Formulation of Geospatial Data Standards - An Inevitable Requirement

P.N. KOUL

ABSTRACT

The digital geographic information (DGI) has provided appreciable flexibility to the users of the data for scientific planning of developmental activities, their implementation and monitoring. The digital geographic information in the form of Digital Cartographic databases possesses a wealth of information for various users of the data. It is convenient to exploit the data for varied applications such as, but not limited to planning, scientistic activities, defence and various other applications. The digital data enables its useful exploitation in respect of integration and value addition which otherwise is not possible in conventional mapping environment.

In such an activity it is important to efficiently convert the analogue map data to digital environment with respect to a pre-defined data model structure to enable subsequent utility of the data by different organizations/ agencies for various GIS applications. An optimum data model structure can withstand pressures of efficient data exchange, which is a prime necessity for data integration of varied data sets. Development of data exchange format and formulation of standards for data exchange are required to enable users to have a transparent flow of data between various GIS environment. The standards should be technically sustainable and viable in the long run. Any ambiguity in formulation of the data standards may lead to a situation where it becomes impossible to implement them.

The existing standards may be adopted and altered/ amended to accommodate various data types which may include cartographic data in vector form, satellite imagery and aerial photographic data in raster form and non-spatial data standards should be formulated in such a way that the approach is simple, unambiguous and exploits the existing standards to the extent possible. The exercise should be aimed for an optimum and an efficient data exchange. The process of standardization of the geo-spatial data leads to uniformity in the generation of cartographic data and the end products of satellite images. It is an important step towards achieving 'interpretability'.

It is very important to identify that the data standards, in respect of a data exchange format, are formulated well in time. The timely formulation of geospatial data standards is an important factor in the whole process and it may not practically serve much purpose if its availability and implementation is delayed. The standards have to be anticipatory and not reactive.

This paper covers concepts related to the standardization of geo-spatial data. The paper also highlights various issues related to geo-spatial data standards, which should be open ended, and flexible to accommodate any future modifications or change.

INTRODUCTION

Prevailing scenario of information explosion has led to information technology environment. As a result appreciable pressure has been built upon data producers to transform spatial data from analogue (paper) maps to softcopy maps and non-spatial data from table domain to spatial domain. This is being efficiently facilitated by advanced computer systems comprising of sophisticated hardware and software modules, which are supported by powerful operating systems. The transformation and sophistication in operating systems from VAX to UNIX and then to Windows has played a pivotal role in giving tremendous boost to the generation of geo-spatial data.

The situation has warranted creation of user-friendly computer environment, which could facilitate an efficient decision support system. For any such system it becomes necessary that various data sets, pertaining to a common location, are merged since combined data sets can provide a decision support system. The decision support system may be required to facilitate :

- Asset management
- Rail/road/naval/airline transportation
- Precision farming
- Emergency response
- Forestry
- Health and public safety
- Disaster management and launching of rescue missions

- Environmental monitoring
- Defence needs which may include military command and control, fleet management and air-traffic control

The requirement of scientific and efficient decision support system is an ardent need and has a global application, may it be a developed, developing or an underdeveloped country. To support such a system there is a strong need to integrate varied data sets generated by various source agencies. The success lies not only in how efficiently the integration of data is designed but also how fast/timely a country is in carrying out the data integration so as to strengthen the system. Any serious delays may come in the way of planning and execution of efficient decisions in respect of the national developmental activities and/or disaster management.

In this paper the author has attempted to bring out the importance of formulating standards of digital cartographic databases at national level. The author has been coordinator, and part of Study Team, of the subcommittee for formulation of National Spatial Database Standards which is a part of Natural National Resources and Management System committee of Cartography and Mapping (NNRMS-SC$C). The Study Team has prepared a draft document on standards/ format, which supports integration of varied datasets and is based on harmonisation of two formats, which are being extensively used in India at present.

