

CONFERENCE OF EUROPEAN STATISTICIANS

UN/ECE Work Session on Methodological Issues Involving the Integration of Statistics and Geography

(Neuchâtel, Switzerland, 10-12 April 2000)

Topic (i): Policy and organizational aspects in GIS and statistics

**CANADA'S NATIONAL GEOGRAPHIC ALLIANCE FOR THE DEVELOPMENT AND
MANAGEMENT OF A NATIONAL GEOGRAPHIC DATABASE**

Submitted by Statistics Canada¹

Contributed paper

1. Over the last several years, government organisations in Canada have devoted significant resources into exploring issues of co-operative geographic data development and data sharing. These activities have energized especially those organizations involved with the production and maintenance of road networks. At this time, Elections Canada and Statistics Canada have successfully implemented a partnership that has created a national road network to serve both their operations; a first in Canada. Building on current momentum, a National Geographic Alliance is described that would build on efforts of Elections Canada and Statistics Canada to date, and expand beyond the original participants. The paper reviews the business model, technical issues and real advantages such an Alliance would present.

I. INTRODUCTION

2. Street network files have wide use in GIS applications. Road patterns, together with their associated road names, address range and other attribute data, make up an important part of mapping, geocoding and many other applications. With the growth of GIS applications in general, the demand for consistent, national coverage of streets in GIS ready formats in Canada has also grown accordingly.

3. Elections Canada (Canada's electoral agency) and Statistics Canada (Canada's statistical agency) both have need of such data, and increasingly are finding that the requirements for national level coverage are growing. The critical applications for street networks at both departments are mapping and geocoding. These organizations have worked together for several years now, building and harmonizing their own spatial frameworks, while sharing the burden of costs and expertise required to put a complete Canadian street network file in place.

4. At the same time other federal, provincial and municipal organizations have moved ahead with their own street network file projects. In recent years, more and more private sector companies have also appeared with street network file offerings, often in partnership with one or more public sector agencies. All this activity is promising for the future of GIS applications in Canada. Nevertheless, geographic street network data that is structured in a standard manner and up-to-date, and can be used for multiple applications, on a national scale, have not evolved quickly.

5. To be efficient, it is clear that an up-to-date national street coverage must take advantage of the wealth of existing data sources and data providers. At the same time, in order to provide a consistent data

¹ Prepared by Danny Wall.

structure upon which to build applications, some mechanism to harmonize the wide variety of data sources within a single framework, is also mandatory. Building largely on experiences gained from the Elections Canada – Statistics Canada partnership a model for creating a single national street network file is proposed; the National Geographic Alliance (NGA).

II. A NATIONAL GEOGRAPHIC ALLIANCE FOR CANADA – PART OF A LARGER INITIATIVE

6. There are a number of factors that enable the creation of the NGA at this time. These include the foundation of the Canadian Geospatial Data Infrastructure (CGDI) initiative, the availability of data from a wide range of sources, and database technology developments. The impact of technological developments on data has been well documented in many other forums, and the availability of street network data in Canada has already been noted. The CGDI initiative is however, worth reviewing here in some detail.

7. The CGDI was established in 1996 to promote the sharing and expanded use of geographic data in Canada. Supported by all major federal and provincial agencies and with participation by private sector and municipal organizations, the CGDI has become an important forum for discussions of geographic data sharing issues in Canada. There are five major thrusts within CGDI: data access, framework data, geospatial standards, partnerships, and supportive policy.

8. CGDI is critically important to the NGA since national street network files touch in some manner on all of these areas. It is expected that NGA members and participants in NGA would also be participants in the CGDI. In this way, NGA members will be in a good position to take advantage of policy directions, technical standards or distribution mechanisms as they evolve under the overall guidance of the Canadian geomatics industry. Conversely, NGA members will have the opportunity to positively influence developments within the CGDI, to ensure that data sharing and standardization with regard to street network files are facilitated.

