



Economic and Social Council

Distr.
LIMITED

E/CONF.74/L.71
24 August 1982

ENGLISH ONLY

Fourth United Nations Conference on the
Standardization of Geographical Names
Geneva, 24 August to 14 September 1982
Item 9 of the provisional agenda*

AUTOMATED DATA PROCESSING

A Program to Develop Automated Processing of Foreign Names
at the Defense Mapping Agency**

Abstract of Paper presented by the United States of America

The management of geographic names and related information by the Defense Mapping Agency to meet internal mapping and charting requirements and to support standardization programs of the U.S. Board on Geographic Names has, over the past several years, increasingly demonstrated the need to incorporate techniques of automated data processing. Accordingly, DMA has instituted a programme that will be able to deal with a multitude of toponymic problems related to standardization and that will provide access to names information to meet various requirements. The first stage of the program calls for the installation of the Geographic Names Input Station in the latter part of 1982.

* E/CONF.74/1

** Prepared by Robert F. Augustine (Defense Mapping Agency), Douglas Caldwell (U.S. Army Engineer Topographic Laboratories), and Diane E. Strife (Illinois Institute of Technology, Research Institute).

The Defense Mapping Agency (DMA) collects, evaluates, and maintains names information of foreign places, as well as Antarctic, undersea and extra-terrestrial features, to meet internal needs for mapping and charting and to support efforts of the U. S. Board on Geographic Names to standardize names for official U. S. purposes. As a result, there is a continually expanding file which serves as a base to support the use and application of geographic names throughout the United States government and for numerous professional organizations and the public as well.

Today, the foreign place-name file consists of names for 2.5 million features with BGN-approved and variant names stored on 4.5 million index cards. These cards are currently organized alphabetically by country, but since 1979, DMA has been sponsoring the development of a prototype Geographic Names Input Station (GNIS). This development is being carried out under a contract to the Illinois Institute of Technology Research Institute (IITRI). The U. S. Army Engineer Topographic Laboratories research problems associated with the digital processing and representation of foreign characters, symbols and diacritics, as well as immediate gazetteer-production requirements. Installation and testing of the GNIS prototype is scheduled for the latter part of 1982.

The processing of foreign-names text information requires special techniques for data entry, storage and display. More than half of the 116 languages or transliterated forms that may be treated at DMA require an extended alphabet with sets of marks not found in the 26 letters of the English alphabet. Marks used in extended Roman alphabets are divided into three categories: diacritics, special characters, and special symbols. A diacritic is a mark that may be placed above or below an alphabetic character, but which does not alter the basic character, for example, š. A special character is created when a mark is superimposed upon or drawn through another character, for example, đ. A special symbol is a non-Roman alphabetic mark that may be used alone, for example, ø. DMA requires over 70 of these three types of marks to represent, as accurately as possible, the names of places and features in countries throughout the world.

The International Standardization Organization (ISO) has adopted or proposed two standards: ISO 5426 "Extension of the Latin Alphabet Coded Character Set for Bibliographic Information (1980)" and ISO/GIS 6937/2, "Coded Character Sets for Text Communication: Part 2: Latin Alphabetic and Non-Alphabetic Graphic Characters." These two standards do not agree completely, and fail to include all the Roman alphabet marks, characters, and symbols required by DMA. Since no inclusive standard now exists, tables must be provided to convert from one format to another. Attempts to standardize data storage formats for interchange among users are still in their early stages.

An intelligent terminal with a programmable keyboard was selected that could achieve the required softcopy display as well as perform the data entry and data retrieval processes. A printer with an 8 by 16 dot matrix display was selected to provide for local and interim printing operations. New characters may be added by way of a matrix editor as character display requirements change or increase.

The major computational and data storage element of this prototype system consists of a 10 megabyte microprocessor. A 9-track magnetic tape drive, two disk drives, and the microprocessor are located in a single unit. An associated digitizer provides raw data to the system in order to permit computation of feature locations.

The prototype GNIS will be upgraded following installation to include greater storage capacity and more efficient interactive capability. It will serve as a possible foundation for an all-digital Foreign Place Names Information System that will make use of knowledge acquired during the development of the prototype GNIS. If the GNIS proves satisfactory, a feasibility test is planned to assess its capability to perform some of the pre-compilation names application processing required for map production.

In addition to integrating and consolidating names production within DMA, the system will provide a greater capability to process toponymic data and disseminate names information to users in both the public and private sectors.