N78-14560

REMOTE SENSING AND GEOGRAPHICALLY BASED INFORMATION SYSTEMS*

Richard C. Cicone**
Environmental Research Institute of Michigan
Ann Arbor, Michigan

ABSTRACT

The incorporation of remotely sensed digital data in a computer-based information system is seen to be equivalent to the incorporation of any other spatially oriented layer of data. The growing interest in such systems indicates a need to develop a generalized geographically oriented data base management system that could be made commercially available for a wide range of applications. This paper reviews some concepts that distinguish geographic information systems and proposes a simple model which can serve as a conceptual framework for the design of a generalized geographic information system.

1. INTRODUCTION

Progress in the field of earth resources analysis, through the use of remotely sensed data, is resulting in the application of the developed analytical tools to meet the needs of a broad community of users. As an outgrowth, one sees the development of extensive geographically oriented data bases populated not only by scanner data, but also by a wide variety of associated data. A centralized data bank is envisioned as a tool to fulfill user information needs. The desire to make use of the information content of these data carries the responsibility to determine an effective means by which the data can be efficiently managed. It becomes imperative, then, to determine an environment, or information system, within which the data can be stored, retrieved, manipulated and displayed.

Information systems can be divided primarily into two categories: (1) object oriented, and (2) spatially or geographically oriented [1]. Object oriented systems include scientific or statistically oriented information systems and management information systems. In a sense, a spatial information system is nothing more than an object oriented system with an added attribute -- locational identifiers. It is just this characteristic, however, that adds to the complexity of storing, retrieving, and manipulating these data. For this reason one would wish to distinguish geographic information systems from object oriented systems.

Much effort has been expended in developing generalized data base management systems (DBMS) which are commercially available and well suited to the employment of object oriented data bases [2]. An advantage of commercially available DBMS is that a basic set of data management programs is made available that provides a starting point for a variety of applications. New users in the market for geographically oriented computer-based information systems (GIS) discover that a basic set of software designed for spatial systems is not commercially available.

*This work was supported by the National Aeronautics & Space Administration through their Earth Observations Division, Johnson Space Center, Houston, Texas, under Contract NAS9-14988.

**The author is a member of the Information Systems & Analysis Department, Infrared & Optics Division, ERIM, Ann Arbor, Michigan.

The user will respond by developing a local system that does not have general applicability, either by adapting a commercially available DBMS [3,4], or by developing independent in-house capabilities [5,6]. Each time such a system is developed effort is duplicated since the user has not been able to take advantage of the fact that there is a commonality to the data sets and data processing algorithms required in a basic computer-based geographic information system. It is precisely this commonality that makes the construction of a generalized spatial data base management system feasible.

The remainder of this paper simply proposes a way of thinking about a geographic information system that prefaces the construction of a generalized system. The paper is organized into four parts: (1) a general overview of geographic information systems that incorporate remotely sensed data, (2) idealized GIS standards, (3) the interrelationship of information system components, and (4) a GIS design model.

2. GENERAL OVERVIEW

Some general observations must precede the more technical discussion of the GIS design concepts. Spatially oriented data sets may be regarded as layers of information. The terminology stems from the common process of overlaying or intersecting layers of spatial data to extract information in regard to the co-occurrence of events. Remotely sensed and associated data can be viewed as different layers of spatial data, inherently grid, linear, or point source in structure. The incorporation of these data in a GIS data base is equivalent to the incorporation of any other spatially oriented layer of data. Assuming that each data layer has associated with it a unique storage structure, the display and analysis of these data require an interface between layers of both linear and grid structure, and of varying resolution sizes. Often it has been the restriction of data to a specific storage structure that has limited a system's scope of applicability. The development of a GIS that incorporates remotely sensed data will require special processing functions, which in turn will result in the need to manipulate non-spatial or object-oriented data. An example is the extraction of cluster signatures, i.e., sets of mean vectors and covariance matrices representing the statistical composition of a remotely sensed data layer. These statistics are not spatially oriented, yet once computed are integral elements of the data base.

The scenario, then, requires no restrictions limiting the data types (they may be spatial or not), data structures (they may be grid cells or not), or processing functions (the design must permit upwards compatible, modular growth so as not to limit the scope of applicability). What drives the system then? Simply answered -- the user. A generalized geographic information system must permit the user of the system multiple views of the data that are independent of the data storage structures. This basic axiom drives the concept of a generalized GIS.

