Reprinted from

Tenth International Symposium

Machine Processing of

Remotely Sensed Data

with special emphasis on

Thematic Mapper Data and

Geographic Information Systems

June 12 - 14, 1984

Proceedings

Purdue University
The Laboratory for Applications of Remote Sensing
West Lafayette, Indiana 47907 USA

GEOGRAPHIC INFORMATION SYSTEMS: QUESTIONS TO ASK BEFORE IT'S TOO LATE

J.L. STAR, M.J. COSENTINO

University of California
Santa Barbara, California

T.W. FORESMAN

Naval Civil Engineering Laboratory
Pt. Hueneme, California

ABSTRACT

A geographic information system (GIS) is a means to analyze and manage spatially-referenced data. There are two different kinds of questions which must be considered in the design of a GIS: technical, and management.

From a technical viewpoint, one must consider data storage capabilities, the kinds of analyses needed, and report generation, among other things. From a management perspective one must consider problems of people, time and money: what kinds of people will use the system, how rapidly must the system be able to complete a task, how much time is available for integrating a new system into the organization, what money is available for both developing and maintaining the system in the long term?

Both classes of questions must be asked in the design phase. Interaction between the ultimate users, system design personnel, and management is essential in order to address these different points of view. We present a case study where developing particular analysis scenarios assisted in designing a GIS.

I. INTRODUCTION

A Geographic Information System (GIS) is a specialized form of management information system. The design of a GIS is non-trivial, in terms of both technical and management issues. There are a great number of issues to consider when designing a GIS - particularly when the system is to be used operationally (in contrast to a research or evaluation system).

Resolution of these issues requires a constructive interaction between the ultimate users and the technologists. Unfortunately, this does not always take place. Making decisions in the design phase requires a bit of predicting the future. When the implementation staff asks the project managers "what is the finest spatial resolution you might want?", they are tempted to say "the best you can provide". When the design people ask "how rapidly must we be able to turn around a given problem?" they should not be answered "immediately if not sooner". These questions and their answers are the results of sloppy thinking, and can lead to unpleasant surprises when the system is supposed to be working.

Some time ago we were called to a meeting, to try and help design a system for geographic analysis and information storage for a particular user. After a day of discussing the project with the manager who would eventually use the system, we changed our approach. We began to develop a set of scenarios that force the local experts to show us what they really want to do, and specify many of the constraints in the process. In the material which follows, we provide a brief description of geographic information systems, and some important aspects of our analysis as an input to system design.

II. GEOGRAPHIC INFORMATION SYSTEMS

A. WHAT ARE THEY?

According to Short (1982), "A GIS is a georeferenced system for the acquisition, storage, retrieval, and manipulation of data." Before taking each of these in turn, it's helpful to make sure our use of terms is clear.

A database is an integrated and shared repository of stored data. We'll restrict our discussion to digital databases. A database management system is the software that handles access and modification of the database.

Spatial data are data with positional attributes. Entries in the system have accompanying geodetic coordinates. Spatial data may also have nonspatial attributes, and often do. For example, the entry in a database may have a latitude/longitude description for its position, as well as the nonspatial attributes of elevation, soil type, and vegetative cover. Spatial databases tend to be large, and often include many kinds of data structure (i.e., raster, vector, and scalar) which are not easily intercompared. Converting diverse kinds of structures, on different scales, recorded in terms of different systems of geocoding, into a common information base is one of the principal tasks of a GIS.

Acquisition. There are many kinds of data, in terms of structure and function - each may require a different kind of input. Digital imagery, analog imagery, maps, vector/polygon lists, point files, and so forth all will arrive in different forms and scales. The information must be converted to a digital form when necessary, and spatially identified.

Storage. The underlying physical structure of the data should be hidden from the user. Storage is typically magnetic tape and disk libraries, with archived copies on magnetic tape.

Retrieval. This is a part of the database management function. We wish to take only what's needed from the database, in the form we choose.

Manipulation. This includes updating information, deriving new dataplanes, rescaling and changing projections, statistical analyses, mensuration, modeling and simulation.

B. WHAT ARE THE COSTS?

Costs may be broken down in a number of ways, but not until certain management issues are decided:

*Who will provide the service? An outside contractor? In-house staff on a vendor-supplied system? Or will you develop and support a system in-house?

*If a system is to be built in-house, is it a commercial system, one developed by another public agency, or one developed in-house? Each of these has different start-up and long-term costs, as well as different requirements for in-house staff.

*What are the start-up costs? In addition to straight hardware and software costs, does your organization already have the kind of people able to set-up and manage a data processing/database management system/computer aided design environment, or will staff have to be developed? Is there an existing pool of information about the site which can be put into the new system?

*Who will create the database and maintain it?

*What will it take to integrate the system into the existing organizational structure with minimal disruption.

*What are the long-term support and operation requirements, in terms of data acquisition and update, hardware/software and staff?

The technical issues which drive the design of a GIS are fairly well known, and revolve around operational requirements. How often will the system be used? How big are the "normal" problems? What response time is acceptable? How might the system load grow over time? What kinds of analysis must be performed? How much and what kinds of data need to be acquired over time, and what is the associated storage? What kinds of tasks are being performed now? What tasks can't be performed now?

