MEETING OF THE WORKING GROUP ON TRAINING COURSES IN TOPONYMY

### Presentation to the International Cartographic
### Association on the Canadian Geographical Names
### Data Base (CGNDB)

Submitted by Kathleen O'Brien, Canadian Permanent Committee on Geographical \
Names (CPCGN) Secetariat, Canada.

# PRESENTATION TO THE INTERNATIONAL CARTOGRAPHIC ASSOCIATION ON THE CANADIAN GEOGRAPHICAL NAMES DATA BASE (CGNDB)

This shortened version of *The Canadian Geographical Names Data Base (CGNDB): its past, present and future* by David Fraser, Barbara Bowler, Jocelyne Revie, Kathleen O'Brien, and Paul O'Blenes was presented by Kathleen O'Brien, with the assistance of Jacqueline St-Clair, at the International Cartographic Association conference, on 17 August 1999. The full text of the paper appears in the *ICA 1999 Proceedings*.

## Overview of the CGNDB
The Canadian Geographical Names Data Base (CGNDB) is the authoritative national data bank of Canada's geographical names. It has a twofold purpose:
- to provide a national repository for geographical names that have been recognized by the Canadian Permanent Committee on Geographical Names (CPCGN), and
- to make these names available for both government and public use

Approximately 67% of the CGNDB's names records are officially approved geographical names to be used in government publications (e.g., maps, reports, etc.). The other names in the CGNDB are considered 'unofficial' names, and include formerly approved names, variants of the current names, and names that have been rescinded.

The CGNDB provides official names for various applications including: production of Canada's 1:50 000 and 1:250 000 national topographic maps; maintenance of the toponymy layer of Canada's National Topographic Data Base; gazetteer production; and a World Wide Web site.

Representatives of each province and territory and federal departments concerned with toponymy form the CPCGN. It is served by a Secretariat supplied by Natural Resources Canada. The CGNDB is maintained by the Geographical Names Section, which supports the Secretariat.

## Links to other databases
CGNDB names records were matched to Statistics Canada 1991 place name census records. The resulting digital file provides a link between official CPCGN place names and Statistics Canada population data.

The CGNDB is also used by federal departments as the authoritative source for geographical names, for example, in the filing of environmental impact reports. Plans have been developed for the regular import of records into the CGNDB from the Undersea Features Data Base, managed by the Canadian Hydrographic Service, a CPCGN member.

# The evolution of the CGNDB

## 1. Card-index registry to RDBMS client/server

In the late 19th century, the need for a single authority to which questions of geographical nomenclature and orthography could be referred was apparent. Such an authority was required because errors and inconsistencies existed in both the spelling and application of Canada's geographical names. As well, resource mapping beyond settled areas, together with the heavy immigration that was occurring at that time, made it imperative to deal with Canada's toponymy and geographical naming procedures.

As a result of this need for a toponymic authority, the Geographic Board of Canada (GBC) was established in 1897. Its mandate was to undertake the standardization of geographical names in Canada and to provide expert advice to federal departments and agencies on the origin, spelling and use of geographical names.

The Board created a card-index registry. It served as Canada's national repository for officially approved geographical names. As names were approved, index cards were created. Information on these cards included such items as: the official name; the province or territory in which the named feature occurs; the date that the name was officially approved; and any other names that had been used to refer to the place or feature. Information on these cards was organized according to data categories. Card formats varied over the years, but the information on them was the same.

## 2. First digital data base

During the 1970s, it was recognized that this analogue card-index registry could no longer satisfy the business requirements of the Geographical Names Section.

The amount of toponymic data being manually input into the card-index registry was increasing by 20% per year so maintaining the data was time-consuming and labour-intensive. It was also difficult to satisfy requests for names information on time.

It was decided to automate the names data base so DATABOSS/2, a data base management system, was purchased and the card records were encoded into it.

## 3. Second digital data base

In the late 1980s, a major overhaul of the CGNDB was required because of problems associated with the existing system:
- it was not user-friendly and not fully interactive
- its functionality did not have sufficient flexibility in both customizing names information and cross-referencing information in the data base
- its applications were unilingual , and
- it was installed on out-of-date hardware and was slow.

In 1988, the system was moved onto an Oracle-based system. Based on both relational DBMS technology and SQL, it provided significant flexibility. It also allowed for menu-driven user interface applications to be developed. And, it provided fast and easy access to the data.

## 4. The need to re-engineer the system

In 1997, an evaluation of both the business and technological issues facing the Geographical Names Section, as they relate to the CGNDB, was performed. The evaluation found that the CGNDB, in its current state, satisfied the business requirements of the Geographical Names Section. But the CGNDB was several versions behind most recent releases of Oracle Server and the application software was not supported by recent versions of Oracle Server.

As a result, the Geographical Names Section issued a contract to re-engineer the CGNDB. The major requirements were:
- to upgrade the version of Oracle Server used to manage the CGNDB
- to redevelop the CGNDB user interface, and
- to analyse and revise the CGNDB data model.