III. STREET NETWORK FILES – IMPORTANCE TO NATIONAL SCALE GIS APPLICATIONS

9. Before discussing a business model for the NGA, the use of street networks in national scale GIS applications will be reviewed. While street network files may support any number of applications, two may be considered of primary importance, especially in controlling very large field programs, as is the case at both Elections Canada and Statistics Canada. The applications are geocoding – the linking of individual addresses to exact locations, and mapping.

10. Geocoding is surely one of the most widespread uses of GIS street network data. At Elections Canada, street network data is used to geocode their Register of Electors. At Statistics Canada, street network data is used to geocode the Address Register, a national scope list of dwelling addresses. Both registers contain millions of records, and form a crucial part of each agency's operational activities. The geocoding of such large data volumes is often a difficult task. To facilitate the process of geocoding, it is imperative that street network attributes, such as road names and address ranges, are maintained in a standard manner. Software solutions to geocoding non-standardized data holdings are possible but can be cumbersome and often are unreliable.

11. Mapping applications are considered here in a general sense, to mean both cartographic outputs and the maintenance of boundaries, both of which are prevalent at Elections Canada and Statistics Canada. Street networks are important in these applications in that they provide one of the key (and most important) cartographic components or reference features for mapping. Mapping applications impose a very real constraint on street network files – street networks must be positioned correctly, in at least a relative sense, to all the other reference features used in mapping applications, like water features and railways. This concept of vertical integration of different geographic feature themes is so important that

mapping applications are simply unworkable if it is not respected. To meet mapping requirements, relevant cartographic features must be considered along with road layers. At Elections Canada and Statistics Canada, they are maintained within the same database environment as the road layers.

12. The multiple uses of and constraints upon street network data at Elections Canada and Statistics Canada are likely typical of other national scale applications. At other federal agencies, there may be a focus solely on one or the other of these applications. However, such an approach may limit the growth of GIS applications beyond their supported application. To be more widely usable, a national street network must support more GIS solutions in a more general manner.

13. Compounding the complexity of building a standard national scale street network file, is the fact that there are multiple data sources from which it would be built and maintained. At first glance it may not be obvious why multiple sources may not simply be used 'as is' to support the applications discussed here. However, experience at Elections Canada and Statistics Canada has proven that without a consistent street network and supporting cartographic infrastructure, applications that must use such data are unwieldy and prone to error in outputs. This is unacceptable given that these applications support electoral and statistical activities.

14. The National Geographic Alliance proposes to resolve this issue by coordinating the maintenance of a national street network that is designed to meet the mutual requirements of its members. By explicitly concerning itself with issues of vintage and standardization, the Alliance will provide a geographic data resource that is coherent, up-to-date and usable across a broad spectrum of applications. Moreover, National Geographic Alliance members will share the cost of maintaining such a resource, with the result being a better street network than any member could have built alone.

IV. A BUSINESS MODEL FOR THE NATIONAL GEOGRAPHIC ALLIANCE

15. Given that the purpose of the NGA will be to build a shared street network resource usable by all members, the business model should respect certain shared values. As currently envisioned, several of these principles are stated here:

- the NGA will endeavor to maintain a file that as much as possible is useful to all partners within the Alliance; it must meet the mutual requirements of all partners;
- the NGA should encourage data providers to submit data to the NGA, by structuring their activities and acquisition methods in such a way as to consistent treatment and feedback to data providers;
- the NGA will not dictate how data should be collected or held by data providers, rather, it will react to and encourage consistent delivery of such data inputs. The NGA will avoid whenever possible data collection activities;
- the NGA, through its members will reflect directions indicated by the CGDI, and in the same way influence CGDI directions;
- the NGA will facilitate sharing of geographic data amongst partners by providing a common geographic framework based on the jointly maintained database and associated attribute information.