BEYOND MAPPING

The latest trends in mapping have forced a transition from traditional mapping to 'Mape-matics', that is, Maps as data, maps as numbers, which is achieved by creation of digital cartographic databases (DCDBs). The data in digital form provides flexibility to relate and integrate data originating from various diciplines such as satellite imagery, cartography, geology, geomorphology, forestry, land use, census etc. The multifold advantages facilitated by such a scheme of integration of varied data has lately drawn the attention of cartographers, space scientists, engineers, geologists, foresters, planners, computer scientists and other professionals. The valuable cross interaction between the above-mentioned community of scientists has led to **Geomatics** which primarily supports integrated land resource management - a prime requirement for sustainable growth and infrastructure development. Geomatics encompasses the varied data sets, originally generated in various scientific disciplines, and deals with their acquisition, integration, organisation, management, infrastructure, standardisation, archival and dissemination.

Geomatics not only encompasses management of standard databases but also includes data modeling while integrating spatial and non-spatial data related to topography, geology, hydrography, vegetation, land use, census, wild life etc. The situation, leading to complex modeling procedures, has compelled the data producers and data users to **look beyond Mapping**. As a result the demarcation between the cartographers and data users has been rendered blurred in the context of the prevailing information explosion. It has now become necessary for the cartographers and scientists from various other fields to look beyond their domains, closely interact with each other and the data users so as to understand each other's requirement. That appears to be the only way to generate meaningful digital cartographic databases for accurate and successful execution of developmental activities at the grass root level.

REQUIREMENT OF STANDARDS

For an efficient data integration and data sharing it is necessary to develop standards, execute their acceptance and implementation. In the context of integration of varied data sets the proposition is decidedly complex but needs to be seen through. Appreciable activity is going on internationally, that it makes sense to examine and get educated about as to what is going on in the standards field. The formulation/ implementation of standards is inevitable for the integration of geospatial applications and main stream database technology.

The sources of varied data sets are different and so are their standards and accuracy specifications mainly because the source organisation have worked in isolation. In India, about a decade back, it became apparent and necessary to develop data exchange format and an exchange format called DVD (Digital Data Exchange)-1 was developed and is in vogue today. This format was highly relevant at the time of formulation but it may not meet the aspirations of mapping community in the prevailing information technology environment, which decidedly warrants integration of vector/raster data, satellite imagery data, thematic data and nonspatial data sets.

As of now a wide disparity of standards exists between various data sets mainly because of two reasons. Firstly, the source organisation have generally worked in isolation and secondly, there may be valid differences because of different needs in which case different standards may be desirable and acceptable as long as the standards are harmonised [1]. For example a standard designed for topographic database may not need to carry the same data element as one designed to represent a hydrographic element but what is important is that both should have the same data structure and be compatible when they overlap.

The following objectives need to be kept in mind while formulation of standards towards open sharing of geo-informational resources :

a. Development of appropriate standards towards modeling and moving data in general, that is, ability to deal with both geo-spatial and traditional information.

b. Enable handling of any kind of geo-spatial data including the data with extensive attribute description and those with geometry defined by vector or raster structures in two/three dimensions.

INTEROPERABILITY

As we move forward towards open sharing of geo-spatial information resources, it becomes pertinent to keep in focus versatility in data exchange i.e., how effectively geo-spatial information residing in one system is being utilised by other system. This leads to the concept of interoperability, which means the extent to which users, software and data can move between different computer environments. An efficient interoperability has to support travel of display and analytical functions generated in a particular system to the requester's destination ensuring compatibility, a Java applet application. [2]

SALIENT SPATIAL DATABASE STANDARDS

The list of such standards may be elaborated but some salient standards are as under :

a. Projection system

b. Scale factor for various levels of databases

c. Co-ordinate system

d. Accuracy specifications for raster and vector data with respect to scale,

e. Minimum spatial unit (raster/vector)

f. Data format which provides in-roads for raster/ vector data

g. Likely use (level of hierarchy)

h. Metadata standards

i. Data dictionaries

j. Lineage

Map projection

Map projection system is basically a mathematical model that transforms the location of features from the earth's to a map. Some projections preserve shape, others preserve accuracy of area, distance or direction. Any such transformation projection distorts some parameter of earth's surface-be it distance, area, shape or direction. A suitable projection needs to be selected vis-a-vis the scale factor and

likely purpose such as macro level planning, regional planning, planning for rural and urban development, water shed development, planning for urban infrastructure, engineering design etc.

