Iowa Department of Transportation Statewide Coordinated GIS

William G. Schuman, Tim Strauss, Dan Gieseman, and Reginald R. Souleyrette

This paper details a project conducted by the Iowa Department of Transportation (DOT) and the Center for Transportation Research and Education (CTRE) to construct a statewide coordinated GIS. This effort was undertaken in response to increased pressures on transportation agencies to improve their efficiency and accountability through performance-based planning. The first step in response to this pressure was to better integrate the disparate data throughout the Iowa DOT. The project team conducted interviews of data collectors, managers, and users. They then identified key data elements, determined the system architecture needed to support and maintain the DOT's data, and created the necessary data conversion procedures to populate the GIS database. As most transportation data are spatial in nature, GIS was the logical choice for data integration. Implementation of the Statewide Coordinated GIS is progressing at the Iowa DOT using GIS, data warehousing, Local Area Network (LAN), and Wide Area Network (WAN) technologies. Internet and CDROM solutions will be used to distribute data to organizations external to the Iowa DOT. While these methods facilitate the sharing of data, it is very important that the DOT design a comprehensive data maintenance, metadata, and distribution strategy so that well documented data are provided on a regular cycle. A graphical user interface will provide a straightforward method of accessing the disparate data sources without requiring the end user to be familiar with the complexities of the underlying data sources. An interface to provide online metadata to the end user is also part of the system development. Key words: management systems, GIS, database, warehouse, metadata.

INTRODUCTION

The Iowa DOT has a wealth of information stored in many database systems throughout the organization. Some of these systems are well established and maintained, and others are smaller data sets collected by an individual for a specific office. In either case, while data are valuable in isolation, integrated data sources provide a much broader view of the information and the interrelationships of the data.

W. Schuman, Iowa Department of Transportation, 800 Lincoln Way, Building #5, Ames, Iowa 50010. T. Strauss, D. Gieseman, and R. Souleyrette, Center for Transportation Research and Education, Iowa State University, 2625 N. Loop Drive, Suite 2100, Ames, Iowa 50010.

Data integration is necessary for successful implementation of many national transportation initiatives. The Intermodal Surface Transportation Efficiency Act (ISTEA) initially mandated the creation of systems that would have required the development of integrated databases. While no longer mandated, the requirements of these systems heightened an awareness of the value of coordinated data to transportation professionals and management. In addition, Iowa's Blue Ribbon Task Force and the IowAccess initiatives have emphasized taking advantage of technology to provide more efficient methods of making decisions and interacting with Iowa's large and growing information infrastructure.

In most instances, the spatial component has been the element for facilitating the integration of these data. Many advances have made it more practical to provide integrated spatial data. The hard technologies, such as faster computer systems and better computer networks, have allowed the data to be made available on the user's desk. Easy to use software, such as Microsoft Windows GIS applications, have also played an important role in the ability of an organization to get the needed tools into the hands of the users. Even more important than these technological advancements may be the affordability of the technologies. GIS, until recently, was a tool only available for larger, well-funded organizations due to its high costs and learning curve. Lastly, the opening of proprietary data structures has made it easier to share data between systems, thus facilitating the sharing of data between or within organizations using different software products.

EXISTING SITUATION

The Iowa DOT has many offices that collect and manage data. While these data are important in the development of a Statewide Coordinated GIS, many other data sources exist outside the DOT that would aid in more efficient and better decision making. These internal and external data sources are described below.

Internal Data Structures

The data maintained by the DOT are related in many ways. The relationships can be direct, such as using a specific value from certain data entities, or indirect, such as a spatial relationship (e.g., a road crosses through a wetland). Figure 1 demonstrates the complex data relationships that exist at the DOT.

This situation is not unique to the Iowa DOT. Many DOTs in the United States have identified the need to better integrate their data and to identify the redundant data collection and maintenance

