

RESEARCH RESULTS DIGEST

June 1992

Number 186

These Digests are issued in the interest of providing an early awareness of the research results emanating from projects in the NCHRP. By making these results known as they are developed, it is hoped that the potential users of the research findings will be encouraged toward their early implementation in operating practices. Persons wanting to know the project subject matter in greater depth may do so through contact with the Cooperative Research Programs Staff, Transportation Research Board, 2101 Constitution Ave., N.W., Washington, D.C. 20037.

MATERIALS	ACT	USE
MRS Supv		
Materials		
20-7-46		
20-7-18		
Geol Staff		
Proj Dir		
Staff Engineer		
Pavement		
QC Coord		
Asst & Struc		
Exam & Asph		
Soils & Asph Mix		
TESTS		

Responsible Staff Engineer: Ian M. Friedland

Areas of Interest: 11 Administration, 12 Planning, 24 Pavement Design and Performance, 25 Structures Design and Performance, 40 Maintenance (Highway Transportation)

Data Interchange Standards for Bridge Management Systems and Integrated Highway Information Systems

An NCHRP digest detailing recommendations for data interchange standards developed under NCHRP Project 20-7, Task 46, "AASHTO Guidelines for Bridge Management Systems," conducted by The Urban Institute, Mr. William A. Hyman, Principal Investigator.

INTRODUCTION

Bridge management systems are one of many infrastructure systems that have been implemented, or are likely to be implemented, in state, regional, and local agencies throughout the country over the next decade. Proposed federal legislation would require that each state and metropolitan planning organization in large urban areas have bridge, pavement, congestion, and safety management systems. In addition there are legislative proposals that each state also have a traffic monitoring system.

Besides these initiatives, many governmental agencies throughout the country are implementing geographic information systems. Some agencies are introducing executive information systems to provide chief administrative officers with information on project and program status plus a wide variety of performance indicators. The U.S. Department of Transportation, Transportation Research Board, American Association of State Highway and Transportation Officials, and the states are engaged in a joint information retrieval system on a wide area network called Value-Added Network.

These management and information systems are being imposed on a wide variety of existing planning tools, decision-support systems, databases, and communication systems that reside in various agencies. Generally, the management tools and databases are not integrated. Many different kinds of databases exist, and the data are frequently stored in incompatible formats. Additionally, many databases serve just a single organizational unit.

One response to the lack of linkage or communication among databases has been the development of integrated highway information systems. A number of states have sought to tie together in a single database management system, files pertaining to pavements, structures, maintenance, accidents, finance, and traffic.

Such integrated databases are highly desirable and strongly recommended, but they are not a panacea. It is not always economically efficient to store data centrally. Even in client-server environments, where portions of a centralized database are fed to a specific user, it is sometimes more efficient to maintain certain data locally. Also, new databases and application

programs will continually emerge in response to the needs of individual organizational units.

Part of the solution to the problem of compartmentalization and inaccessibility of data is the development of standards for data interchange. Whenever desired data are not easily accessible from integrated files, the information must be transferred from one data structure to another. Demands for data transfer are expected to grow rapidly with implementation of a variety of infrastructure and traffic management systems including intelligent vehicle-highway systems. Moreover, there will be growing demand for transfer of data not only between organizational units of an agency but also between different levels of government. Data transfer will increasingly involve telecommunications and local and wide area networks.

This digest provides a summary of existing and emerging standards for data interchange that might be adapted to bridge and other infrastructure management systems, and contains a brief account of what might constitute a universal standard. The material in this digest was prepared by William Hyman as part of the final report on NCHRP Project 20-7, Task 46, "AASHTO Guidelines for Bridge Management Systems." The bridge management guidelines developed under this task will likely be published and distributed by AASHTO in early 1993.

FINDINGS

Overview of IHIS

An integrated highway information system (IHIS) is a computer-based system used by a transportation agency to organize all the data and information associated with a network of transportation facilities. This system provides a convenient, efficient, and cost-effective way of data storage, retrieval, reporting, and exchange. In general an IHIS belongs to a class of information systems referred to as data base management systems (DBMS).

An IHIS helps integrate different specific functions within a highway agency that reside in various individual units or divisions. In the past, separation has resulted in customized management and information systems, independent of systems used by other units. For example, road maintenance, bridge maintenance, and traffic operations exist as independent sections within the highway department. Separate bridge, pavement, and traffic management systems are used not only to store data relevant to each

operation but also to aid in developing short-term and long-term solutions to existing problems. From the individual unit's point of view, this approach may be an ideal management perspective. However, highway agencies have suffered from slow, costly, and tedious management of information and decision-making processes at the strategic network level resulting from the incompatibilities among information and management systems, and the duplication of data in the task-oriented tactical levels. This problem is the motivation for developing an integrated systems framework.

There are several issues that need to be addressed in setting guidelines for an IHIS, many of which have been discussed previously [Briggs and Chatfield, 1987; Paterson and Scullion, 1990]. They include organizational considerations, computer hardware and software, highway database contents and organization, and other administrative issues. The states that have implemented an IHIS have had different experiences developing systems mostly because of their organizational structure, their existing management and information systems, the specific needs of their organization, and their budgetary constraints. Because the actual implementation will vary from state to state, it is only possible to discuss the most general components and functions of an integrated highway information system. It should be noted that most existing IHIS are primarily data management tools with limited decision-support functions.

Components of IHIS

An integrated highway information system includes a collection of databases that are grouped according to their role in overall transportation planning and management. For use in network planning, programming, and budgeting on a state level, IHIS databases contain information on the road system as a network of roads with jurisdictional, functional, traffic demand, physical, and cost characteristics. Local and state highway agencies require this information to administer the road network.

For an IHIS to be functional, it should not only have databases but also software to control operations on the databases. This control system, called a database manager, allows files or records in the files to be linked, updated, searched, deleted, or accessed by other systems used by individual highway units or bureaus that perform specific functions and operations. Therefore, aside from the data, the IHIS has the following components [Connecticut Department of

Transportation, 1989]: (1) a reference system that allows data contained in each file to be referenced to a common point; (2) a data-collection system with all data files computerized and linked to the reference system; and (3) a system with the ability to manipulate data files within the system (IHIS) to develop desired analyses and appropriate reports.

The successful implementation of IHIS depends on how the above components are put together to satisfy not only the needs and requirements of the highway agency for which the integrated system is most beneficial but also the particular needs of individual units within the agency, which may or may not adapt to the centralized information management system.

In light of this it is necessary to identify the standards or rules pertaining to IHIS and its components. This may help some highway agencies plan the development of their own integrated systems or enable those with existing integrated systems to evaluate them to determine future courses of action.

Highway Reference System

All maintenance activities and operations on a highway system are referenced by their locations in the road network. Hence, location is a common attribute used by management systems to associate information such as traffic volume, pavement condition, accident rate, degree of curvature, railroad crossings, and so on. The most common reference systems used by highway agencies are the route-milepost system and the link-node system. However, specific units in many highway agencies are using other location reference systems that suit their needs. The Michigan Department of Transportation, for example, has identified 38 reference systems, classified into coordinate, area identifier, segmental location (length), and point location systems, which reflect the varying degrees of detail and the levels of aggregation used by different units in the department. A study to determine the standard to use for an integrated system in Michigan has recommended the control section-milepoint method which is a variation of the route-milepost system [Briggs and Chatfield, 1987].

The route-milepost system uses the route number assigned to different roads in the network as the key or primary reference; the mileage or milepost is measured from a specific point on the route and used as secondary attribute. A modification of the route-milepost system uses an inventory number to identify the road segment. The link-node system, in

contrast, defines a link or segment of road as bounded by two nodes at its ends. The nodes may represent intersections or points of transition from one link characteristic to another. While the route-milepost system uses points as references, the link-node system uses lines as references.

