



Economic and Social Council

Distr.
LIMITED

E/CONF.85/L.78
3 August 1992

ENGLISH ONLY

SIXTH UNITED NATIONS CONFERENCE ON THE
STANDARDIZATION OF GEOGRAPHICAL NAMES
New York, 25 August-3 September 1992
Item 5 (d) of the provisional agenda*

NATIONAL STANDARDIZATION: ADMINISTRATIVE STRUCTURE OF
NATIONAL NAMES AUTHORITIES

Project: "Inventory and Organization of Geographical Names"

Submitted by Brazil**

* E/CONF.85/1/Rev.1.

** Prepared by Sonia Luiza Terron, Brazilian Institute of Geography
and Statistics, Rio de Janeiro.

Brazil, which represents 46 per cent of the land surface of South America has, according to the 1991 census, a population of 146 million, distributed over 27 states, 4,490 counties (municípios) and about 15,000 populated places. There are eight major hydrological basins, divided into 796 minor drainage-basins, and 1,545 main watercourses, covering all the Brazilian territory. These figures give an idea of the dimensions of the field of geographical names in Brazil.

Brazil is the only country of Latin America whose official language is Portuguese. Differing from other countries of the continent, the idiom is almost uniform, having no dialectal forms, but it is rich in regionalisms and in influences from indigenous languages such as Tupi-guarani. The Tupi-guarani became a generally spoken language in the eighteenth century, with more than 3 million users, outnumbering Portuguese-speakers in the ratio of three to one. The African languages, mainly the Bantu and specifically the Kimbundo dialect, also had great importance, since at the beginning of the nineteenth century 50 per cent of the population was composed of Africans. Italian, Spanish and German immigrants brought great contributions to the Brazilian idiom. All these influences can be found on geographical names, either in generic terms or in the proper names of populated places and physical features.

In the recent past, a specific interest in the standardization of geographical names has been shown in studies, recommendations and resolutions of several governmental agencies.

In 1938 the Brazilian Institute of Geography and Statistics (IBGE) established a Linguistics Geography Sector, later called Geographical Names Sector (1948), which organized and published the first volumes of the Geographical and Toponymical Brazilian Dictionary. In the 1950s, IBGE published a list with more than a thousand toponyms, revised in accordance with the Brazil-Portugal Orthographical Agreement, concluded in 1945. In 1971, the index of the International Map of the World on the Millionth Scale (IMW-BR), revised by the philologist Antenor Nascentes, was published with 36,000 toponyms. Later, the IMW-BR catalogue, already built on a computer, was concluded. The Brazilian Territorial Division, which periodically lists all the cities and villages, had its 1980s edition revised by the Philology and Vocabulary Committee of the Brazilian Academy of Letters, a further valuable contribution of IBGE to geographical names study.

Besides IBGE, other governmental agencies, such as the Brazilian Academy of Letters, the Brazilian Association of Technical Norms, and the National Museum, have worked in the field of geographical names, although not permanently or systematically.

In spite of the efforts made over decades, the resulting projects and studies are isolated. Even at IBGE, despite many attempts, toponymical information remains scattered in various publications, conventional or computerized files and databases, maps and charts. For purposes of standardization, it will be necessary to assemble them on a single unbiased

and up-to-date database. Moreover, the statistical and geoscientific sectors of IBGE's activities would be improved by such a database, and the database would in turn be enhanced by them.

Thus, considering the current status of geographical names enterprises in Brazil, the need for two complementary programmes of simultaneous accomplishment must be reiterated. The first programme might be concerned with systematic collection, treatment and storage of toponyms and correlated data. The second might have regard for geographical names standardization, without which most of the correlated technical and operational activities become unreliable. In addition, the sociocultural damage its absence may cause to the nation by the dissemination of wrong denominations and exotic orthographical forms should be considered.

The project "Inventory and Organization of Geographical Names", the subject of the present document, is being carried out at the Geosciences Directorate according to the above-mentioned first line of action. The project aims at a gradual modular implementation of a geographical names database, reorganizing and complementing the already available toponymic information, and performing systematic collection and treatment of geographical names.

Regarding the legal establishment of a geographic names authority, national action is necessary, and towards that end, IBGE will spare no effort.

Current status of toponymy

The Geosciences Directorate, in view of the need for a specific unit responsible for territorial matters and for establishing a geographical framework for censuses and statistical surveys, has recently created the Territorial Structures Department. This new department, whose activities demand an intensive relationship with various units of IBGE, was considered the appropriate place within which to establish a toponymic section, charged with the organization and maintenance of toponyms.