3. ESTABLISHING SYSTEM STANDARDS

Designing a geographic information system first entails establishing system standards. The flexibility embodied in the basic axiom should also be thought of as first in importance among standards.

1. The user of the computer-based geographic information system must be permitted multiple views of the data independent of storage characteristics.

This simply means that although a data layer may be stored in polygonal format, the user may access it as if it were in grid cells. This imposes a responsibility upon the system to provide the appropriate interfaces. Other important standards include:

2. Applicability of the entire system as an organizational resource belonging to no one user or one application.
3. Applicability of the data to the users' needs.

4. Applicability of the data processing functions to the users' needs.
5. Efficiency of data storage, retrieval, and processing.
6. Provision of a 'user friendly' interactive and/or batch environment.
7. Fulfillment of cross-functional information requirements.
8. Fulfillment of cross-level information requirements.
9. Practicability within computer facility support.

Many of these standards are drawn directly from objectives which commercially available data base management systems aim to achieve [7]. It will become clear that the DBMS environment is required. The spatial nature of the data places special demands, however, on the data management system. The next step is to look at the various components of the information system in this context to establish distinguishing operational features.

4. GIS COMPONENTS

The task here will be to define "information system". The discussion will be in the context of a geographically oriented system. The scope of this subject is broader than the aspects of data base or of data management alone. The entire range of system components affects the manner by which spatial data are managed. For example, the processing of spatial data produces a need for requirements that place special demands on the data management system.

Figure 1 illustrates an interrelationship between the six components of an information system [8]. Two of these, "data specification" and "acquisition" pertain to the process of data creation. "Data management" and "data base" components pertain to the maintenance and retrieval of data in a computer-based environment. The information extraction processes are carried out at the "data processing" and "dissemination" stages. Each component will be discussed individually and, in some cases, interrelationships with other components will be discussed.

Data Specification

Data specification involves four basic processes [9]:

1. The establishment of specific data needs. These data needs may span a variety of data types including: land, environment, population, and administration. The selection of specific data types would be based on the system application.

2. The establishment of cross-level data needs, as well as cross-functional data needs.

3. Categorization of data types and interrelationships by topic and feature.

4. Determination of data update standards, based on the rate of data change and data growth through processing.

A unique characteristic of spatially oriented data is encountered once one initiates the process of data specifications. That characteristic is the "layered" nature of geographic data. That is, every location on the ground can have associated with it a wide variety of characteristics. One system employed at the Environmental Research Institute of Michigan defines 23 variables; like land use, soil and topography, to characterize a location of approximately one hectare in size. The same coordinate references many layers of information [10].

Data Acquisition

Probably the most awesome task confronting the implementation of any information system is the gathering of data in a computer-compatible format. This process takes on special problems when the data are geographically oriented. The tasks at hand include [9]:

1. Establishing data sources, i.e., determining which data are currently available and which data must be measured.

2. Establishing strategies for sampling.

3. Determining data computer compatibility. This may require a complicated digitization process to a standardized coordinate referencing system.

Here we are confronted with a second important spatial data characteristic. The volume of data required for even the small applications may be enormous. Spatially oriented data can exist in any one of three forms; point source data (climatological data), linear data (a street network), and areal data (thematic maps over contiguous regions). These data can be dimensioned not only by their spatial resolution, but also by their temporal resolution, i.e., rate of change as reflected in the frequency of measurement. As a familiar example, remote sensor data gathered by Landsat are segmented into frames. Each frame is 100 nautical miles on a side containing over 28,000,000 bytes of data. These data are measured every 18 days. Approximately 20 data sets are gathered over a given site in a year, representing over 0.5 billion bytes of data. Associated with these data, one may require a variety of other information: elevation from sea level at a point or land use category.

Data Base

By "data base" we mean the collection of pieces of quantitative and qualitative information, in a retrievable format, that measures or describes features of interest. The term "data base" is often misused, as is "data bank", for the information system itself. The information content of each piece of information or datum is three dimensional [8]: (1) thematic, what is being measured, (2) spatial, where it is being measured, and (3) temporal, when it is measured. Each datum can function either as an analytical variable (i.e., a measurement that can take on any numerical value over a continuous or interval scale) or a categorical variable (i.e., a descriptor or attribute that can take on a limited number of values on a discrete or nominal scale) or both. For example, multispectral scanner data are analytical, soil type data are categorical, and topographic information could be either or both.

