THIRD UNITED NATIONS CONFERENCE ON THE
 STANDARDIZATION OF GEOGRAPHICAL NAMES
Athens, 17 August-7 September 1977
Item 11 of the provisional agenda.
 Automated data processing (ADP).

THE COMPUTERIZATION OF GEOGRAPHICAL NAMES

THE QUEBEC EXPERIMENT* 1/

Paper presented by Canada

---

 * Paper prepared by Mrs. Grenon-Roy, geographer with the secretariat of
the Quebec Geographical Commission.

 1/ The French version of this paper appears in a special issue of the bulletin
"CANOMA", vol. 3, No. 1, which will be distributed at the Conference.

ENGLISH/FRENCH ONLY 1/

THE COMPUTERIZATION OF GEOGRAPHICAL NAMES

THE QUEBEC EXPERIMENT*

(paper presented by Canada)

At the Second United Nations Conference on the Standardization of Geographical Names, which was held in London in 1972, various countries underlined the ever-increasing role being played by computers in the field of toponymy. Storage requirements and the problems of data accessibility, combined with the need for place-name standardization, prompted the search for a technique providing instant access to information.

In Quebec, a place-name computerization project has been under way since 1971. It is known as the Projet Répertoire and consists in the recording and computer processing of official place names in Quebec. Two organizations have been taking part: the Geographical Commission of Quebec, which is responsible for official geographical nomenclature in Quebec, and the Gécet, which is a Laval University research group interested in questions of toponymy and geographical terminology.

The objective of the project is to create a permanent bank of toponymic data and to establish an integrated system for utilization and updating of these data. Three steps are involved in the project:

(1) establishment of a toponymic file;

(2) computer analysis of data;

(3) verification and updating of data.

---

* Paper prepared by Mrs. Grenon-Roy, geographer with the secretariat of the Quebec Geographical Commission.

1/ The French version of this paper appears in a special issue of the bulletin "CANOMA", vol. 3, No. 1, which will be distributed at the Conference.

/...

1. *Establishment of a toponymic file*

The years 1971 to 1973 were devoted to the establishment of a place-name file on computer media. The file includes more than 60,000 official place names, which appear in the *Répertoire géographique du Québec*, a 701-page volume published by the Geographical Commission, and in supplements, which have been published in the Quebec Official Gazette.

The computer file is set up in such a way that there are several units of information or "fields" for each toponym. The sample list printout shows the contents (see appendix). The first field shows the date of the name's official publication. In the second column, the eight digits give a sequential code, which was determined mechanically and is used to add or correct toponymic information. The other eight digits appearing on the following line are a language code (English, French, Amerindian, Eskimo, other) and semantic code (historical, dedicatory, descriptive, anthroponymic and other names) for the place name.

The "POSITION" field gives the geographical co-ordinates in degrees and minutes, thus locating the toponym. The following two columns, "NAME" and "ENTITY", give the geographical names, which are arranged in integral alphabetical order of the specific. They consist of two main divisions: administrative entities, that is, those that have been defined by man (counties, cities, towns, villages, and so on), and natural topographic entities, which comprise the range of features of the geographical landscape (coves, streams, bays, mountains, and so on). The toponyms are located according to township, county and area, all of which represent types of cadastral and census divisions. This information appears in the "PLACE" field. The numbers and letters of the last column ("SHEET") refer to map numbers and are explained in greater detail later.

The bank is placed on a medium (a magnetic tape) by means of the APL system. The file is then converted from the initial APL code to the EBCDIC code, which makes it possible to use the French print chain. In addition to capitals, the chain includes accents and lower-case letters, all of which are necessary in writing French and which a conventional computer print chain does not have. This step thus solves the problem of transcribing geographical names with accents.

2. *Computer analysis of data*

Between 1973 and 1975 the operation of the bank made it possible to develop solutions to the original methodological problems concerning the computerization of place names and especially to establish correlations between different parameters such as linguistic and semantic classifications in connection with generics, the distribution of name types according to the various regions, and so on. This analysis, in addition to showing the difficulties inherent in the definition and recording of geographical entities, points up the problems associated

with place-name standardization (spelling, homonymy and so on).

3.   *Data verification and updating*

        Early in 1975 the Geographical Commission announced its inten-
tion to publish a second edition of the *Répertoire des noms géographi-
ques*. This will include the updated and official geographical nomencla-
ture of Quebec. Publication is slated for the spring of 1978.

        At the moment the file comprises all the names of the first
Répertoire and the supplements, that is, all names that have been made
official since the time when the Commission was first created (1912) up
to the end of 1974. It is realized that place names must be reviewed
systematically, since over the years changes occur (different spelling
and usage, more specific location, and so on).

        The methodology that has been developed makes it possible for
geographical names to be easily found on basic cartographic documents and
for the data to be checked and corrected if necessary. The geographical
co-ordinates that already appear opposite each place name in the file
make it possible for the computer to determine map numbers. These cor-
respond to the national mapping system (see the "SHEET" field in the
sample list printout). Proper programming makes it possible to obtain
list printouts of all place names grouped according to the numbers to
which they belong. For example, sheet 22 D/09 gives all place names un-
der that number that have been made official. The names are reviewed
and corrected if necessary. Through coding, the corrections may be re-
corded on lists prepared for that purpose. These are then entered in
the data file by means of an APL console, which is directly connected to
the computer.

        In conjunction with the name verification process, the file is
updated continually. It currently contains more than 75,000 toponyms,
which represent the sum of official geographical nomenclature to date
in Quebec.

        Mechanized operation of this permanent bank of geographical
names has several advantages. In addition to considerably reducing the
number of potential errors,it provides quick access to a large volume of
information. The flexibility of this system also offers many possibili-
ties for compiling names. Various lists have been selected to date, in-
cluding one of generic toponyms, for example, "river" and "creek", one
of inhabited places, and one of names appearing in the first *Répertoire*
arranged, according to language and origin.

        This technique has also made it possible to conduct research
on problems of defining and recording generics. A list of 5,000 names
was selected and compiled, the names being arranged in alphabetical
order of the generic. The list was used to determine systematically the
correspondence or non-correspondence of generics and geographical en-
tities. For example, it was possible to verify the different meanings
of the generic term "coulée", which in certain areas designates a *ra-
vine,* in others a *stream,* and is also sometimes used in the sense of
small channel.

By verifying terms in their toponymic and geographical context, we are able to establish accurate terminology while taking into account the vocabulary currently in use to describe geographical phenomena.

Computer analysis of generics helps to improve the presentation and accuracy of Quebec nomenclature. In the next edition of the *Répertoire des noms géographiques du Québec,* the toponym (generic and specific) and the geographical entity will be separated. This method of recording data meets one of the recommendations that emerged from the second conference on the standardization of geographical names.

By the end of 1977, the name verification, updating and analysis processes will be completed. The permanent bank of toponymic data will thus be used for the new edition of the *Répertoire des noms géographiques.* It will be updated continually and new information (historical data, variants, and so on) will be added as required. Partial or complete lists of names may be produced upon request.

Quebec's experiment in the computerization of place names has been very positive. In addition to solving data recording and storage problems, the project has made it possible to develop a new, more effective methodology to overcome difficulties in the definition, spelling and selection of toponyms. In short, the project has solved the problems inherent in the standardization of geographical names.