More states have adopted the route-milepost system because it is easier to convert field locations in terms of mileage. However, the link-node system may be more suitable for some situations. For example, in traffic engineering, route-milepost is ideal for calculating time mean speed, but space mean speed is the average speed over a given length, for which the link-node system is more appropriate. There are many examples in pavement management, network equilibrium problems, bridge management, and other applications for which both systems are required. It will be ideal for an integrated system to use the route-milepost system to define specific nodes on the road network and then establish linear relationships between these nodes to define the links. A management and information system that provides interface between the two systems would allow conversion from one to the other and would resolve this problem. The other alternative is to use a data structure that can associate or link two nodes. This concept will be discussed in the conceptual models of databases.

A coordinate reference system, which underlies geographic information systems (GIS) and uses geodetic measures such as latitude and longitude, is used by the United States Geological Survey (USGS) and other national and private agencies to identify the location of highways, bridges, rivers, and other natural and man-made features on the surface of the earth. A coordinate reference system can provide a connection between local, statewide, and interstate network analyses to achieve not only national transportation goals but also those related to social, economic, and environmental data that are referenced by geographic location. The level of precision offered by GIS in locating specific points of interest on the road depends on the resolution of the digital maps and their inherent accuracy. The resolution of many digital maps is very coarse—for example, at the 1:100,000 scale. Resolution at the USGS quad scale (1:24,000) is approximately 20 meters, but quad maps have proven to be of questionable accuracy for more detailed highway applications. Aerial photographs of higher resolution and accuracy can be digitized and converted to base maps in GIS. The creation of highly accurate cartographic base maps for

GIS generally depends on the improved mapping technology, especially those based on satellite global positioning systems (GPS). GPS provides three levels of accuracy: one for military use (< 15m, 95 percent of the time); one for civilian use that is similar to military accuracy but currently no longer available; and one for civilian use that is degraded (100 m, 95 percent of the time). The last level is called selective availability. Even with the most degraded satellite transmissions under selective availability, accuracies on the order of one to five meters can be achieved using a GPS receiver in conjunction with a base station located at a precise point and with differential processing. Indeed, satellite surveying, which requires fairly long dwell times at the survey site, can achieve accuracies that exceed 1 cm. Levels of accuracy of one to five meters are more than sufficient for most road and bridge management applications. Kinematic (mobile) GPS data collection methods recently demonstrated in the AASHTO 38 State GPS/GIS project revealed that roadway alignment and many other inventory features could be located at accuracies of two to five meters at highway speeds of roughly 55 mph. Expensive technology was used that may not be cost effective in many states, but the principle that GPS can be used to develop a highly accurate cartographic database describing the roadway network has been established. Accurate updating of roadway feature inventories can be accomplished fairly inexpensively with a GPS receiver, a supplementary base station, and differential processing.

Database Organization

We have identified the reference systems that can be used for linking highway information data files. It is now necessary to identify the groups of files that will constitute the integrated database and how they are linked in the system. Highway agencies should maintain a database of at least the following information for their road networks: road inventory, traffic volume, traffic accident, maintenance, finance, and structure.

Road inventory data contains the physical and geometric characteristics of the section of the road. These include its functional classification, construction parameters such as pavement type and thickness, aggregate or disaggregate condition of the pavement, number of lanes, alignment, and curvature among others. Some states have developed a separate pavement database because of the large portion of

highway funds allocated to pavement management compared to other highway maintenance and management operations. Traffic volume data are obtained from traffic counts and surveys and weigh-in-motion equipment, which collect information on road use and composition of traffic. Traffic accident data are usually reported and recorded by motorists and the police. These data include detailed information on accidents such as number of fatalities and injuries, number and types of vehicles involved, and the factors contributing to the accident such as weather conditions and the nature of the roadway. Maintenance data refer to types, extent, and cost (labor, equipment, materials) of maintenance performed on the section of the road including periodic, routine, improvement, and emergency maintenance operations. Maintenance includes snow removal, pavement overlay, bridge deck replacement, sign rehabilitation, grass cutting, and other actions. The maintenance database can provide highway agencies with past and present information associated with the maintenance of the highway and its appurtenances. Finance data, such as the type used by Pennsylvania Department of Transportation, link improvement, maintenance, and other funds to specific locations on the road. These data enable the state agency to allocate or attribute costs and benefits to each road section to determine the economic returns from different types of improvements and maintenance. Finally, structural data include inventory of, and management information related to, bridges and railroad crossings that are part of the road network.

Highway agencies in different states maintain different sets of databases as a result of management systems not being uniform across states. For example, most states do not have a finance database. It should be emphasized, however, that each state is in a different stage or level of integration and variability in organizational structure among highway agencies makes it difficult to define a standard database organization that will suit the needs of all states. A typical existing integrated highway database consists of roadway data, traffic data, accident data, structural data, and physical conditions data as shown in Figure 1. Efforts should be made by states to link the other files identified previously to their current highway information systems.

The linking of data files in the IHIS databases is performed by a file linkage system that, for most states, consists of three functions, namely (1) a database management function that links data from several files, (2) a report function that produces

Highway Information System

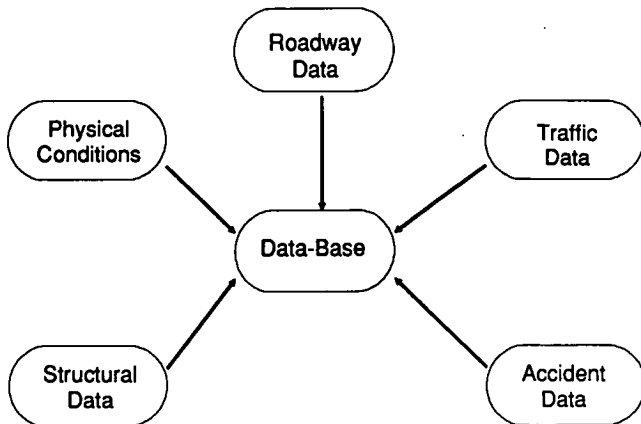


Figure 1. Typical IHIS databases (Source: FHWA)

descriptive reports and tables, and (3) a statistical function which yields comparative statistics [FHWA, 1987]. Again, a reference system is required to link the data files or the records across data files. Many user-friendly programs that are menu-driven have allowed users with very little computer skills to perform specific tasks, such as preparing reports and summaries obtained from the database. Other database management systems can create graphs or digital maps for data representation. The interface between DBMS and application programs is discussed later in this digest.

Database Models and Structures

The database manager allows the user to retrieve or store data in the database. The user usually requests information or access by a functional call to the database manager. The overall configuration of the information in the database is a conceptual model or schema, but the database manager frees the user from knowing the exact details of how the data are stored and manipulated [Tsichritzis and Lochovsky, 1977]. What the user sees is a view of the data. The different models used to represent data in the database are flat file, relational, network, hierarchical, and object-oriented.

Flat Files. Highway data in the 1960s was formatted using card image files, otherwise known as flat files. These files use a standard data format consisting of two logical units of data: the data item or attribute, which is the smallest logical unit of data;

and a record, which is a collection of data items. The flat-file model specifies that data are to be organized according to records and data items, and each data item in the record or punched card is assigned to a column or field (Figure 2). This file system of data management is obsolete as magnetic tapes and disks have replaced punched cards as storage media. It does not have the efficient management capabilities of existing database systems because each file or record is independent of other records. Most programmers, though, retained their notion of data as being grouped into flat files and many application programs still require data to be organized into flat files. Modern world database models include relational, network, and hierarchical models which, like the flat file model, group data into records and attributes but allow associations and interdependencies among records and data items.