The logical design of a spatially oriented data base includes the determination of the data layers or attributes, data interrelationships, and due to the potential volume of data, a data segmentation strategy.

Physical storage characteristics of geographically oriented data include two basic types: (1) regular cells or grid encoded data and (2) irregular cells or linearly encoded data, though each can be encoded in a variety of ways [11]. Traditionally, systems are of one type or the other. The fact that should not be compromised, however, is that certain layers of information fall naturally into one storage type or the other. The optimum system can manage both forms of data. Let us discuss the concept of the spatial data structure a little more fully.

Data structure refers to the manner in which data sets are arranged and the formats in which data elements are stored. The data structure enters at a variety of levels. Data can be found in each of the following structures:

1. Raw Data Structure: The form in which data are acquired, eg., soil map or MSS CCT format.

2. Computer-external Data Storage Structure: The computer-compatible format in which the data base resides outside of the computer processing unit.

3. The Computer-internal Data Storage Structure: The format in which the data reside within the computer processing system.

The external storage structure is the format from which data are initially retrieved before processing and by which data can be disseminated to various users. As mentioned, the two basic geographic information system external data structure organizations are line encoding and cell encoding.

In line encoding, spatial features are defined using nodes and connecting line segments. Point form data are described using only nodes; linear form data consist of nodes and connecting line segments; and areal data consist of nodes and line segments forming closed regions, i.e., polygons. Polygons need not be contiguous nor completely cover the scene of interest. The organization of the encoded nodes and line segments is generally handled through lists. Linear encoding techniques include: (1) location lists [12], (2) point-dictionaries [12], (3) DIME files [13], and (4) chain/node encoding [14]. Line encoding offers the most general type of geographic data representation [15] and is particularly advantageous in terms of computer storage requirements in describing: (1) large uniform regions of data such as state or county boundaries, i.e., regions that are large in area in comparison with the basic data cell size, (2) regions of irregular shape, and (3) features that are characteristically linear.

Cell encoding is a special form of line encoding of areal data. Cells are rectangular polygons and are usually square. Because of the regularity of the shapes, and since they generally cover an entire scene, cells can be stored as an array, rather than in a list. This form of encoding can permit an efficient way of retrieving certain kinds of data since access is done through coordinate referencing, i.e., indexing into the array, rather than searching through a list.

Grid structures include three encoding techniques:

1. Sequential. Data values are entered into cell after cell along rows or columns.

2. Compact Sequential. Repeating data values are not stored for each cell, but stored along with a length attribute.

3. Complete Coding. Each data value has a locational vector associated with it.

A third data storage structure type that is not always considered integral to the geographic data base are data that are not necessarily geographically oriented, but list oriented. Yet these data are so integrally related to the processing of geographic data that they should not be separated from it. Most systems manage these data in associated flat files. As previously mentioned, these data could include statistical characterizations of a particular layer of geographic data. Tables of aggregated statistical information that correspond to features of interest to the user of the system form another integral part of the geographic data base. These data types lend themselves to a DBMS data organization more readily than the geographically oriented data. However, due to their analytical nature, the data values fall in a continuum and are thereby not easily retrieved using inverted lists, which function best in an environment of discrete data values. A relational environment is more appropriate [16].

Data Management

We have seen, so far, that geographic systems are characterized by: (1) the spatial orientation of the data, (2) the layered characteristics of the data, (3) the potential volume of data, and (4) the variety of optimal data structures. These characteristics create a few problems for the subsystem responsible for the management of these data. Let us speak of data management in the CODASYL sense [17]. That is to say, the data base management subsystem is aware of a logical data structuring, and is responsible for the retrieval of these data in a manner that assures data integrity and application programs that can remain independent of data storage.

