontario	110K drill holes
Que	SIGÉOM Diamond drill holes			Que.	NAD 83	UTM			c					c	Quebec	100K
NB	Drill Core Storage				ATS 77	NB Stereo LL									New Brunswick	20 MB

	Name or Description of data set	Metadata Standard	Data Model Standards	Data Format Standard	Spatial Refernce	Coordinate System	MS Access	Ingres	Oracle	SQL Server	ArcInfo	Excel	Other	Internet	Coverage	Approx. size of holdings
NS	Drill holes Database				NAD 27	UTM					c		Origin Advanced Revelations	c	NS – Onshore	5 MB
Nfld	Drill Core			NFLD	NAD 27	UTM	c						Point Shape files		All of Nfld and Labrador	5 MB
GSC-Sidney																
GSC-Pacific																
GSC-Calgary	Checkshot & crystal cable velocity surveys				WGS 84	UTM LL						c			Canadian Arctic Islands	3 MB
	Sonic, gamma & density logs			LAS	WGS 84	UTM LL							LOGM		Canadian Arctic Islands	20 MB
GSC-Ottawa																
GSC-Quebec	Surficial geology & hydrogeology				clrk 66	UTM	c						GIMS			40K entries
GSC-Atlantic																

Legend: b = in beta f = future
c = currently used p = partial

Note: The absence of an entry does not necessarily indicate that the agency does not collect borehole data. Also, the given entry may not represent all of the borehole data collected by the agency. This table only represents those data sets (databases) reported in the survey.

Table 9. Biostratigraphic databases standards, software, and extent.

	Name or Description of data set	Metadata Standard	Data Model Standards	Data Format Standard	Spatial Refernce	Coordinate System	MS Access	Ingres	Oracle	SQL Server	ArcInfo	Excel	Other	Internet	Coverage	Approx. size of holdings
Yk																
NWT																
Nunavut																
BC																
Alta																
Sask																
Man	Manitoba Stratigraphic DB	FGDC	PPDM		NAD 83	UTM LSD			c					f	SW Manitoba	500 MB
Ont																
Que																
NB																
NS																
Nfld																
GSC-Sidney																
GSC-Pacific	Paleontology DB		Georecords (in part)			UTM LL							Superbase		BC & Yukon	15K-20K point data records
GSC-Calgary	Paleo		Expect PPDM		NAD 27	LL			c			c	Paradox 4.5 Expect ArcView & ArcIMS			
GSC-Ottawa																
GSC-Quebec																

	Name or Description of data set	Metadata Standard	Data Model Standards	Data Format Standard	Spatial Refernce	Coordinate System	MS Access	Ingres	Oracle	SQL Server	ArcInfo	Excel	Other	Internet	Coverage	Approx. size of holdings
GSC-Atlantic	Dinoflaj	FGDC (in Metadb)							c					c	Worldwide	62.7 k records
	J-K Foraminifera								c						Worldwide	>3K records

Legend: b = in beta f = future
c = currently used p = partial

Note: The absence of an entry does not necessarily indicate that the agency does not collect biostratigraphic data. Also, the given entry may not represent all of the biostratigraphic data collected by the agency. This table only represents those data sets (databases) reported in the survey.

Table 10. Other reported databases standards, software, and extent.

	Name or Description of data set	Metadata Standard	Data Model Standards	Data Format Standard	Spatial Refernce	Coordinate System	MS Access	Ingres	Oracle	SQL Server	ArcInfo	Excel	Other	Internet	Coverage	Approx. size of holdings
Yk																
NWT	BIBLIO. EGS papers	FGDC		wpd, doc, pdf									WordPerfect Word Adobe Acrobat	f	DIAND publications	
Nunavut																
BC																
Alta																
Sask	Geochronology				NAD 27	UTM	c							f	Mostly Northern Sask.	1 MB
Man																
Ont	OGS Maps	GOITS 72.0		JPG											Ontario	8,000 raster 1,000 vector
	Abandoned Mines Information System (AMIS)						c								Ontario	
	Geology of Ontario	GOITS 72.0		SHP E00	NAD 27	Lambert LL									Ontario	1:250K 1:1M
	Publication database	GOITS 72.0		Tiff	NAD 27	UTM Lambert LL			c						Ontario	20 GB 12,500 reports
Que																

	Name or Description of data set	Metadata Standard	Data Model Standards	Data Format Standard	Spatial Refernce	Coordinate System	MS Access	Ingres	Oracle	SQL Server	ArcInfo	Excel	Other	Internet	Coverage	Approx. size of holdings
NB	BIBLIO. Geoscience publications				ATS 77	NB Stereo LL	c								New Brunswick	60 MB
	Peatland Rights				ATS 77	NB Stereo LL		c							New Brunswick	1 GB
	Peatland Resource				ATS 77	NB Stereo LL	c								New Brunswick	100 MB
NS	BIBLIO. NovaScan			Geo Scan std.	NAD 27	UTM	c				c		STAR	c	NS – Onshore, limited offshore	14K documents
	Abandoned Mine Openings				NAD 27	UTM							FoxPro	c	NS - Onshore	
	DEM 1:10K & 1:50K				NAD 27	UTM					c				NS – Onshore	500 MB
	Peatland Inventory				NAD 27	UTM					c			f	All of NS	5 MB
Nfld	BIBLIO. GEOFILES – Bibliographic Assessment (see Table 3.4.1)															
GSC-Sidney																
GSC-Pacific																
GSC-Calgary	SAMPLE. Research Sample Collection		PPDM		NAD 27	LL			c					f	All Canada. Some foreign samples	250 MB
GSC-Ottawa	BIBLIO. Publications Directory	FGDC				LL	c						AutoDesk & MapGuide	c	All Canada	150 MB

	Name or Description of data set	Metadata Standard	Data Model Standards	Data Format Standard	Spatial Refernce	Coordinate System	MS Access	Ingres	Oracle	SQL Server	ArcInfo	Excel	Other	Internet	Coverage	Approx. size of holdings
GSC-Quebec																
GSC-Atlantic																

Legend:

b =	in beta	f =	future
c =	currently used	p =	partial

Note: The absence of an entry does not necessarily indicate that the agency does not collect other types data. Also, the given entry may not represent all of the other types data collected by the agency. This table represents only the information reported by survey respondents.