Relational Model. In a relational database model, data are represented as relations or tables. Each relation is a record type with attributes (columns) representing the data elements in the entity set (record) and a series of rows or tuples corresponding to the attributes (Figure 3). Every relation has a unique identifier or primary attribute (key) that differentiates the relation from other relations. Other attributes may also be used as keys to associate the relation to other relations. These secondary keys are different from the primary key in that they may assume similar values or have no values. Database management systems using the relational model (also called RDBMS) allow the user to search and update the data through a relational data language such as DSL (Data Sub-Language) Alpha or SEQUEL (Structural English Query Language) [Tsichritzis and Lochovsky, 1977]. RDBMS is the most common

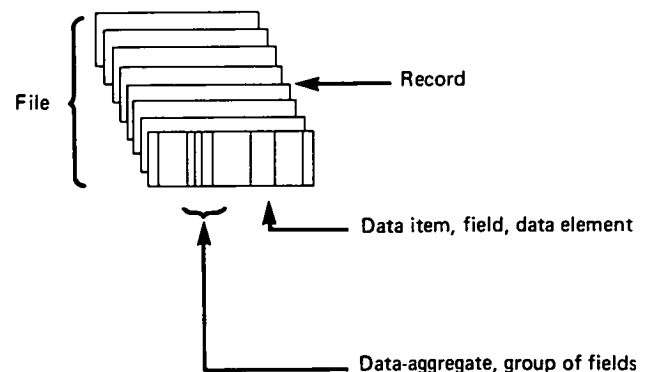


Figure 2. The flat-file database model (Source: Martin, 1983)

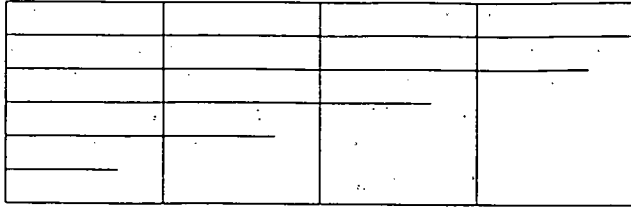


Figure 3. *The relational database model* (Source: ORACLE)

DBMS product environment in the market today. SEQUEL is the basis for the relational database language SQL (also pronounced "sequel") that has been adopted as a standard language for RDBMS (see discussion on Emerging Standards).

Network Model. The network database model uses specific links to associate attributes of the record types. These links (or arcs) allow search or data selection across linked record types. The record types and arcs are represented as nodes and links, respectively, in the model representation, as shown on Figure 4. The languages used to operate on network data models are based on those proposed by the Data Base Task Group (DBTG) of the Conference on Data Systems Language (CODASYL) and include DML for COBOL and EDBS (Educational Data Base System), which uses the APL language. Another system, called Adaptable Database System (ADABAS), resembles a relational system but can be considered a general network system.

Hierarchical Model. The fourth type of database model, called hierarchical or tree model, is a special case of the network model. In the hierarchical model, record types are also represented as nodes and the associations as links but there is an ordered relationship between record types (Figure 5). One node or record type is identified by its root node and the other nodes above it in the tree (hierarchy). Unlike in the network model, a record type in the hierarchical model cannot exist independently unless it is a root node, and a search for a data element in

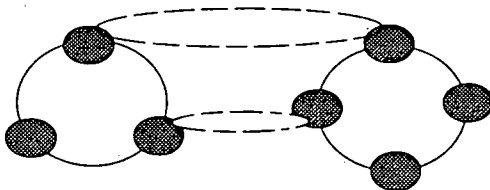


Figure 4. *The network database model* (Source: ORACLE)

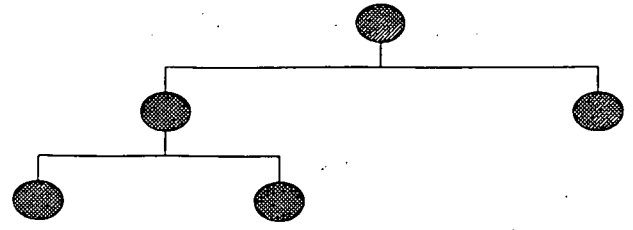


Figure 5. *The hierarchical database model* (Source: ORACLE)

each record has to go through a unique path defined by the hierarchy. In a sense the hierarchical model is a restrictive model because it does not allow deletion or addition of record types in the hierarchy without affecting the other nodes. Examples of systems that use the hierarchical data model are the Information Management Systems (IMS) developed by IBM and SYSTEM 2000 (NATURAL), which is an "English-like," interactive query language [Martin, 1983]. Network and hierarchical database management systems are called procedural systems because the structures of the database models determine how data processing is to be done.

Object-Oriented Model. The object-oriented database model is a recent scheme for representing data. It defines an object as a data set or entity consisting of not only attributes but also procedures. This model allows different levels of data abstraction by building objects from other objects (Figure 6). Inheritance is specified by defining an object as an instance of its class. More complex data structures can be built or added to existing data structures. Therefore, an object-oriented database is a complex structure consisting of not only the data but also the procedures with which to share, link, or operate the data. There are only a number of object-oriented database management systems (OODBMS) vendors to date. The most popular object-oriented languages are C++ and Smalltalk. The object-oriented concept is fairly new, and there is still a lot of research going on in this area.

```

Object Road (
  int route_number; /* declaration of object road */
  char func; /* the route number of the road */
  object segment { /* functional classification */
    coord* leftnode; /* segment attribute of road */
    coord* rightnode; /* coordinate of beginning milepost */
  } /* coordinate of the ending milepost */
  char pavmt; /* type of pavement */
}

object coord {
  int longitude; /* a point object represented by coordinates */
  int latitude; /* geodetic vertical location */
} /* geodetic horizontal location */

```

Figure 6. *The object-oriented database model of a road*

The IHIS Database Models. Existing database management systems used by different states differ because of the way the data are stored or recorded and the system used to link these data. Each unit in the highway department may maintain its own database, using its own data format, which is most convenient for its purpose or which has been used for many years (like the card image format that is still used by many agencies). The complexity of developing a uniform integrated highway information system for adoption by all units within the highway agency, and perhaps all state highway agencies, lies in the definition of a standard database model that would link data files including roadway inventory, traffic volume, traffic accident, maintenance, finance, and structural data by using reference keys or attributes like route-milepost, link-node, or coordinates. Whatever form this model and database management system takes would require changes in the way the individual units within the organization record and manage their data.

IHIS and Applications Programs

Various application programs are used by units and divisions within the highway agency for their specific purposes. In an integrated information and management system environment, these programs must be able to draw, process, and store data in the database. As explained earlier, the database manager allows the application programs access to the database. The application programs use predefined function calls to request information from the database just as a user of the DBMS would communicate directly with the database manager. The different types of applications that exist within the highway agency and their relationships to the other components of IHIS are shown in Figure 7. Every application is an information and management system in itself, but is limited in scope compared to the IHIS. Access to one or more databases may be required by each application, and the database manager should be able to provide the information required by the application. The application programs perform independently from the integrated DBMS and can do anything to the data such as create a summary report or generate maps or graphs for the user. Therefore, the database management system not only links databases for highway information management (reporting, etc.) in the network level but also provides the interface for retrieval, storage, updating, and processing of information by external applications. In this way the

INTEGRATED HIGHWAY INFORMATION SYSTEM

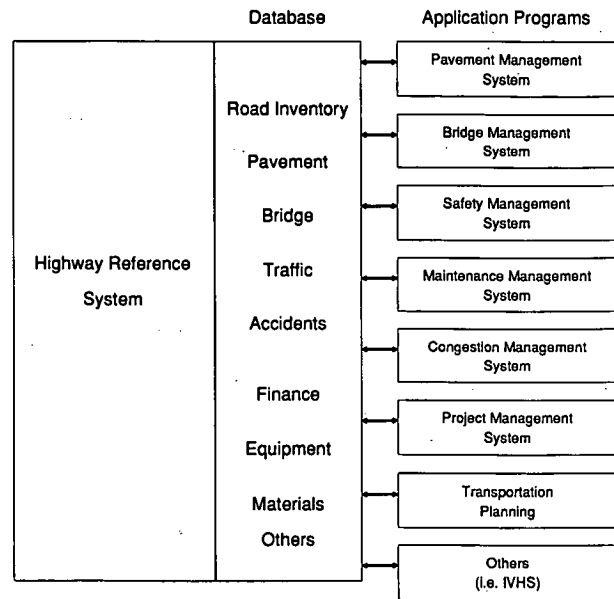


Figure 7. *IHIS application programs* (Source: Paterson and Scullion, 1990)

DBMS serves as an interface to all the application programs that operate on a common collection of data sets and files. The result is very efficient data storage and management.