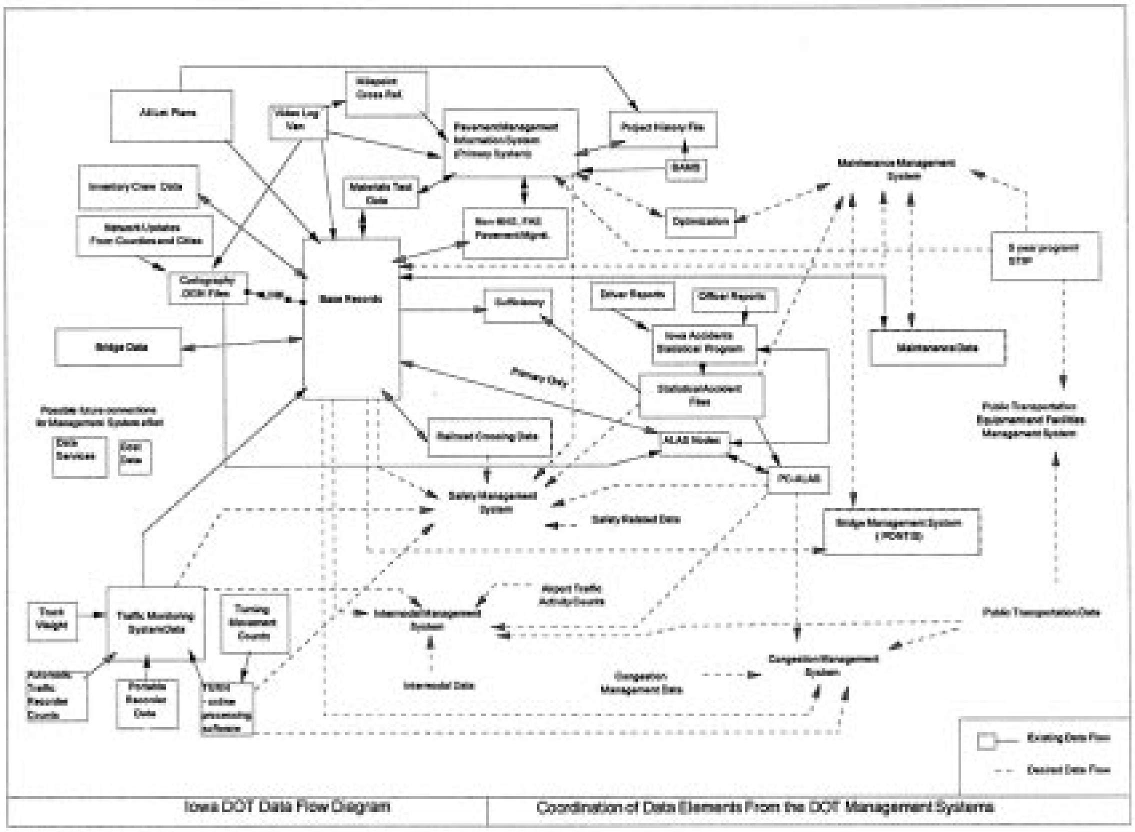

FIGURE 1 Iowa DOT data relationships.

problems within their organizations. Many DOTs, such as Maine, Texas, Virginia, Oregon, and Mississippi have data integration and GIS implementation projects currently underway.

Examples of Iowa DOT data include road, bridge, accident, and environmental data. These data are maintained in various mainframe and PC databases.

External Data Structures

In addition to the data collected and maintained by the DOT, several other data sources are available for use and are needed by the DOT users.

The Iowa Department of Natural Resources (DNR) has an established GIS with many useful layers of information. Examples of DNR data useful to the DOT include soil types, wetland information, location of underground storage tanks, and topography. The DNR's data are maintained in Environmental Systems Research Institute's (ESRI) PC Arc/Info software.

Several other organizations make data available to the public through Internet sites or by ordering the data on CDROM. The organizations that provide data useful to the DOT are numerous, but include the United States Geologic Survey, Census Bureau, National Wetlands Inventory, Bureau of Transportation Statistics, and several private organizations. These organizations may also provide aerial and satellite imagery that can be used by the DOT to enhance the accuracy of the existing data or to provide better information about the surface of the earth.

DATA INTEGRATION EFFORTS

Due to the large number of databases that store information at the DOT, it is impractical to consider translating all the existing data to one format. Additionally, numerous legacy systems were developed to maintain and query the data in those databases. Rewriting all the maintenance and query interfaces would be an insurmountable task for the DOT. For that reason, data warehousing concepts were identified as a means of integrating the data without having to merge all the data into one database management system for maintenance.

Data Warehouse

The data warehouse architecture allows data to reside in its current legacy system for maintenance, yet provides an environment where the data can be housed for better integration and used in GIS queries. Figure 2 illustrates the data warehouse integration process.