The retrieval demands of a GIS preclude the direct employment of a commercially available DBMS. Retrieval functions include retrieval based on nominal data characteristics, coordinal data designation and relational techniques. Furthermore, the analytical and continuous nature of certain data causes the inverted list approach to the retrieval of the data to become inefficient. The commercially available DBMS encounters other obstacles. Whereas spatial data may be stored as a polygon, a structure that is not supported by commercial DBMS's, the user may require the data in grid format requiring the DBMS to invoke a data base procedure called, say, point-in-polygon, which will convert the irregularly shaped data into a grid cell matrix (responsive to the basic design axiom), again not supported by available DBMS's. Furthermore, the user's point of view may require a data resolution different from that available, resulting in a complicated resampling of these data. The result of these inadequacies of commercially available systems in the development of various spatial information systems has been the local implementation of data management systems. Often features that make the CODASYL systems of value were lost, especially program/data base physical storage independence. It should be re-emphasized that, though commercially available systems fail to meet the data management needs of a geographic system, the characteristics of such systems can be designed into a spatial data management system.

Data Processing

The processing of spatial data quite often is analytical in nature and generates new layers of spatial information that must be maintained by the data base management subsystem. Data base growth, therefore, comes not only from the specification and encoding of raw data types, but also from the processing of encoded data. Whereas a data management system attempts to preserve data base and application program independence, the nature of the application programs employed may affect the data supported in the data base.

Processing of spatial data falls logically into three steps: preprocessing, processing, and post processing.

The intent of data preprocessing is not to extract information from the encoded data, but to modify the data in such a manner as to make the extraction of information more feasible or efficient.

Data preprocessing generally deals with such items as transforming raw data into some standard coordinate referencing system like Universal Transverse Mercator. This activity is termed geometric correction. A second preprocessing activity might involve the analytical transformation of data. For example, many forms of spatial data, in particular those measured using remote sensors, are multivariate in nature. A principal component analysis of the data may warrant a transformation to compress the data into fewer dimensions with axes oriented in the direction of the principal components. In effect, a new layer of data is produced.

Data processing of the spatial data pertains to the information extraction process. This involves three basic operations:

1. Feature extraction in response to a user's query within a layer of data, eg., a discriminant analysis to determine physical characteristics of the data.

2. Feature extraction between layers of data; commonly this is accomplished through co-occurrence analysis or overlay processing; here layers of data are "intersected" to determine geographic regions that satisfy a user specified query that may be algebraic in nature.

3. Inference modeling; the information content of the data is used in conjunction with a mathematical model to project changes that may occur, eg., an ecological system over a period of time under a set of circumstances.

Data post-processing pertains to the aggregation of information extracted in the data processing stage. Statistics are gathered into a format compatible to some report or tabular display. Oftentimes the sequence of data processing efforts is an interactive one. The post-processing of the data may warrant another processing approach to extract new or different forms of information.

Dissemination

Dissemination pertains to the delivery and maintenance of data and information extracted from data to the users of the system. At a local level, data are disseminated to the users through some sort of hard or softcopy interface in the form of a map or a report. That report may be generated as a response to a query language interface with the system user. It may take the form of a table, chart or graph. The standard vehicles designed to transport these data would include a line printer, a table plotter, or a video terminal with associated hardcopy unit.

Dissemination of spatial data and information both begins and ends the cycle of a geographic information system. The information extracted may be the computation of a new layer of data which is in turn re-entered into the system for further processing and analysis, or may be a final report describing the results of data processing and analysis.

5. COMPUTER-BASED INFORMATION SYSTEM MODEL

The preceding discussion addressed the concept of a spatial information system. The intent was to indicate the special features of such a system that make the management of the data components unique through definition of the variety of system components. An attempt was made to indicate the special data management requirements that do not fit into the scope of commercially available data base management systems. The contention was made, however, that a spatially oriented data base can be managed through a general system that is designed especially for spatial data, and at the same time remain within the philosophical structure of a DBMS as defined by the CODASYL Data Base Task Group [18]. The following proposed model attempts to adhere to this philosophy.

The basic principle followed in the proposed model pertains to the interface between user and data. The user will be allowed multiple views of the data, independent of the physical storage of the data. For example, if Landsat data 'A' and aircraft data 'B' are to be processed, the user may specify:

GET A, B GRID RESOLUTION = hectare

indicating that each data set is retrieved in a grid format at a hectare size resolution. Alternatively, the data labeled 'B' could have been polygonal data. The same request would have necessitated the employment of a polygon to grid algorithm. The user sees grid data of similar resolution, even though the storage of these data are not necessarily grid in structure nor similar in resolution. Structurally, (see Figure 2) the system is a Data Base Management System with the addition of a user/system interface through a batch or Interactive Processing Language (IPL). Each element will be examined more closely.