The software and hardware environments with which DBMS operate and interface with users and application programs can be personal (often called flat-files because they usually operate on one record at a time, which also makes them slow); micro-computer (allows multiple files to be tapped at one time and also the development of custom applications); host (usually mainframe and minicomputer systems that serve several end-users at a time who might be working on dissimilar computers and operating systems); and network (provides client-server functions under a robust network operating system, and data management services to client workstations using a variety of operating systems). Data communications at the network level are required for integrated highway information systems because the information is distributed to several end-users who are using a variety of application programs on different computer operating systems at various locations (Figure 8).

Data Communications and Telecommunications

Access to database management systems by end-

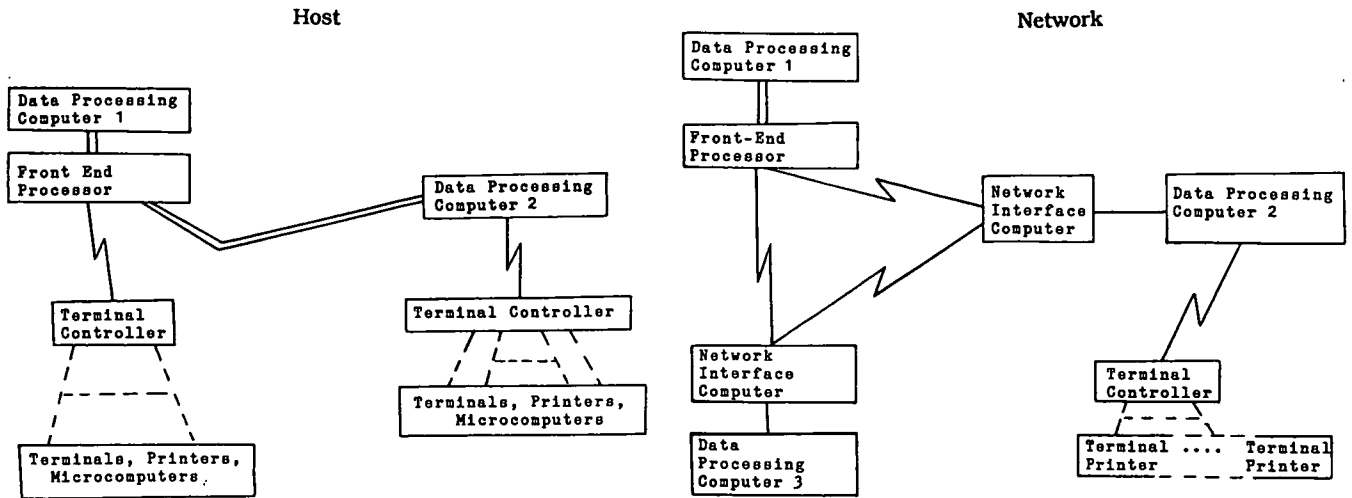


Figure 8. Host versus network database server (Source: Lientz, 1988)

users and application programs is usually done through a computer interface. The user may access data directly from a microcomputer or host computer in which the database resides. However, several users may require access to the same information at the same time, as is the case with an integrated highway information system. A computer network is a system wherein multiple computers share the resources of all computers among various users that are connected to the network via a high-speed data communications link [Lientz, 1988]. A computer network is usually controlled by network control software that resides in the central computer services office (CCSO). In a centralized database management system, all the information is stored in the CCSO and each user has to share the same data with other users. This system is ideal for IHIS. However, several difficulties are associated with a front-end centralized information system including slow access rate resulting from a large amount of information, high update rate by application programs and end-users, and issues related to privacy and security of data. Databases can be decentralized by grouping them in a way that will require the least adjustments from different users and application programs. This is the basic concept of distributed databases and architecture. The computer network provides the facility for databases stored in separate locations to be accessed by different users working with dissimilar computers in different places and allows these users to communicate with each other or transfer information through the network.

Computer networks are classified in many ways. Depending on the area of coverage, networks can be classified as local area or wide area. Local area networks (LAN) provide service to users in a building or campus environment. These networks are usually privately owned by business organizations or universities. An example of a local area network would be the computer network used in the city office to allow different users in the building to access and transmit information within the organization.

Wide area networks (WAN) encompass a larger geographical setting allowing telecommunications across local boundaries. This type of network may be privately owned (e.g., phone companies) or publicly owned, which is also another way to classify networks. While there is more flexibility in network architecture and protocol for privately owned networks, public domain intelligent networks have been more or less standardized. The exchange of information between private and public information vendors has been difficult because of the diversity in network architectures. It is the incompatibility among various network architectures that led to the development of the open systems interconnection (OSI) reference model, which will be discussed later.

Computer networks may also be classified according to their switching and transmission module. This function specifies the medium to use in transmitting multiple forms of information (data, voice, video images) in two directions. The specification includes the path and circuitry of transmission and can be classified as *switched*, *packet-switched*, or *channel-*

switched. A detailed discussion of these architectures is provided [Ahamed, 1988].

At present, neither wide nor local area IHIS computer networks have been implemented at the state level. Existing IHIS environments and database servers use host computers within the department that can only serve a number of end-users and application programs. However, a contract was recently approved for implementation of a Value-Added Network (VAN), which will provide connectivity between the mainframe computers of participating AASHTO member departments including the U.S. Department of Transportation and the Transportation Research Board. VAN will feature an electronic mail system, access to AASHTO bulletin boards and publishing services, access to certain databases maintained by participants and selected vendors, and a

wide variety of potential applications including a national Comprehensive Highway and Planning System (CHIPS), which seeks to integrate databases that support full infrastructure management systems, GIS, and video logging.

Current and Emerging Standards for Data Interchange and Their Implications in BMS and IHIS

In many transportation agencies, technical data interchange standards are set by an uppermanagement chief information officer, and those decisions become the basis for a wide range of policies, procurements, and technical decisions that affect bridge management systems (BMS) and practically every other agency data processing application. Although it is beyond the scope of this report to recommend or mandate agencywide data processing policies, it is helpful to describe the industrywide standardization efforts now underway and to show how these events influence the implementation of BMS and IHIS. This section describes some of the current and emerging standards for data interchange at the national and international level. These standards concern models and procedures for data representation, database languages, data transfer and communications, and computer hardware/architecture, all of which are relevant to the implementation of integrated information and decision support systems for highways and bridges. The organizations involved in setting these standards are identified first.

Standard Setting Organizations

The National Institute of Standards and Technology (NIST) is a government organization that, among other things, issues standards for equipment sold to the federal government. Standards developed by NIST for processing and exchange of information are collectively known as Federal Information Processing Standards (FIPS). Another standards organization in the United States is the American National Standards Institute (ANSI), which, unlike NIST, is a private sector standards coordinating center. Its functions include identifying industrial and public needs for national consensus standards and coordinating private sector standards activities and development. ANSI is the recognized U.S. member of International Organization for Standardization (ISO).