Also recently, IBGE, as the coordinator of both the statistical and cartographical national systems, has revised the general plan of statistical and geographical information. This plan, which embraces a wide range of programmes and projects, has established the basic guidelines for IBGE's toponymic activities in the project "Inventory and Organization of Geographical Names". All the geoscience programmes profiled by the plan will be presented to the user community during the I National Geosciences Conference, in 1992.

The project has as its purpose the systematic collection, organization and consolidation of the physical and cultural named features on Brazilian territory, in order to shape the onomastic memory of the Brazilian geographic and physiographic features.

The recently created toponymic section, simultaneously with plans for the project, is continuing the collection and computer storing of toponyms from topographical maps series, formerly done by the Cartography Department, in order to publish lists of toponyms from topographic maps. This activity was initiated in 1987, with the compilation of the toponyms from the 46 sheets of the IMW-BR. In mid-1989, compilation from topographic maps on the 1:250,000 scale was started. There are 550 sheets covering Brazil at this scale, and among them, 397 are published by IBGE. Up to now, 136 sheets have been collected, with nearly 40,000 toponyms. However, because of difficulties in handling diacritical marks in the computer's storing process, all files are now being revised and corrected.

Also with regard to geographical names collection, the 1991 demographic census provided a complete and up-to-date set of toponymic data related to populated places, political, statistical and geographical areas. All such information, albeit needing treatment under the toponymic focus, is stored in the Territorial Structures Database and in the 1991 demographic census files.

The Territorial Structures Database was designed to store information on the Brazilian geographic structures of interest, along with associated data, and it is continuously updated. Although it is not as large as the 1991 census files, its structure is more complex because there is not a single hierarchy that suits all levels of the unit's relationship; moreover, there are many references needed for the genealogical research. As a sample of the toponymic potential of this database, some unit levels are quantified below:

States	(27 Federal first-order political units)
Counties	4,490 " <u>municípios</u> ": first-order political and administrative divisions of each state, as of 1 September 1991
Districts	Counties (<u>municípios</u>) may divide their territory into one or more districts, as necessary for administrative purposes (8,770 districts, municipal administrative divisions)
Subdistricts	Administrative areas, boroughs (<u>bairros</u>) and similar units (around 3,500 legal urban planning or administrative areas)
Populated places	Around 15,000 cities, villages and rural settlements
" <u>Favelas</u> " and similar areas	About 3,200 irregular urban settlements (shantytowns)
Special areas	There are about 1,000 environment protection and Indian communities preservation areas

/...

The files are built to provide geographical support for the 1991 demographic census; in particular, the Census Sectors Boundaries File, has also stored a large number of toponyms. The Census Sectors Boundaries File was built to guide the collection of data by providing, for each one of the 163,261 enumerated areas, called Census Sectors, the physical and cultural features that define the boundaries of the area to be canvassed. Aiming at a later application, these feature names were stored between pre-established special characters, defining alphanumerical strings of automated retrieval.

These toponymic data are still unexplored, but it is expected that, as a starting point, nearly 2 million toponyms may contribute to building the toponymic files. The greatest advantage of these data is that geographical locations through geocodes are already assigned.

Another source of geographical names, also derived from the 1991 Demographic Census, are the so-called "collection forms", filled in by the enumerator during data collection. On these forms, the address of every housing unit and business establishment was listed for each enumeration area. Although the forms were devised for data collection control, they indirectly provided a wide array of toponymic information, especially that related to populated places and street names. Once the forms are in a computer-readable form, their treatment and storage in a computerized file, as a prototype of an address master-file, is being planned. Simultaneously, it is intended to evaluate the feasibility of their use as a source for the project.

Project overview

The project "Inventory and Organization of Geographical Names" has been divided into two major activities: Cadastre of Toponyms; and Research into the Origin and Historiography of Toponyms.

The first one refers to the collection, classification and organization of the toponymy that identifies the physical and cultural features on Brazilian territory that are needed for the implementation and maintenance of a geographical names database.

The scope of this activity has led to the planning of carefully measured steps, using all the toponymic information from existing files as a starting point, in order to take advantage of the work done over the past years.

Combining needs and immediate demands with the availability of time and resources, the accomplishment of the Cadastre of Toponyms has become an innovative operational approach. Instead of taking as first step the collection of names and attendant information from current and large-scale topographic maps, the task of completion will be attacked in a second phase. The first phase, then, involves the data organization, collection and treatment for the major named features and areas, as well as the development and implementation of the geographical names database and retrieval system.