Metadata Standards

Formulation of metadata standards for cartographic databases has an appropriate relevance in the context of data integration. Metadata means data about geo-information such as, but not limited to, documentation of data containing details about data source organization(s), authorised contact persons, accuracy reports, reliability of data, data model information, spatial referencing including projections and co-ordinate systems. In short, metadata may be defined as a catalogue which serves as a authoritative reference in respect of geo data and may involve one or more data producing agencies.

Data Dictionary

It contains a catalogue of all the attributes for a data set along with all the constraints placed upon attribute values during the data definition phase. It may contain information about files, records, attributes rather than just data.

Lineage

It is a narrative of source materials and procedures used to produce the product as well as a record of updates and changes. This is a component of data quality.

WORLD SCENARIO

Advanced countries have formulated standards for digital cartographic databases. So critical is the geographic information., its standardisation and portability from one system to another that in USA President Clinton issued executive order to establish National Spatial Data Infrastructure (NSDI) and commissioned the Federal Geographic Data Committee (FGDC) to document it, to give it operational form and to provide it a management structure designed to help it grow and become a focus for combined efforts of both public and private sectors concerned in any way with development and use of geo-spatial information. Already significant strides have been made in such countries towards standardisation and data integration, which includes description of content of geo-spatial data collection (metadata). However, progress in respect of interoperability has been comparatively slow mainly because the problem is complex. Traditionally GIS packages, whether commercial or public, have created highly structured and architecturally closed operational environments with proprietary spatial database designs, which tend to make these packages operationally monolithic. The serious draw back comes in the way or their efficient use [1]

However, substantial work is being done at present to enable transparent barrier free access to heterogeneous geo-spatial data sets in a distributed environment by using standardised interfaces for the generation of queries in real time followed by constant retrieval of query derived data sets to the users destination.

As regards India, we have a data exchange format Digital Vector Data (DVD)-1 in vogue for the last one decade. The necessity to develop standards at national level is now being felt intensely than ever before. As mentioned, the Study Team, under Natural National Resources and Management System committee on Cartography and Mapping (NNRMS-SC$C), has designed standards/format in such a way that all types of data (vector, satellite imagery data, other raster or DEM data) can be integrated under the same roof. The study group has basically attempted to harmonise DVD-3 data of Survey of India and NRIS (National Natural Resources Information System) data format of Department of Space and developed a fresh format. The draft document provides in-roads for the following type of data :

a. Base category wise data (vector data)

b. Thematic data (vector data)

c. Raster multi-band imagery (band wise)

d. Single band raster image

e. Single band coded raster image

f. DEM data

The format once accepted and adopted will facilitate data integration desired under the prevailing environment.

CONCLUSION

There is a strong need to create an environment, which facilitates transparent data handling and accessibility. This can be achieved through harmonisation of varied data sources and their representation in unambiguous and clear manner. The requirement warrants design of a set of standards since it is the standard, which basically make the data portable. In this paper the author has elaborated various aspects of an efficient data standard/format required at national level which can provide in-roads to various raster/vector data sets and thus facilitate data integration. This may be a leading step towards interoperability of digital data, which poses one of the major cartographic challenges at the dawn of the new millennium.

REFERENCES

Valerie Hume - Geomatics Volume 52, 1998 Edition, Canadian Institute of Geomatics

Trip Report - http//www.aft.pfc.forestry.ca/features/spatial-data-report.html

Spatial Database Standards for National Cartographic Databases (A Report of NNRMS Study Team) by Lt Col P N Koul (Survey of India) and Sh. R.K. Goel (Space Application Centre)