ISO is an international agency whose function is to define standards on a broad range of areas. Its pur-

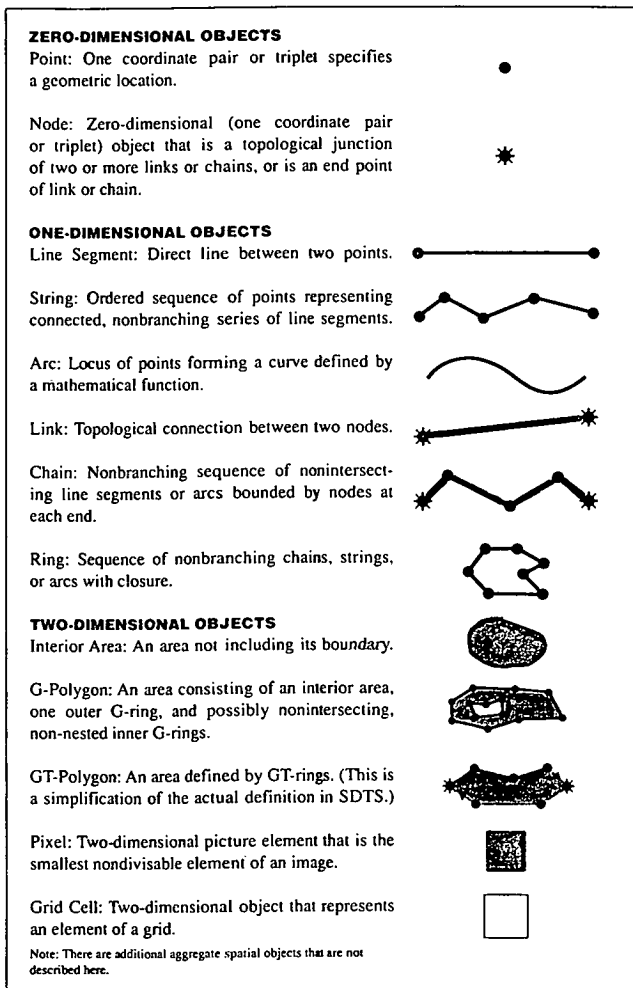


Figure 9. Definition of SDTS data objects (Source: Tosta, 1991)

pose is to promote the development of standardization and related activities to facilitate international exchange of goods and services and to develop cooperation in the sphere of intellectual, scientific, technological, and economic activity [Stallings, 1990]. Another international organization dealing with standards is the International Telegraph and Telephone Consultative Committee (ITTCC). Its goal is to standardize, to the extent necessary, techniques and operations in telecommunications to achieve end-to-end compatibility of international telecommunication connections, regardless of the countries of origin or destination.

These are some of the standard setting organizations involved in specifying various components of information systems and data communications protocols, which can be applied to the IHIS environment.

Standards for Representing Spatial Data

The general concepts of database management systems and the existing database models that can be used to implement an integrated highway information system were discussed under "Components of IHIS." However, many highway agencies have developed their own information systems using one or more database models. Efforts have been made at the national level to define a uniform standard with which all spatial data can be represented and transferred. For highways, these spatial data include links and chains to represent highways, sections and interconnections of highways, and two-dimensional features such as area. The USGS has drafted Spatial Data Transfer Standards (SDTS) as a Federal Information Processing Standard [Tosta, 1991]. These standards use the object-oriented database model to represent real world entities that can be transferred among and interpreted by dissimilar computer systems. The objects are called spatial objects and consist of points, nodes, lines, arcs, strings, chains, rings, pixels (picture element for images), grid cells, and other two-dimensional entities (Figure 9). These objects represent spatial phenomena, such as highway, pavement, road, and other geographic features that can be referenced to National Geodetic Control Points or the Topologically Integrated Geographic Encoding and Referencing (TIGER) System developed by the Bureau of Census. The SDTS can be used on the highway agency level in representing roadway inventory and other spatial data. The implementation of SDTS in the ANSI/ISO 8211 format (which specifies the actual encoding of

data fields and subfields on the transfer media) is also included in the USGS draft proposal.

Standards for Computer-Aided Design and Drafting

Computer-aided design and drafting (CADD) is widely used in many transportation departments. CADD application programs are regularly used to produce maps, plans, and specifications including supporting engineering calculations. Many agencies are also using CADD programs for geographic information systems. CADD tools frequently operate within a client-server environment where the host database operating system, typically a mini or mainframe computer, serves up selected information and the programs perform their operations upon it. Some data may be stored in dedicated, user-oriented databases. An industry group of CADD developers has created a standard for CADD databases, known as Initial Graphics Exchange Standard (IGES). This standard uses the ASCII file format and serves as a neutral database format for exchanging and processing graphics entities. Many CADD software systems (e.g., AUTOCAD) have routines to convert their proprietary database structures to IGES and vice versa. The standard has also been adopted by ANSI

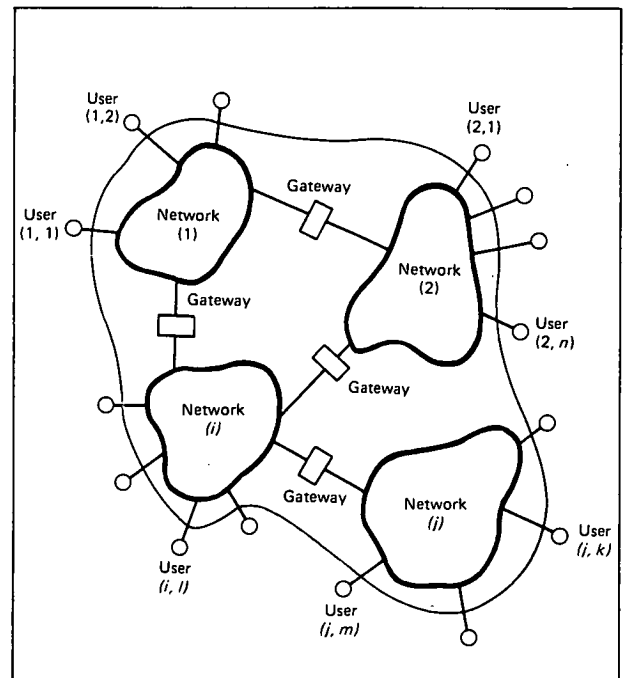


Figure 10. *Conceptual representation of the open system interconnect (OSI) environment* (Source: Ahamed, 1988)

Table 1. The different layers of the OSI model (Source: Ahamed, 1988)