/...

Once each geographical name is intricately tied to other data identified with a particular landscape feature, the content of the database might stop there, though always remaining open to the addition of new elements or data categories, and aiming at the possibility of files merging. The elements previously selected include:

Written form of the name

Designation of the kind of geographical feature

Identification and extent of the named entity by geographical coordinates (the coordinates are the entries for the map names files)

Location by state and county (municipio)

Variant names

Other data elements for specific files, such as population, elevation and codes referring to census, postal and fluvial areas.

Owing to operational facilities, toponymic basic files will be first built on a supermicro-environment, while the final depository of geographical names is developed for the mainframe computer.

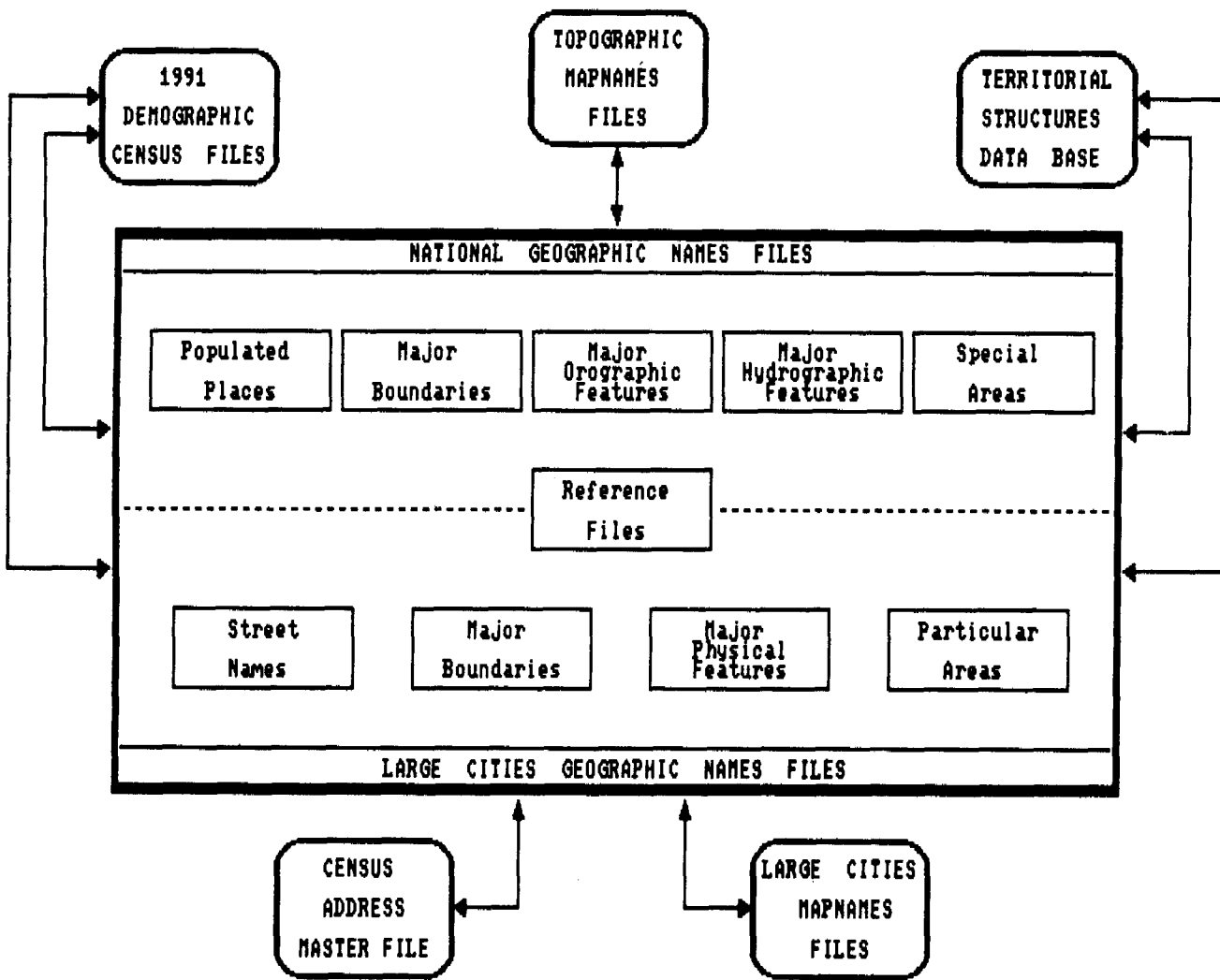
These basic files (figure I) were initially designed to form two groups: the National Geographical Names Files and the Large Cities Geographical Names Files, both involving selected major named features, except for the populated places file and for the street names files, where a relative degree of completeness is desirable.