16. Figure 1 "A Conceptual Diagram for the National Geographic Alliance" shows how the NGA might be structured. The diagram shows three major participating groups with the NGA: the NGA Data Providers, the National Geographic Database (NGD) unit, and the NGA Partners. It should be noted at this time that much of this structure has in fact been put into place. However, it is the contention of the current partners that both the number of partners and data providers can be expanded as the NGA evolves. Finally, it should be noted that the NGD includes a structure by which partners may share their business layers with each other. This 'deposit library' will allow partners to deposit and retrieve time-stamped version of their business layers (such as electoral or statistical boundaries) directly from the NGD, rather than arranging such transactions between partners.

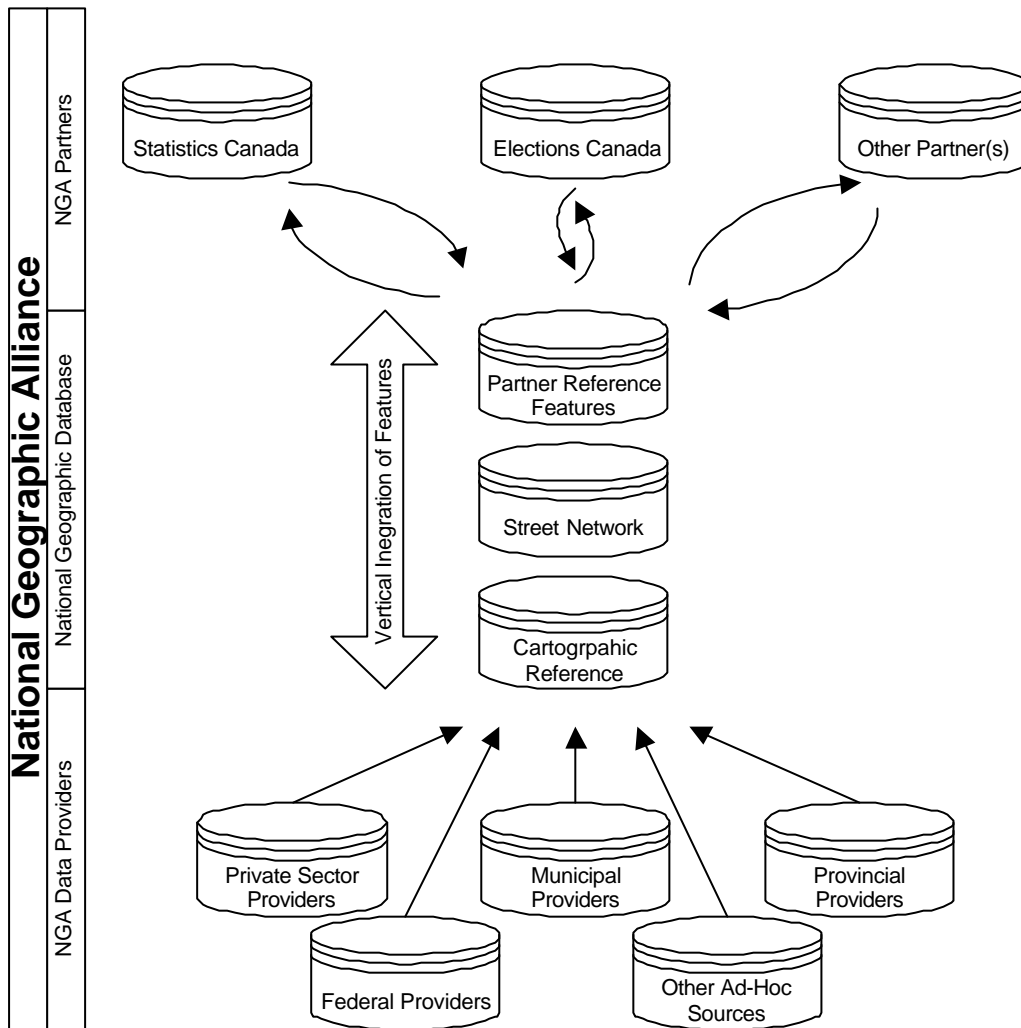


FIGURE 1 – A Conceptual Diagram for the National Geographic Alliance

17. The foundation of the NGA would be the data providers. These NGA members, whether through cooperative or business arrangements, will be encouraged to submit their data together with appropriate documentation for use in the NGD. Data providers would be strongly encouraged to be consistent in their deliveries. All data sources, including those with overlapping coverage will be reviewed in a consistent manner from the perspective of data content, quality, coverage and cost. This evaluation stage will be conducted by the NGD technical team, and will use a set of standard data characteristics, that would represent the NGD street network model and standards. Required modifications or manipulations of source data would be developed and become part of the ‘data profile’ for that provider.

18. The NGD unit would evaluate data sources, administer cooperative or business arrangements with data providers, develop integration solutions and maintain the NGA database. This unit would handle all of the technical aspects concerning data integration and data access by partners. Some of the technical challenges facing the NGD project team are discussed in the following section.

19. NGA partners, the last component in the model presented here, would provide resources required to operate the NGD unit, in addition to participating in its operation. The partners would guide activity within the database unit through two mechanisms, an oversight committee and a technical advisory committee. The oversight committee would provide final approval of data provider purchases or partnerships, and would manage the allocation of resources to the database unit. The technical committee

would ensure that the content, methods and processes in use within the NGA database unit met partner requirements and expectations. It is expected that partners would provide access to NGA data holdings through their own databases, rather than directly to the NGA database.

V. TECHNICAL ISSUES WITHIN THE NATIONAL GEOGRAPHIC ALLIANCE

20. The complexity of managing a single coherent street network updated and augmented from a wide variety of data sources cannot be underestimated. A quick review of some of the prominent technical issues is presented here.