OSI layer	Type of functional intelligence ¹
1. Physical	PHL: Activation, maintenance, and deactivation of the physical connection. The electrical and mechanical characteristics for physical interface and the transmission media. (See ISO 2110, CCITT V.24, V.28 or EIA RS-232-C, also EIA RS-449, CCITT X.21, V.35; EIA RS-422A, CCITT V.11, X.27 for balanced voltage; EIA RS-423A, CCITT V.10, X.26 for unbalanced voltage; CCITT I.430 for (2B + D) ISDN and CCITT I.432 for 24 B or 30 B channels.)
2. Data link	DL: Synchronization/framing, error detection and recovery, and flow control for information transmitted over the physical link. (See ISO DIS 8886, 1745, 2111, 2628, 2629; CCITT X.212, X.21 for basic mode; ISO 3309, 4335, 6159, 6256; CCITT X.25, X.75, X.71 for HDLC; CCITT Q.920/Q.921 for ISDN.)
3. Network	NL: Establish, maintain, and terminate switched connections. Addressing and routing functions. Service TL, independent of DL and PHL. (a) Connection: network connection, data transfer, optional expedited data and receipt transfers, reset, and connection release. (b) Connectionless: UNITDATA. (See ISO 8348, CCITT X.213; protocols CCITT X.25, 1984, packet. ISO DP 8878 with X.25 for connection oriented network service; ISO 8473 for connectionless internetworking; CCITT Q.930/Q.931 ISDN.)
4. Transport	TL: Selection of network service. Evaluation of the need for multiplexing. Selection of the functions from the lower layers. Optimal data size decisions. Mapping of transport addresses to network addresses or the end-point users and negotiated. Data flow regulation between the end users. Segmentation and concatenation. Error detection and its recovery. (See ISO 8072, CCITT X.214; ISO 8073, ISO DIS 8602-connectionless, CCITT X.224.)
5. Session	SL: Provision to transfer data and to transfer control in an organized and synchronized manner. User may define the degree of control and synchronization that the session layer will provide. (See CCITT X. 215, ISO 8326; X.225, ISO X.8327, T.62.)
6. Presentation	PRL: Assures the delivery of information to the end users in a form that is usable and understood. The information content (semantic) is retained, even though the presentation format and language (syntax) can be altered to suit the source(s) and destination(s) of the information. (See ISO 8824 for abstract syntax notation ASN.1, and ISO 8825 for encoding, CCITT X.409.)
7. Application	AL: Facility to serve the end user. Provision of the distributed information service. Communication management between the AL and PRL. Service to the user, the application (SASE), and the group of applications (CASE) served. Authentication of user IDs, destination IDs, authority to exchange information. Determination of service quality from the lower layers. Data integrity, error recovery, and file transfers are also assured. (See ISO DIS 8649, protocol ISO 8640, CCITT X.400 message handling system.)

¹The nature and amount of intelligence required at each level and the established standards for specifications and protocol.

as a standard data structure (ANSI/SAE J1881-AUG88).

Another *de facto* industry standard, which is especially important for graphic representations of spatial data, is DXF or Data Exchange Format. This format was originally the proprietary format of a particular CADD developer, but most of the major Geographic Information System developers now offer the capability to transfer data to and from this format.

Standards for Database Languages

The relational database model and the languages used to operate on the relational databases were described under the section titled "Relational Model." ANSI has adopted SQL as the standard language for relational database management systems [ORACLE, 1988]. This language has also been adopted by the ISO and by the U.S. Government as a Federal Infor-

mation Processing Standard, which made it the universal database access language. Because of the widespread use of RDBMS and the adoption of SQL as its standard language, programmers and vendors of database management systems have created an object-oriented SQL called Object SQL, which can manipulate object databases [Loomis, 1991]. Object SQL can be modified to process relational databases as well, but building such a hybrid language that can process both relational and object databases is as complex as what relational database programmers are doing to create relational models that will have the functionality of object models. There is no standard yet for object-oriented database languages; however, proposals have been presented to ANSI by industry groups to make Object SQL the standard. Indeed, bridging the gap among database models is still an open area of research. This issue will be discussed in the section on "Conversion, Migration, and Integration."

A strong industry standard that competes with SQL is the xBase language, which is used in a plurality of personal computer database applications. Developers of xBase products have maintained the viability of this standard by encouraging a strong third-party market of books, development tools, and off-the-shelf applications, as well as maintaining a strong speed advantage over personal computer SQL implementations. One sign that the xBase standard may be weakening in favor of SQL is the fact that several

major xBase developers are now offering SQL translation facilities in their systems.

It should be noted that the choice of database language standard implies nothing about the underlying physical storage format of data. Most database management systems have their own proprietary file formats. One important feature offered by SQL and practically all other database languages is the ability to translate data from the proprietary format into ASCII, a human-readable standardized text format.

Standards for Data Storage and Retrieval

Computer memory is an array of randomly accessible data identified by a unique address. Computer main memory systems include read-write memory (RWM), read-only memory (ROM), and variations of these two types. RWM allows data to be stored at any address; it also allows the stored information to be read back at any time. In contrast, data in ROM are only stored once (when the ROM was created), but the contents of any address can be read any time. The advantage of ROM over RWM is that it is non-volatile; that is, the contents of the memory are not erased even if the computer power is removed. However, ROM types have been modified to allow data to be stored and erased.

Compact disk ROMs (or CDROMs) are cost-effective media for robust mass data storage (however, semiconductor memory may become less expensive in the future). The ISO has adopted a standard file structure for CDROMs called ISO 9960. This standard has a hierarchical file structure and is now generally used on most CDROM disks sold. The standardization of CDROM disks can make the CDROM medium compatible with a wide variety of application and computing platforms.

Standards for Specific Transactions

Transaction standards comprise an important class of data exchange standards. A transaction standard pertains to the transfer of a class of data used in a specific type of transaction; for example, shipment orders between a group of manufacturers and a group of material suppliers, or the transfer of funds among financial institutions. Various organizations under the auspices of ANSI, such as the X12 Committee of the Data Interchange Association, develop a wide variety of data transaction standards primarily for industry, but also for government. These standards are characterized by agreed-upon calculations or operations

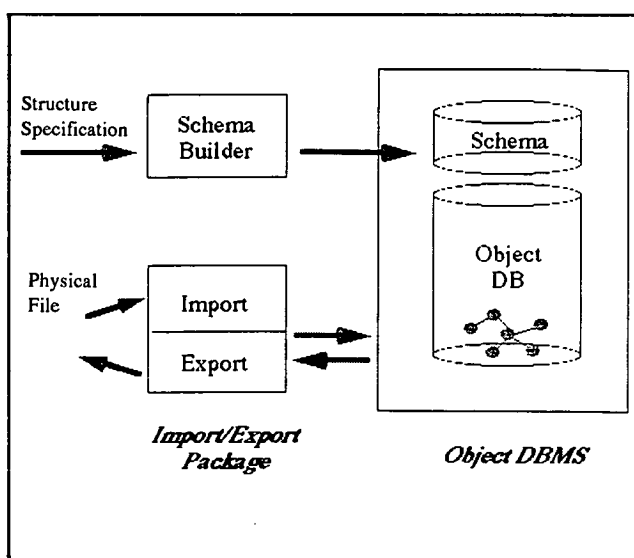


Figure 11. *Import/Export mechanism* (Source: Loomis, 1991)

automatically performed on a certain type of data exchanged between two or more parties. Thus, transaction standards are concerned not only with data interchange but also standardized operations on the data transferred.

One way to standardize the transfer of certain types of data from one related highway database to another is through a series of transaction standards. Examples of highway-related transaction standards are the standardized format and procedures for annually transmitting Highway Performance Monitoring System (HPMS) data, motor vehicle accidents (ANSI D16.1-1989) data, and the National Bridge Inventory (NBI) data to the Federal Highway Administration. Transaction standards might also apply to data interchanged between highway-related databases within a state transportation agency. For instance, the most recent traffic count, classification, and weigh-in-motion data might automatically be adjusted to the current year and then automatically transferred to the databases driving bridge and pavement management systems. A more elaborate transaction standard might also involve calculation and transmission of indicators of traffic data quality such as mean and variance and sample average, frequency and size.

Standards for Pen-Based Systems

Recently, a new breed of portable computers—an electronics hybrid of a clipboard and a drawing tablet—has been introduced into the market and is being rapidly commercialized. The first generation of these new computers, part of a class called pen-based systems, is intended to replace paper forms by capturing and storing data as soon as it is entered, so that the data can be directly transferred to a host database. Pen-based systems are designed for indoor and outdoor field data collection that can feed information into an IHIS database and can potentially be used for data retrieval in addition to data collection. Some models have add-ons that permit real-time data exchange using radio frequency data communication (RFDC) links.