The National Geographical Names Files were designed to organize a nation-wide selection of toponymical data. The preliminary view includes five basic files: populated places, major boundaries (political, geographical and statistical named areas), major orographic features, major hydrographic features and special areas (legally established areas for the protection of the environment and Indian communities).

The Large Cities Geographical Names group has been limited to the capitals of states, the large cities or metropolitan areas and some other cities of relevant interest. The preliminary view involves four kinds of basic files: street names (street, avenues and important highways), major boundaries (main administrative and planning named urban areas), major physical features (orographic and hydrographic) and particular areas (public parks, recreation areas, historical sites, "favelas" and others).

For supporting the needs of the two file groups, the Reference Files might record the categories of features, their definitions and their cross-references to generic terms. They also might record bibliographies of all source material used for building the basic files.

/...



The diagram (figure I) summarizes only the initial designed structure, actually counting on auxiliary files not shown. Around the core represented are files and databases, already implemented or planned for future development, which have some data intricately tied to those of the geographic names files. For these data, consistency would be continuously verified.

IBGE has a core of geographic support staff in each of its 27 regional offices. As this staff works actively with state, county (municipio) and local officials, data collection and preliminary treatment will be performed by these units under the technical coordination of the Territorial Structures Department.

For building the basic geographical names files, a storing, retrieving, editing and updating system will be developed to operate on the regional offices' super microcomputers. The same system will also support the centralized data consolidation and enhancement process on similar equipment in the Territorial Structures Department.

Simultaneously, a database, with an efficient toponymical data retrieval system, will be developed for the mainframe computer, as the final depository of geographical names. This database will be loaded, step by step, with the consolidated data obtained during each project's phase.

As the first phase of the Cadastre of Toponyms was based upon a selective and thematic approach, the second might deal with the completeness of physical and cultural features. In this phase, collection of names and associated information will be done from current and largest-scale topographic maps, on a state-by-state basis. The operational accomplishment might be similar to the first phase, e.g., on a decentralized acting line.

Although, with the completion of the Cadastre of Toponyms, several needs will be met, other requirements demand research into the origin and historiography of toponyms, whether for analysing the historical, cultural and legal evolution of the nation, or for the implementation of a national geographical names standardization programme.

While the Cadastre of Toponyms concerns itself with geographical names in current use, research into their origins and historiography focuses also on historical and obsolete names, variant applications of a name, and other names once used for currently named places, features and areas. It will also be responsible for the loading of the geographical names database with information about the physical and cultural history of a named feature, with special attention to name origin, as well as information for background research.

Owing to the nature of these activities, priorities and limits will be established to prevent an excessive amount of time spent on research. The first step will deal with the selection and acquisition of sources likely to have the greatest potential for providing historical information. Subsequently, when the basic geographical names files become available, the named features and areas to be researched will be selected. Priority will be

/...

given to populated place-names and geographical names of features related to legal matters concerned with Brazilian territorial division.

In systematizing and performing these activities, the first goal of the project "Inventory and Organization of Geographical Names" will be met. The immediate demands are planned to be accomplished in a relatively short time, and a medium and long-term project chronogram is being established.

Conclusion

While the systematic collection and standardization of geographical names, as well as their publication, are the goals of every country, the ways and means of reaching them may vary. Specific constraints related to the history, language, economy, culture and politics of a nation must be considered when planning any enterprise.

The project "Inventory and Organization of Geographical Names" has met Brazilian constraints, particularly those related to the scope and complexity of IBGE's mission.

While the building of an automated and multipurpose geographic names information system, meeting national needs as accurately and completely as possible, is desirable, there are, however, some basic steps to be taken before the challenge can be met. The toponymical data already collected must be organized and treated to a pre-established level of precision; operational processes and methods must be revised and adapted, regional cores of staff established able to work in the field of geographic names.

Thus, in the first phase of the project, past attempts, current needs and future requirements must be carefully measured against financial resources and the possibilities presented by new technology. The second phase should be one of enhancement if the preliminary results prove the feasibility of a long-term project.

Finally, beyond all technical purposes, aims and intentions, the accomplishment of the project "Inventory and Organization of Geographical Names" attests once more IBGE's tradition of making an abiding contribution to the sociocultural development of the nation.
