

## ESTABLISHING A CORPORATE GIS DATABASE FROM MULTIPLE GIS PROJECT DATA SETS

Timothy R. Johnson  
Karen C. Siderelis  
Land Resources Information Service  
P.O. Box 27687  
Raleigh, North Carolina 27611-7687

### ABSTRACT

The Land Resources Information Service (LRIS) has operated the geographic information system (GIS) for the State of North Carolina since system implementation in 1977. LRIS was established to provide the State with a geographic data capture and analysis capability to support natural and cultural resource management. LRIS has conducted several hundred projects over the past 12 years utilizing the capabilities of the GIS to build specialized data sets and perform various analyses with those data sets. LRIS has reached a point where its data resources and the associated investment in the capture and maintenance of that data have become very significant and the use of data is at times unwieldy. The need to establish a statewide "corporate database" of geographic information has become an important issue affecting further growth and the ability to respond effectively to new projects. This paper addresses the background and issues related to the decision to establish a corporate database, the methodology used to design the database, the implementation status to date, and the anticipated benefits of completing such an effort.

### BACKGROUND

Recognition of the need for an information system to support geographic-oriented decision making in North Carolina arose out of the North Carolina Land Policy Act of 1974 (Tribble and Siderelis, 1988). This legislative mandate led to the establishment of LRIS as an agency within state government and to the purchase of geographic information system hardware and software from Comarc Design Systems in 1977. Over the years, LRIS performed a wide variety of project work on a cost recovery basis with only a minimal level of state appropriations. These projects can generally be described as one-time efforts for diverse geographic areas (e.g., individual counties or application-specific portions of the state) using a wide variety of base maps ranging in scale from 1:24,000 to 1:250,000. The data layers captured as part of these projects also varied considerably. Typical base layers such as transportation, hydrography, and political boundaries were captured as well as more application-specific layers such as primary and secondary nursery areas for fish and historic and archaeological sites.

During the course of growth of LRIS, GIS technology changed tremendously with advances in both the power and utility of hardware and software to support spatial problem solving. LRIS eventually outgrew the capabilities of the Comarc software and purchased ARC/INFO in 1986. This decision prompted the organization to look at its investment in data captured using the Comarc system and the desire to maintain as much of that data as possible within the ARC/INFO environment. It also was an opportunity for LRIS to develop initial concepts for the design of a corporate database.

A corporate database is one that is established as a corporate or organizational resource as well as one that meets several criteria. The criteria used to guide the design of the corporate database are given below.

- The database should be practical to maintain in terms of institutional and digitization cost considerations.
- The database should consist of data that satisfy common needs.
- The database should be structured to maximize its utility.
- The database should not include all digital data holdings (i.e., not every data layer in digital form should be part of the database).

## GOALS AND ISSUES

LRIS is a service organization whose mission is to maintain a digital database of geographic information for the State of North Carolina and to provide GIS services to a variety of users including agencies of state government and other levels of government as well as the private sector. As a service organization, LRIS attempts to provide services as cost effectively as possible. The cost recovery nature of the work requires that LRIS look carefully at improving the means by which users are served. The most important issues involved in establishing a corporate database for LRIS are those which are related to providing a high quality of service at the lowest possible price for users.

The primary goals which LRIS hopes to attain in establishing a corporate database are to: (1) improve response to user demands for data and services; (2) maintain a high degree of data quality; (3) minimize extensive new data capture costs, a significant component of total GIS operating costs; and (4) develop better organization of data resources. Current and future user needs are the cornerstone for the design of a corporate database. Even with the best of intentions, a database designed and developed with the wrong data at the wrong resolution for the wrong users yields a data resource which will not be used as intended. A thorough examination of the characteristics of the LRIS user community is therefore necessary as a first step.

Data quality is a major issue as one begins to identify the components of a corporate database. The quality of service that LRIS provides starts with the quality of the data captured and used to perform spatial analyses. If those data are to be used successfully beyond a single, short term project, a high degree of quality must be assured. These data must also be available over the entire state at a level of resolution which is meaningful for analytical needs. LRIS has a substantial amount of data developed under the Comarc data structure. Cost/benefit assessments of quality, currency, and extent of coverage are parameters necessary to evaluate whether it is most beneficial to convert this data into a structure usable by ARC/INFO or to recapture or in some way obtain new data to replace the old Comarc data. These parameters, either individually or together, determine whether a particular project data set will be useful for conversion.

Data availability is another issue of concern. The corporate database is intended to represent statewide coverage for each layer of data selected for inclusion in the database. For some layers, data coverage is extremely incomplete and would thus require a substantial digitization investment. The best alternative would be to obtain digital data from other reliable sources to minimize the cost of data acquisition. LRIS has worked with the USGS, for example, to obtain digital data in a timely and efficient manner to fulfill some data needs.

In order to establish a corporate database and simultaneously meet the challenge that it entails, LRIS needs to ensure that the database contains data that are practical to maintain over the long term. This issue involves institutional considerations as much as, if not more than, technological ones. Data layers that are to be routinely maintained require the backing of the institutions that were the sources of those layers in the first place. The institutions must be committed to providing updates to the data on some reasonable schedule. Without this kind of support, such data become impractical for storage and maintenance in a corporate database.