The current generation of pen-based systems permits one to enter data with a pen by handwriting, by touching the pen tip to a pick list of options, by drawing, or by furnishing a script annotation or signature. Essentially two types of data are captured: alpha-numeric fields or digitized imagery. The first generation of pen-based systems has a handwriting recognition algorithm that can identify block letters and numbers. Alpha-numeric data entry may also be

achieved by selecting an option displayed on the screen. Digital imagery in the form of sketches, photos, signatures, and script annotation can be stored separately.

In recognition that pen-based systems represent the next major evolution of portable data collection and computing, Go Corporation has developed an operating system for these types of computers [Landry, 1991]. This operating system is likely to become the *de facto* standard for pen-based systems in the United States, although Microsoft Corporation has introduced a competitive operating system. Both operating systems address all normal file handling capabilities, including the processing of data types and formats, which pen-based systems capture, export, and import.

Standards for Networking

The open system interconnection (OSI) reference model mentioned earlier was developed by the International Standards Organization (ISO) and is now widely used in the network marketplace. The goal of the model is to define a standard architecture for networking. The protocol for the OSI consists of 7 layers or modules that are arranged in a hierarchy based on increasing levels of abstraction. Each layer performs a function that is dependent on the more elementary lower layer. The system is "open" in the sense that the layers are not tightly specified, which would allow different networks to interconnect or link easily (Figure 10). This model is now followed closely by many networks. The elements of the OSI protocol from lowest to highest level are as follows [Lientz, 1988]: (1) physical—transmits information bits across the network; (2) data link—transmits data frames across network links; (3) network—supports connection of multiple network links; (4) transport—transfers data along a complete network path from an origin to a destination; (5) session—supports communication between application systems; (6) presentation—provides for application systems to be independent of the form and representation of the data; and (7) application—provides network access to application systems. Existing protocols are used in networking based on the first three layers of the OSI model. For example, X.25, which is a standard for packet switching, appears in layers 2 and 3. Table 1 shows how various telecommunication standards fit within the OSI model.

For the integrated highway information system, layers 1, 2, and 3 define the functions that will be used by network providers, such as the U.S. Depart-

ment of Transportation or a state highway department, to connect or link the computer networks used by different highway agencies in order to transmit information within states (e.g., among districts) or among states. Layers 5, 6, and 7 represent the functions used by the network users and application programs such as pavement management systems (PMS), bridge management systems (BMS), and traffic management systems including IVHS. Finally, layer 4 serves as a medium or liaison between the U.S. DOT and a particular highway agency and the users of application programs within that agency.

Conversion, Migration, and Integration

Earlier discussion pointed out how difficult it is for highway agencies to adopt uniform standards for data interchange among information systems because of the differences in their data organization and management. Unfortunately, highway administrators have to live with the reality that information systems exist in the operational and tactical levels of organization and a top-down approach to data management will be met with obstacles and resistance. The role of the highway manager, then, is to coordinate with different units of the organization and the highway network administrators, in order to successfully implement the integration process. Two approaches to evolve from the existing independent file systems to an integrated database world are: (1) convert old systems to the standard database models, or (2) continue the existing file system but build a bridge to the database system [Martin, 1983].

Each approach has specific advantages and disadvantages and should be assessed based on the current information systems used by each state. The options identified above can be used to integrate highway databases or enable the IHIS to work on dissimilar data structures. For example, all data in flat file format can be converted to relational, object, or any other format chosen as the standard.

The simplest standard is conversion to the lowest common denominator. If data files contain only alpha-numeric fields, regardless of whether they reside in flat, relational, hierarchical, or object databases, they can be converted to universal standard ASCII format. An even lower common denominator is to convert all data to bit streams, with each data item (field, object, etc.) containing a header and a trailer as in the ANSI/ISO 8211 format. However, this is computationally inefficient. Higher level

interfaces are more desirable because they are easier and faster to process.

Conversion to one universal standard database format is desirable only if all application programs that access the databases are also to be rewritten using one database language, or if individual units within the organization have enough resources to modify or buy new application programs. In many cases this is not an economic alternative.

The other option is to create a higher level neutral interface through an import/export mechanism that converts one database structure to another, depending on the programming language used by the application program. Figure 11 shows such a mechanism for an OODBMS that has to work on a non-object database. A higher level import/export mechanism serves as an interface among different programming environments but still does not integrate them.

An even more ideal integrated system allows application programs to process the databases directly, regardless of the database structure. This means that a program written in C++ or Smalltalk can process relational databases as if they were object databases. An example of such a tool is the SQL Gateway which allows programs written in an object language to access a relational database as if it contained objects instead of tables [Loomis, 1991]. This tool converts Smalltalk and C++ messages to equivalent SQL statements and converts tables into objects and vice versa. Figure 12 shows how the SQL Gateway links an object application to a relational database.

The foregoing techniques identify some examples of the directions that a highway agency can take in order to migrate from its existing highway information system to a more integrated data management system. Again, the agency has to evaluate and choose, based on its objectives and resources, the most economic alternative. Each implementation alternative has various costs associated with hardware, software, personnel, and other technical and administrative requirements.

Candidate Standards for Data Interchange

The OSI model shown in Figure 10 and Table 1 can serve as the framework for standardization of data interchange within and between various transportation agencies. Specific data interchange standards that pertain to programs driving infrastructure management systems would lie within the application, presentation, and session layers of the OSI model.

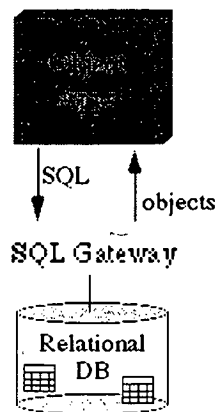


Figure 12. *The SQL gateway* (Source: Loomis, 1991)

Application standards will specify how the users of IHIS and application programs can access the network facilities and databases. The application process is characterized by user elements accessing OSI application services. These application services may be common or specific, depending on the application process. Examples of common application services are establishing, maintaining, and terminating network connections, while specific application services include file transfer, job transfer, message exchange, or remote terminal access. For example, a user of IHIS has to identify himself through a user account and password before gaining access to the system or an application program. Standards that can be adopted for this layer may include those identified in Table 1.

Standards within the presentation layer of the IHIS pertain to those that specify how users and application programs would exchange or transfer information (e.g., standardized procedures that states use to send National Bridge Inventory data to FHWA or that a planning unit in an agency uses to send digitized maps to GIS application programs). These standards might take a number of forms; for example, those that identify the syntax or structure to represent highway data (e.g., IGES, ISO 8211, TIGER, SDTS); those that identify the semantics or languages that can be used to process the data (e.g. SQL and Object SQL); and, those that specify the functions to use in processing data by different application programs using dissimilar data structures (e.g., SQL Gateway, Import/ Export Mechanism, Conversion Routines).

Standards that would support communication between application systems are required by presentation layer functions. These can be found in the session layer and might consist of standardized

linkages among databases (e.g., requiring that a standard type of location reference point be attached to all data that have a spatial attribute). These upper layer entities are illustrated in Figure 13 for two IHIS application entities.

Standards for system architecture and other communication protocols would occupy the lowest four layers. An example of this is the widely used X.25 packet-switched network standard.

Interfacing a BMS with Agency Standards

This digest has concentrated on industry events and trends of a technical nature, which impact the agency policies and standards within which a BMS must be implemented. Transportation agencies who wish to develop their own BMS should explore these issues of standardization as part of the choice of platforms, tools, and languages for the system. Agencies which choose to acquire and customize either the recently completed FHWA Pontis BMS or the system now under development by NCHRP do not necessarily have to grapple with these issues, but they do have to arrange for interfaces between these systems and those that already exist within the agency.