Data Storage Management

Geographic data are distinguished, as mentioned previously, by the methods employed to retrieve the data. This in turn is reflected in the data's physical storage structure. The data base manager is responsible for those storage structures. It is informed of the particular structures and data sets active in the data base through the data definition created by a data base administrator through the data definition language (DDL).

Data Base Manager (DBM)

The data base manager is a set of software programs that interfaces between a user or program request for data and the physical representation of those data in storage. This software needs to be aware not only of the data sets that comprise the data base but also of the permissible methods of retrieval. For example, a user may not request signatures to be retrieved in polygonal format, though he may retrieve polygonal data in grid format. The DBM passes the user request for data and invokes format service routines (FSR's) available to establish a working data set which will be processed by the application programs. The DBM learns the data structure through the data definition language (DDL) and communicates with the processing system through the data manipulation language (DML).

The Data Definition Language (DDL)

Before a data base is created, the administrator of the data is responsible for determining what data types will be employed. Joined with this responsibility is the need to establish the physical storage characteristics permitted for each data type, the permissible retrieval mechanisms, and whatever logical linkages may occur between data types. The definition of a data type must here be distinguished from the occurrence of a data type. The administrator does not load the data base, at this point, with real data, but defines the types of data that are permitted in the data base. These data definitions are then communicated to the data base manager using a data definition language. A data definition would include as a minimum: (1) the data type name, (2) the physical storage specification, (3) permitted linkages between other data types, and (4) permitted retrieval formats. The data base manager would invoke a data definition generator which would construct internal tables designating the permissible data types and structures.

Data Manipulation Language (DML)

Interaction between application programs and the data base is through a data manipulation language that is interpreted by the data base manager. Typically, a DML would consist of five or six verbs hosted in another language through subroutine calls. DML verbs could include the following (or their equivalents):

STORE -- which would create the occurrence of a data type, i.e., load the data base

GET -- to retrieve data

FIND -- to locate data

MODIFY -- to alter data

DELETE -- to remove data

Each verb would in turn be modified appropriately to supply the data base manager with sufficient information for successful data interface. For example, a prototype GET command may consist of:

GET dataset(s) mode in-location modifiers

where one or more data sets would be retrieved in grid, polygonal, or list mode and stored at "in-location" as modified by "modifiers" (eg., resolution, region).

Data Processor

Five basic needs arise: (1) a grid data processor, (2) a linear data processor, (3) a signature processor, (4) list data processors, and (5) data display mechanisms, eg., graphics. This paper does not address engineering of these needs. The significant point to be stressed is that the data set processed will be that set termed the "active data set" and prepared to satisfy the current user's view of the data in the data base.

The User

The processing system and user interface through batch-mode operation or interactively through an interactive processing language (IPL). This is to say that the typical user is not a programmer, hence the user is supplied with a very high-level language interface. The supplied vocabulary depends, of course, on the processing functions available in the system.

6. CONCLUSION

A geographic information system is viewed as an organizational resource that serves more than one user. However, not every user's view of the data is the same. In order to support multiple views of the same data, the data base manager is provided as the data/program/user interface. The same skeleton system and supportive software can be supplied to any user employing geographically oriented data. Specific modules can then be developed to resolve the needs of the particular application. Those modules, being data storage independent, need not concern themselves with data base formats. This design provides the following advantages:

- Permits multiple views of the data
- Separates data from data processing functions
- Provides integrated collection of data
- Provides centralized and efficient control of data
- Provides independent management of data security, quality, and integrity
- Minimizes duplication of data
- Automates data filing efforts
- Provides high-level interface for a wide variety of users

Most importantly, a generalized approach to the definition and design of a geographic information system can provide a tool adaptable to different users and different applications.

ACKNOWLEDGEMENTS

The author would like to acknowledge William Malila of ERIM, Professor Waldo Tobler, Professor Richard Phillips, Professor Alan Merten of The University of Michigan, and Ronald Shelton of Michigan State University, for their advice and assistance in researching this paper.