## Inputs into the CGNDB

The naming of geographical features in Canada is, today, the responsibility of each province and territory. Features located on Indian Reserves, national parks, and military establishments, require joint approval between the federal department and the appropriate provincial/territorial names authority. Names decisions approved by the provincial/territorial names authorities are sent to the Secretariat to maintain the national registry.

The CGNDB is not a self-contained system. Geographical names updates come from the provincial/territorial names authorities. They can either be in hard copy or a digital file.

The decisions are entered into the CGNDB. Manuscript copies of the National Topographic System (NTS) maps are updated to show the extent of the features and stored in files containing all correspondence material related to the approval of the names.

There are over 500 000 geographical names records in the CGNDB. About 14% of the records are populated places/administrative areas, etc., 63% are water features, and 23% are terrain features (e.g., mountains and peninsulas).

Over 30 attributes are stored for any name. They include: a unique identifier; various codes to indicate the name's status; the feature type; the region or territory in which the place/feature lies; and several location fields. In some cases, historical information about the origin/history of the toponym is also included.

## Business functions supported by the CGNDB

The CGNDB is used by the Geographical Names Section to perform its business functions including:
- maintaining accurate, up-to-date toponymic information for Canada
- maintaining the geographical names classification system – each name is assigned several codes. The generic code indicates if a feature is a river, lake, city, etc. Region code indicates the province or territory where a geographical feature occurs. Status code

indicates if a name is officially approved, was formerly approved but now rescinded, a cross-reference to another name, etc.

- generating NTS names lists – these are lists used in producing Canada's 1:50 000 and 1:250 000 national topographic maps and maintaining the toponymy layer of the National Topographic Data Base
- updating the Internet version of the CGNDB – this version of the CGNDB can be accessed for information about official and formerly official geographical names in Canada
- producing the *Gazetteer of Canada* series – this contains volumes for individual provinces and territories and, in 1997, the first *Concise Gazetteer of Canada*
- creating digital names products which clients may purchase as standard or customized digital files, and
- linking geographical names with World War II casualties. A registry of Canada's World War II casualties is contained within the CGNDB. The Geographical Names Section uses this registry to search for features that have been named after World War II casualties or to see if a particular casualty has been commemorated.

A contractor was chosen through the competitive bidding process. The Request for Proposal specified the technology required for use in this project but not methodology. The contractor recommended that Oracle's Custom Development Method be used as the development methodology. It would produce the information system in a timely manner and it would ensure that the resulting system satisfied the Section's stated business goals and objectives.

Oracle Designer was selected to work with Oracle CDM because:
- it is repository-based
- it supports the whole development process
- it provides reverse-engineering functionality
- it allows for multi-user access to information in the repository
- it provides a graphic modeling capability
- it can generate applications that can be deployed in both client/server and Web environments.

The re-engineering project was performed between November 1998 and July 1999, with a post-production support period from August – October 1999.


**Results of the re-engineering project: The new CGNDB system**
The work performed for the re-engineering project included:
- a complete redevelopment of the user interface
- a substantive revision of both the system data model and the physical database design
- an upgrade of the version of Oracle Server from 7.1.4 to Oracle 8.0.4; and
- a conversion of data from the existing system into the new system


**The CGNDB: its past, present and future**
The CGNDB has changed considerably from both a technical and a business perspective. The changes include:

- architecture of the CGNDB

Since its initial conception, the architecture of the CGNDB has evolved from a card-index registry to a completely digital system

- production of nameslists

Before the automation of the CGNDB, nameslists production was a manual, labour-intensive process. The use of an SQL RDBMS has dramatically decreased the time required to respond to a nameslist request and the resources required to create nameslists

- mechanisms for obtaining information about Canada's geographical names

The Geographical Names Section maintains an Internet version of the CGNDB. This reporting data base allows for the querying of geographical names and currently receives over 3,000 hits per day.

It should be further emphasized that both automating the CGNDB and providing access to it over the Internet has had three major impacts on the Geographical Names Section:

- it has dramatically decreased the time required by the section's staff to perform technical/clerical operations
- it has allowed staff to focus on the qualitative improvement of the names information in the CGNDB; and
- it has allowed the Section to better respond to the information needs of its clients

**Screen for inputting generic codes in the re-engineered CGNDB**



As in the past, changes in both technology and business will impact the future of the CGNDB. Methods of disseminating data files created by the CGNDB will change as technology does. Data exchange and partnerships will be important elements in this dissemination of information. Cooperation between federal, provincial, and territorial governments will be crucial. In addition,

the CGNDB will continue to preserve Canadian toponymic heritage as it has for over one hundred years. All of these factors will affect the future of the CGNDB, and the Geographical Names Section and its stakeholders.

Kathleen O'Brien
CPCGN Secretariat
November 1999