III. CASE STUDY

A. MARINE CORPS COMBAT CENTER, TWENTYNINE PALMS, CALIFORNIA

The U.S. Marine Corps requested that we consider (and recommend if appropriate) a geographic information systems approach to natural resources management at their facilities within the United States. The specific test case was the Combat Center, Twentynine Palms, California. The U.S. Marine Corps Combat Center is approximately 600,000 acres, which covers portions of 42 U.S.G.S. 7.5 minute topographic quandrangle maps.

Part of our task was to consider remotely sensed data as an input to a geographic information system, We believe that commercial systems are deficient in the integration of remotely sensed data, but rapid progress is being made in a number of areas.

B. ACQUISITION

For natural resource management at Twentynine Palms, we can identify minimally thirty separate data planes, including elevation, geology, soils, land use, and habitat for a number of plant and animal species. Many of these are now archived as flat maps. Digital products are available for elevation and some forms of remotely sensed data, the latter as a repeating source for updating other data planes.

C. STORAGE

A 12.5 meter minimum mapping unit has been selected by the natural resource manager, as an acceptable balance between costs and required precision. Given this resolution, a raster data plane covering 600,000 acres has approximately 16 million pixels. For an average data plane this might represent tens of megabytes of digital data. This is unmanageable for a manual system or a microcomputer system with floppy disks, but well within reason for a minicomputer system.

D. RETRIEVAL

Land use management decisions at the base are made almost daily, constraining the response time of any system. For example, if an outside contractor were providing the needed services, arrangements must be made to provide overnight response - which would be most unusual. An on-site interactive analysis system is the more reasonable alternative.

E. MANIPULATION

One of the manipulations we studied as a typical processing task was a proximity analysis. We considered preparing a map of the base to identify locations of rare and endangered species which are within 500 meters of a traffic corridor. This is a complicated task, involving a number of data layers (known roads, tank trails, and land use boundaries, and known species habitat) and a spatial operator.

Such an analysis would be difficult for any system where entire data layers for the area cannot be accessed at once; this permitted us to recommend against the floppy disk/microprocessor system which had been under consideration. This kind of analysis would also be difficult if done manually, due to both the volume of data and the complexity of the analysis.

F. RECOMMENDATIONS

Based on our evaluations, we suggested that Twentynine Palms acquire a commercial system. A number of companies can supply minicomputer-based GIS's which are able to handle most of the Combat Center's needs. Generally, such systems are only limited by cost in terms of the amount of data they may access, have well known and understood costs of installation and ownership, and have extensive support available on a fee-for-service basis.

The most important issues to be addressed are costs of integrating the new system into existing practice, and staff development. We also noted that the GIS area is relatively immature; new products and capabilities are appearing rapidly. Future upgrades and enhancements must be considered when developing long range plans and when choosing a particular company or group.

IV. CONCLUSIONS

A scenario approach is a useful strategy for GIS design, if one recognizes two cautions. First, the scenarios must be chosen with care to stretch the system's capabilities. Four to six specific tasks that are reasonable to expect of the new system, and are difficult or costly to do with the old, should be worked out in detail. Second, the management issues surrounding system acquisition, development and support are as important as the technical issues.

V. ACKNOWLEDGEMENTS

We gratefully acknowledge Dr. John Estes, Dr. Tom Mace, and Dr. Ralph Brown for their help.

VI. REFERENCES

Green, W.B., Digital Image Processing - A Systems Approach. Van Nostrand and Reinhold Co. (1983).

Moik, J., Digital Image Processing of Remotely Sensed Images. NASA SP-431 (1980).

Pavlidis, T., Algorithms for Graphics and Image Processing. Computer Science Press (1982).

PECORA VI Symposium. Remote Sensing: An Input to Geographic Information Systems in the 1980's. 1981 Proceedings, American Society of Photogrammetry (1982).

Pazner, M., Improved Querying of Large Spatial Databases Utilizing Artificial Intelligence Techniques. Unpublished manuscript.

Short, N.M., The Landsat Tutorial Workbook - Basics of Satellite Remote Sensing. NASA Reference Publication No. 1078 (1982).

VII. BIOGRAPHICAL DATA

Dr. Jeffrey L. Star received his bachelor's degree in Environmental Studies from the Massachusetts Institute of Technology, and a doctorate in Oceanography from Scripps Institute of Oceanography, University of California, San Diego. He is a Development Engineer and Manager of the Remote Sensing Research Unit, Department of Geography, University of California, Santa Barbara, His research interests are in environmental science and applications of remote sensing.

Mr. Michael J. Cosentino has a bachelor's degree in Forest Resource Management from the University of California, Berkeley, and a master's in Biogeography from University of California, Santa Barbara. He is a Staff Research Associate in the Environmental Studies Department, University of California, Santa Barbara. He specializes in applications of remote sensing to wildlands resource management.

Mr. Timothy W. Foresman received both bachelor's and master's degrees in Biology from San Diego State, and an additional master's degree in Environmental Engineering from University of Southern California. He is an ecologist and environmental engineer for the Environmental Protection Division, Naval Civil Engineering Laboratory, Pt. Hueneme, California.