

AGENDA ITEM 17

Automatic data processing

COMPUTER PROCESSING OF GEOGRAPHICAL NAMES*

Report presented by the Federal Republic of Germany

In the Federal Republic of Germany studies on the possibility of processing geographical names by computer were initiated in 1969. In the first period of these studies special attention was given to the requirements of cartographers. Various categories of geographical names were formed, problems of data capture were considered, and the amount of peripheral information required was subjected to research. A first approach was reached at the end of 1970. Early in 1971 a circular dealing with the problems and containing recommendations was distributed to all official agencies in the Federal Republic of Germany for further comments. The answers were discussed in meetings during March and April 1971.

At this stage, the requirements of cartographers were defined. The work so far had been executed by the Institut für Angewandte Geodäsie at Frankfurt am Main, in co-operation with the Institut für Landeskunde, Bonn-Bad Godesberg.

A second period started in June 1971 when the problem was discussed with mathematicians from the Gesellschaft

* The original text of this paper was contained in document E/CONF.61/L.19.

für Mathematik und Datenverarbeitung, Bonn-Birlinghoven. This organization is concerned with computer problems related to federal agencies. It completed an elaborate analysis of the system. Based on this analysis, a third period will begin early in 1972 — the programming of the various phases of data processing for further retrieval. In a fourth period the programmes will be tested. For this purpose a special area around Frankfurt am Main has been selected and the data capture has already been completed.

A report on the subject and the results reached so far will be given at a session of commission III during the International Cartographic Association meeting on "An Integrated System of Processing Geographical Names" at Ottawa in August 1972. The word "integrated" means not only that names are stored for the preparation of gazetteers or alphabetical listings, but also that the system will permit the retrieval of such information as listings based on the number of inhabitants, or locations either in geographical or Universal Transverse Mercator co-ordinates or in the numbering system of the national map series. Listings based on attributes such as the length of rivers, administrative status, heights above sea level etc. will also be possible.

THE APPLICATION OF AUTOMATION TO GEOGRAPHICAL NAMES*

Report presented by the United States of America

A review of automation in the field of work on geographical names reveals that progress has been made but that there still remains much to be done. There have been significant developments in the design of machine systems which are directed to specific name-handling, but most are limited to specialized operations and lack compatibility with general name data-bank concepts.

The development of name data banks by Governments, scientific institutions, business houses, and other organizations has been on the upsurge for many years. Requests for copies of or access to the United States Board on Geographic Names gazetteer tapes are on the increase. The ability to select names according to

* The original text of this paper was contained in document E/CONF.61/L.48.

such criteria as political division, type of feature and map sheet makes such stored data a very valuable resource.

Some progress in the design of systems which will store and print out all desired diacritical marks and special characters has been made. As far as is known, the systems which have this capability are designed for long-range programmes involving a large volume of specialized work — for example, the preparation of library reference cards. All such systems appear to have deficiencies such as difficult and expensive input, incompatibility with other computer equipment, loss of data-manipulation capability or slow print-out speed.

It is generally agreed that input by optical scanning devices is rapid and accurate, and that it allows for the storage of diacritical marks and special characters. However, as with other systems, the cost is high and it is

difficult to document the savings. As with most name files, and for any system, the cost of converting the files to a machine-readable form can be prohibitive.

Further research is necessary in the development of data storage systems which retain the capacity of processing data and interfacing directly with automatic mapping devices. No such systems are known to exist at the present time, although it is evident that the elements exist and the theory of design is well understood. To justify investment in such a system it would have to have the capability of providing names for maps and consequently a storage capacity for scores of millions of names. At this time the selection and application of names is one of the major cost- and time-consuming factors in cartography. If automation of the other aspects of compilation outruns the automation of the naming process, a serious bottleneck will result

For specialized Board on Geographic Names gazetteers such as those for the Republic of Viet-Nam and conventional names, the Board has experimented with the use of computerized photo-typesetting devices. The quality of the product has been excellent; however, the lack of compatibility of the data storage for other applications of name output is another demonstration of the problem inherent in system design. Clearly the computer industry has arrived at a stage where more attention should be paid to industry-wide standardization.

It is recommended that an increased effort should be directed to the expanded application of automation to geographical names, so as to fulfil the needs of a wide range of research and reference users, such as libraries, publishers, and educational and research institutions; and to satisfy the requirements of map production.