V.1 Managing vertical integration

21. Using data from different sources nearly always provides challenges from the perspective of vertical integration. To meet these challenges, the NGA database will maintain road network geometry whenever possible with data purchased from Canada's national mapping agency. This means in turn that provider files must usually be 'fit' (rubbersheeted in GIS jargon) to existing geometry before a scan for updated features can be carried out. This process significantly increases the amount of work required for incorporating data from any given data provider.

V.2 Providing easy data access for partners

22. Allowing partner agencies easy access to updated information is a major objective of the NGD. To better support partner queries, comprehensive change tracking will be put into place. With such change tracking, a partner will be able to retrieve all types of changes to a given area over a period of time or all new streets across the country. Change tracking is critical to the operation of the NGD because it will allow partners to make very specific update queries, at their own discretion.

V.3 Data integration of diverse data inputs

23. Perhaps the most significant technical challenge for the NGD project team will be to integrate data from any number of data providers. Meeting this challenge will involve careful documentation of incoming data, and development of technical solutions for each data source. To the greatest degree possible, such solutions will be re-used on other data sources.

V.4 Difficulty of data 'upgrades'

24. While the National Geographic Database facilitates data sharing amongst partners, it also becomes in some sense the supporting database for many of the partners business activities. Because of this, major data upgrades, for instance replacing large sections of the NGD with better resolution data, may negatively influence partner business activities, if carried out without partner consultation. For this reason, data upgrades must be properly planned.

V.5 Allowing flexibility in partner participation

25. To facilitate NGA member participation or contribution of resources, software systems and procedures should be able to be implemented at any of the partner sites. To some degree, this complicates the design of systems and software, but allows easier participation by the alliance partners.

VI. CONCLUSION

26. Real growth in GIS applications can only occur if data exists to support them. The more consistent and accessible such data are the easier the task of building such applications and consequently the broader such application usage will be. Within the Canadian context, the NGA provides the vision for putting into place one a major data component for national scale GIS operations.

27. While major organizational and technical challenges exist in establishing such a vision, substantial systems and resources have already been implemented. Elections Canada and Statistics Canada, through leading this venture, have put into place a solid foundation for the future evolution of the NGA, and Canadian street networks.