## DESIGN METHODOLOGY

LRIS has identified a methodology for designing the corporate database with the above goals and issues in mind. The design methodology being used for the corporate database consists of a nine-step process which is outlined below:

1. Evaluate current and future needs of the user community.
2. Inventory current data holdings.
3. Perform quality assessment of data holdings.
4. Assess data volume, currency, resolution, coverage, and frequency of use of data holdings.
5. Consider future application demands and derive data requirements to meet them.

6. Identify the most critical data to be stored and prioritize the remaining data for storage consideration.
7. Develop data organization strategy.
8. Prepare implementation plan.
9. Perform implementation of the database.

Each step produces information that is essential in achieving the overall design and implementation goals for the database.

#### Evaluate the current and future needs of the user community

This step identifies the application and data needs of current LRIS users and many potential users including such parameters as frequency of use of specific data layers; provides LRIS with insight into where finite resources should be spent and projects to target in the future.

#### Inventory current data holdings

LRIS has over 60 graphic layers of data and numerous associated tabular data sets in digital form which have been developed over the past 11 years for several hundred projects; simply documenting this data gives the organization a good start on the evaluation process.

#### Perform quality assessment of data holdings

This examination assists LRIS in determining whether specific project data sets are worthy of inclusion in a corporate database, which by definition will be maintained and updated on a routine basis; assessment of quality includes consideration of the source material and digitization methods used, among other factors.

#### Assess data volume, currency, resolution, coverage, and frequency of use of data holdings

These factors are useful in determining the utility of existing data holdings. Data volume provides a first indication of the potential size of the corporate database; data currency addresses the vintage of the data which is potentially important depending on the type of data; data resolution is concerned with the level of detail at which data were captured and indicates its possible usefulness in the long term as well as the relative cost of capturing and maintaining it; data coverage is the geographic extent of the digital data (e.g., statewide, a specific watershed, individual counties); and the frequency of use is an indicator of the value of a data layer or data set and may have implications on the design of the database as a whole.

### Consider future application demands and derive data requirements to meet them

This step is based upon the earlier evaluation of user needs and in light of the results of the assessment of current data holdings. The product of this step is new data needs that should be accounted for in the corporate database design.

### Identify the most critical data to be stored and prioritize the remaining data for storage consideration

This step considers the needs of the user community such as frequency of use of certain data and other parameters to arrive at a set of data storage priorities; it will also provide an indication of the data volume necessary to support users when coupled with size information for current data holdings.

### Develop data organization strategy

This is essentially the detailed design portion of the methodology. Alternative strategies are derived which are workable in the LRIS organizational environment with the design goals enumerated earlier in mind. The strategy that best meets user needs and is acceptable to the LRIS director, production management, system management, project development, and database administration units of LRIS is selected for implementation. The key selection factors are: costs, impacts, and potential benefits.

### Prepare implementation plan

This step lays out the schedule for implementing the corporate database. The plan includes not only the details of the process and the order of implementation but also time and resource estimates for completing the implementation.

### Perform implementation of the database

Based on the plan, this step is the actual setup and loading of the corporate database in accordance with the schedule outlined.

A cyclical process is implied in the methodology whereby data needs and data holdings are continually being matched against one another and refined until a clear picture of the content and design of the corporate database emerges.

## PROGRESS TO DATE

Progress toward development of an LRIS corporate database had been slow until recently due in part to the cost recovery nature of operations.

However, LRIS has been selected as the data management center for the Albemarle-Pamlico Estuarine Study (APES) Program, one of the National Estuarine Programs funded by the Environmental Protection Agency. The APES program is an appropriate vehicle for beginning to develop the corporate database for three reasons: (1) the APES study area includes the entire eastern portion of North Carolina, (2) LRIS is responsible for designing, implementing, and maintaining a GIS database for the study area based on a user needs assessment, and (3) funding is available to perform a thorough, formal database design.

At the time of this writing, the first three steps are nearing completion. The user needs evaluation consisted of a set of over 50 interviews with current and future user organizations to determine both application and data needs. Information from these interviews is being analyzed. The LRIS staff has concurrently developed an inventory of data holdings based on several hundred projects completed since 1977. These data are currently being assessed for their utility as part of the corporate database using data quality and user needs as primary criteria. The remaining steps in the methodology are scheduled to be completed during the 1989 calendar year, including actual implementation of the database.

## CONCLUSIONS

The lessons that LRIS has learned in dealing with GIS project data sets over the past 11 years have been numerous. The design of a corporate database should have been completed sooner, but the cost recovery mode of LRIS operations kept the focus on day-to-day and week-to-week activities rather than the long term. Additionally, based on the experience of the organization in operating a GIS, a broader view of data has emerged in terms of how the data captured fits into an overall database scheme --- i.e., a corporate database.

LRIS expects to attain its goal of implementing a thoroughly designed, useful corporate database to take the organization into the 1990s in service to the user community. The reasons for anticipating this success are the solid methodology identified for achieving the goal, the APES program as a vehicle for boosting LRIS in that direction, and simply the fact that the time has come for the organization to implement such a database. The benefits are numerous and are directly related to providing a higher level of service in terms of cost, timeliness, and efficiency.

## REFERENCES

Tribble, Thomas N., and Karen C. Siderelis. 1988. "Status of the State Geographic Information System in North Carolina". Proceedings of the Urban and Regional Information Systems Association, Los Angeles, Vol I: 48-57.