REFERENCES

1. Steiner, Dieter, and A. Salerno, Manual of Remote Sensing, "Remote Sensor Data Systems, Processing and Management", Chapter 12, Vol. 1, 1975, p. 611.
2. A Buyer's Guide to Data Base Management Systems, Data Pro Research Corporation, 1975.
3. Philips, Richard L., "A Query Language for a Network Data Base with Graphical Entities", CICE Department, The University of Michigan, 1977.
4. Analysis of Computer Support Systems for Multi-Functional Planning, Report III, Management Sciences Staff, Forest Service, USDA, June 1976.
5. Anuta, Paul E., Computer-Aided Analysis Techniques for Remote Sensing Data Interpretation, LARS Information Note 100675, Purdue University, 1975.
6. Zobrist, Albert L., Elements of an Image-Based Information System, Jet Propulsion Laboratory, Pasadena, California, 1976.
7. Fry, James P. and Edgar H. Sibley, "Evolution of Data-Base Management Systems", Computing Surveys, Vol. 8, No. 1, March 1976.

8. Adapted from: Tomlinson, R. F., Geographical Data Handling, Second Symposium on Geographical Information Systems, Ottawa, Canada, August 1972.

9. Information/Data Handling: A Guidebook for Development of State Programs, U.S. Department of the Interior, July 1975.

10. Istvan, L. B., Land Use Interpretation from Aerial Photography for Input to the Section 208 Water Quality Management Program of the Toledo Metropolitan Area Council of Governments, ERIM, February 1976.

11. Deuker, K. J., "A Framework for Encoding Spatial Data", Geographical Analysis, Vol. 4 (1972), pp. 98-105.

12. Warntz, W., The Redbook, Harvard University Press, Cambridge, Massachusetts, 1971.

13. Cooke, Maxfield, The Dime Encoding System, Bureau of the Census, Washington, D.C., Census Use Study Report 4, 1967.

14. Chrisman, Nicholas, Cartographic Data Structures Panel, Proceedings of the International Conference on Automation in Cartography, "Auto-Carto I", Reston, Virginia, December 9-12, 1974, pp. 165-177.

15. Philips, R. L., "Computer Graphics in Urban and Environmental Systems", Proceedings of the IEEE, Vol. 62, No. 4, April 1974, pp. 437-452.

16. Chamberlin, Donald D., "Relational Data-Base Management Systems", Computing Surveys, Vol. 8, No. 1, March 1976.

17. Taylor, Robert W. and Randall L. Frank, CODASYL Data-Base Management Systems, Computing Surveys, Vol. 8, No. 1, March 1976.

18. CODASYL Data Base Task Group, ACM, New York, April 1971 report.

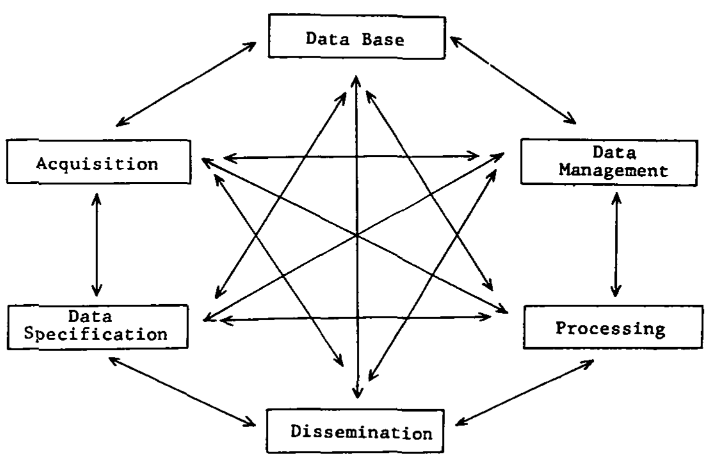

FIGURE 1 SCHEMATIC REPRESENTATION OF THE RELATIONSHIPS BETWEEN COMPONENTS OF AN INFORMATION SYSTEM

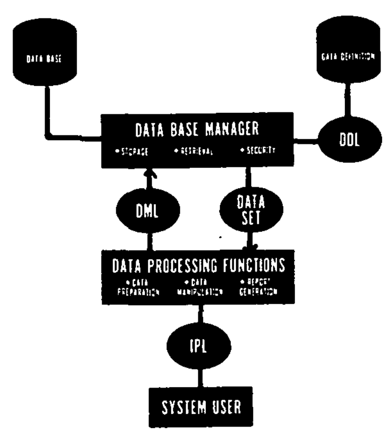

FIGURE 2. A GIS DESIGN MODEL