The data in the existing databases will be updated in the data warehouse based on user needs and maintenance cycles. For ex-

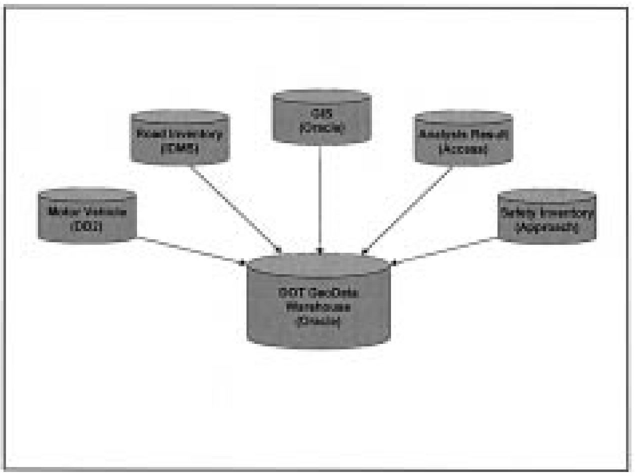

FIGURE 2 Data warehouse integration process.

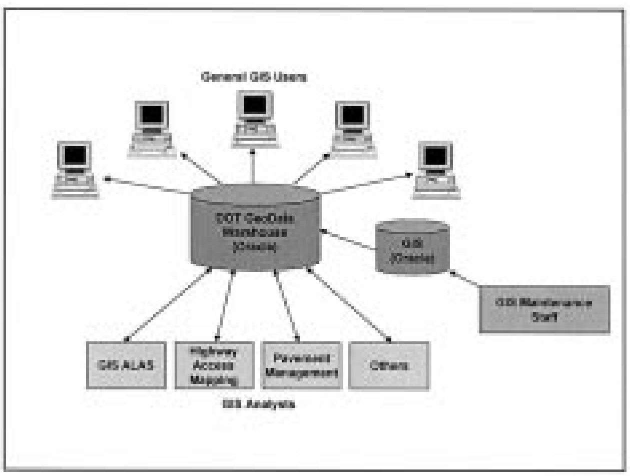

FIGURE 3 User access to the data warehouse.

ample, data updated once every year in December does not need to be updated in the warehouse on a nightly basis. Conversely, construction site status might need to be updated several times per day, if necessary for Intelligent Transportation System development. Real-time data replication processes can also be utilized if the need for dynamic data exists for certain systems (e.g., incident management applications). Storing the data in a warehouse creates a data storage environment that would accommodate all the GIS users at the DOT, whether they are highly skilled GIS analysts or occasional users of the data.

GIS Analysts (shown in Figure 3) have both read and write access to the data warehouse. They will use the data in the warehouse to perform their analysis, and may post the analysis results back to the warehouse for use by the rest of the DOT's users. These users will utilize whatever GIS software best suits their analysis needs.

General GIS Users will have read-only access to the data in the warehouse. These users will be provided a custom interface that allows them a method to find the data they need without having to be knowledgeable of the warehouse structure. This user group also has the ability to do analysis, but can not post their analysis results to the warehouse. The analysis results can be stored only in the user's personal storage spaces such as a local machine or personal network directory.

In order to minimize the customization needed to provide this functionality, the "off-the-shelf" software selected to access the warehouse has the ability to merge several GIS structures. This is a requirement due to the desire to read the warehouse in its "native state" without having to translate the internal and external data formats to a single GIS format, or transform all the maps to a single map projection. Intergraph's GeoMedia software was selected for the *General GIS Users* because it provides the capability to handle the multiple formats and map projections invisibly to the user. In subsequent development efforts this software will be more tightly intertwined with other systems, such as video logging and records management.

Metadata Development

When large amounts of data are provided in an environment such as a data warehouse, many questions arise about the source, accuracy and meaning of the data. To address those questions, metadata about the data in the warehouse will be maintained. Metadata development can be quite involved, but the DOT's initial efforts in metadata development will store the source, accuracy, date of collection and map projection of the data, in addition to a description of each field and the allowable values that may be found in that field. For example, the field may be called *light_cond* in the warehouse, but the metadata may describe that as, "The lighting conditions that existed at the time of the crash." The values stored in that field may be *1,2,3,..,9,* but the metadata will describe each value as, "1 - Sunlight, 2 – Twilight, 3 – Dawn, etc." This information will aid the users in their ability to build queries and understand the reliability of the results they get from a query. Future developments in the metadata will utilize the metadata standards that are defined by the Federal Geographic Data Committee.

IMPLEMENTATION ISSUES

In an effort as large as the departmental implementation of GIS and the development of a data warehouse, it is inevitable that implementation issues arise. Three key areas of concern are addressed here: training and resources, equipment acquisition, and institutionalization of the project.

Training and Resources

The training of the personnel involved in the use and support of the data warehouse and GIS is key to the success of the Statewide Coordinated GIS. It is often obvious that end users need training in the use of a new software tool, but as important is the training of the staff involved in support of the underlying architecture of the GIS and data warehouse. Since GIS is a relatively new tool for the DOT, it is necessary to train support staff in Oracle (the selected

FIGURE 4 Data warehouse interface.

relational database), the GIS, and the software used for customization of the GIS. The new training is often a difficult task to accomplish since GIS implementation is often added to the numerous responsibilities of the support staff. As a result, it is necessary for the DOT management to prioritize the need for GIS with the other efforts that are underway or needed. If different priorities are given to GIS in each Division of the DOT, Department-wide implementation becomes more difficult.

Equipment Acquisition

The acquisition of additional hardware and software to support the data warehouse and GIS is also key to the success of the Statewide Coordinated GIS. The project is on track for hardware and software needs due to effective planning, and the current network installation efforts. As users identify new data requirements, it is estimated that the storage capacity for the data warehouse may reach the five-hundred gigabyte range in the next two years. Thus, it is important to continuously keep management aware of the progress and the successes obtained to ensure future staff and financial support for the system.

Institutionalization

It continues to be important to consider the user's daily workflows when developing the system. The success of the Statewide Coordinated GIS will be measured mostly by its ability to aid the users in their business functions. For that reason, not only is it necessary to train the users in the use of the software, but in the use of the software to do their jobs more effectively. This is a difficult process because the trainer must be knowledgeable in not only new systems, but also in the users' workflows. For that reason the training process will utilize expertise from CTRE, who has been involved in the workflow analysis, and DOT personnel that have a basic understanding of the GIS technologies and data to help train additional DOT users. It is also important that the current procedures at the DOT be reviewed to include GIS where appropriate.

FUTURE PLANS

The current plans for the Statewide Coordinated GIS include the efforts to enhance the user interface of GeoMedia for the DOT's business, development of a single Department-wide Linear Referencing System (LRS), utilization of Global Positioning Systems (GPS) technologies, and integration of other systems being developed at the DOT.

Enhanced User Interface

While the interfaces provided by the GIS software are useful, it is more productive to display the data to the user in a format oriented toward the user's business. Improvements would include clear aliases for database field names and literal verbiage for the values in the field. For example, the new user interface would reformat the old display of "*county_num:85*" as "*County Name: Story*." Another desirable enhancement would allow the user to enter values into two unrelated database tables and the custom software would automatically create the necessary spatial query to accommodate the request.

Linear Referencing Systems

The Statewide Coordinated GIS databases contain a large amount of linearly referenced data. To more efficiently utilize this data, a single LRS needs to be developed for the DOT that accommodates the many users and data collection efforts. The creation of this single LRS will alleviate the current method of "pre-segmenting" the DOT's Base Record road attributes based on the definition of a unique road segment. This would allow for simpler maintenance of the base road map and the road attribute data, and would allow users to utilize dynamic segmentation processes to query the data.

Utilization of Global Positioning Systems

GPS systems continue to become more accurate and require less expense and effort for a given accuracy. It is anticipated that this technology will allow the DOT's end users to add additional features to the base map for analysis and that the personnel maintaining the base map will be able to increase its overall accuracy. Standards and procedures need to be established for these processes to ensure the system integrity and to educate the users about the relative accuracy of the data they are using.

System Integration

Several information systems are currently under development or are being considered for development at the DOT. Three of these systems lend themselves to integration with the Statewide Coordinated GIS efforts: the Records Management System (RMS), the digital video logging system, and the Highway Closure and Restriction System (HCRS). Each of these systems utilizes attributes that are available in the Statewide Coordinated GIS. In an effort to minimize the work required of the end-user, these systems should be integrated to allow the end-user to process a query on only one of the systems and access data from the other systems. For example, a bridge engineer might query for the bridges over 20 years old with substandard deck inspections. After the results are displayed, it would be more efficient if the user could access the video log location of the bridge and the construction documents from the RMS without having to requery the other two systems again. This integration is more easily attained if the systems are designed with the integration in mind at the early stages.

CONCLUSIONS

It is evident that the integration of the data and the creation of the data warehouse will provide the end users with more data to use in their decision-making processes. The DOT continues to recognize the Statewide Coordinated GIS as an important part of improving the efficiency of data distribution at the DOT and additional resources are being devoted to the support of the project.

The current analysis processes depend too heavily on a small group of analysts that have access to the data. This system will allow the users to have more flexibility to apply "what-if" scenarios to the data and develop better solutions to problems they encounter.

The early stages of the project pointed out the need for a more tightly integrated maintenance and distribution process. The DOT continues to implement data warehousing and network technologies to make access to the data easier and provide "unprecedented" access to the Department's data.