One strategy for building this interface is to set up a client-server database which uses standardized telecommunications hardware and software to link the state's DOT headquarters with district offices and perhaps even local governments. The bridge database and related geographic, traffic, and other data may reside on a mini or mainframe computer in the headquarters office, while the main BMS software resides on personal computers and engineering workstations statewide. An SQL-based client-server database manager automatically handles the transfers of data among the systems. In this kind of system, local BMS users can electronically update their local databases on demand or on a regular schedule by placing an electronic request to the server. The server extracts relevant data from its various files and sends these to the local computer. Periodically the server can handle computationally-intensive operations, such as updating bridge deterioration models, and send the results to all of the client computers. To update the master database, clients may send new inspection data to the server, which performs automatic quality checks, merges the data with what has been received from other clients, and schedules an opportunity for a human inspector to review the results. After this, the data are posted to the master database.

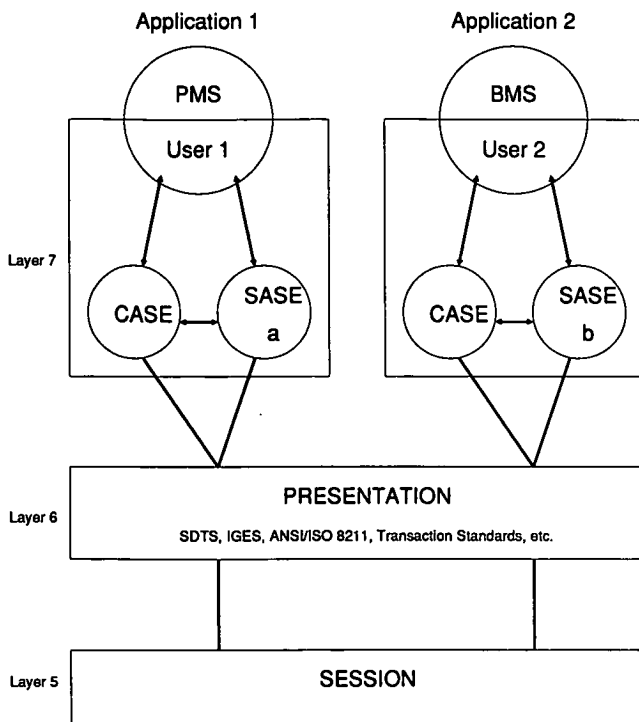


Figure 13. Upper layer architecture for IHIS

This kind of interface is very attractive because it clearly separates the BMS from the "infrastructure" of the networking and database management systems. It is not necessary for the BMS to be written in SQL or for its data to be physically stored in the SQL database manager's proprietary format, as long as the BMS can periodically convert its data to and from a format that the SQL system can access. All of the multi-user networking features, such as security and quality control, can be handled by the agency's choice of standardized systems, which are statewide and not BMS-specific. Because of the clear separation between systems, support services and problem solving can easily be delegated to the most appropriate support staff. On the BMS side, the choice of languages and systems can be made more on application-specific grounds; for instance, the C language might be chosen to give faster performance on PCs and engineering workstations.

Many engineering and technical decisions are most readily tackled by strategies of modernization or "divide and conquer." Client-server architecture is one way of dividing the large problem of data interchange standards into smaller, more manageable ones. It helps BMS to exploit the benefits of standardization while still accomplishing its own specialized analytical and performance objectives.

CONCLUSIONS

The proliferation of databases and management systems within an agency, and the desirability of easily transferring data among different levels of government suggest that there is a need for data interchange standards. Data interchange standards can greatly facilitate internal and external data communication and achieve significant savings to each agency that uses them.

But standards are a two-edged sword. If they are too stringent and short sighted, few organizational units and agencies will choose to comply. Barriers to adoption include money, time, and staff costs, contentment with the status quo, and the constant evolution of software and management tools that tend to render standards obsolete. If standards are too general and lax, they will not have enough specificity and force to ensure that data can move easily among units and agencies.

There can be local standards and universal standards. Local standards might apply within a single agency and universal standards might apply to all levels of government. Universal standards should accommodate local standards. Local data interchange standards can be avoided if an agency establishes suitable linkages among data files or a single integrated database. An example of a file linkage between bridge, roadway feature inventory, traffic, and accident data files would be common location reference points. In the absence of suitable linkages or an integrated database, there must be standards for a common (neutral) interface or for specific data transactions.

These same fundamental options apply for data exchange between different levels of government. Unless there is a cooperative agreement among government agencies to share a BMS and other infrastructure management systems, common databases or linked files among different governments will be infeasible. Again, the principal options will be standards for a common interface or for specific data transactions. An example of a standardized data transaction is the procedure states use to transfer NBI data to the federal government.

Data that moves between linked databases or exchanged through a standardized interface or transaction must be according to agreed on procedures and format. Although most data handled today by BMS and other decision support systems are simple alphanumeric fields, these fields may be stored in different formats ranging from integer to scientific notation and

include uncompressed to compressed data. The databases that store this information may be flat, relational, or hierarchical. Some standardization can be achieved by stipulating that all data be exchanged in ASCII format and that interfaces take the form of some common denominator. If units or agencies have the same type of database software, file structures, and formats, they can serve as the common interface. If not, a lower common denominator will be needed or a powerful method of encompassing and handling the disparate data will be required.

The character of data is becoming increasingly complex as a result of the emergence of geographic information systems, computer-aided design and drafting, optical imaging systems, and voice recognition. In addition, highway data collection methods now include a wide variety of measuring devices and sensors, such as infrared, laser, and ultrasound, whose raw data are usually recorded in analog form and then digitized for computer processing. In the future, data exchange will include the transfer of not only typical alpha-numeric fields but also graphic entities, imagery, text, and voice. For example, it is easy to imagine that a lower level of government might wish to transfer a modification of a cartographic (map) database for GIS to a higher level of government.

A universal standard for data exchange that accommodates a variety of different types of information ranging from traditional data fields to graphic entities, imagery, and voice must be quite flexible. Various efforts are currently occurring within the computer hardware, software, and telecommunications industries to develop standards that can support open system architecture and free exchange of data among diverse users handling many different kinds of data. This report described those efforts and briefly outlined a universal standard based on the well-known open system interconnect model.

REFERENCES

Ahamed, S. V., "Intelligent Networks." In *Encyclopedia of Communications*, Editor: Robert A. Meyers, Academic Press, Inc., San Diego, Calif. (1988).

American National Standards Institute, ANSI Catalog.

New York, NY (1990-1991).

Briggs, D.W. and B. V. Chatfield, "Integrated Highway Information Systems." *National Cooperative Highway Research Program Report 133*, Transportation Research Board, National Research Council, Washington, D.C. (October 1987).

Connecticut Department of Transportation, *Development of an Integrated Highway Information System*. Bureau of Planning (1989).

Federal Highway Administration, "Review of State Comprehensive Computerized Highway Information Systems." U.S. Dept. of Transportation, *FHWA-TS-87-220* (May 1987).

Landry, R., "PenPoint Puts It in Writing." In *PC World* (March 1991).

Lientz, B.P., "Computer Networks." In *Encyclopedia of Communications*, Editor: Robert A. Meyers, Academic Press, Inc., San Diego, Calif. (1988).

Loomis, M. E., "Objects and SQL: Accessing Relational Databases." *Object Magazine* (September/October 1991).

Martin, J., *Managing the Data-Base Environment*. Prentice-Hall Inc., Englewood Cliffs, N.J. (1983).

ORACLE. *SQL Language Reference Manual*. ORACLE RDBMS Version 6.0. (1988).

Paterson, W. O. and T. Scullion, *Information Systems for Road Management: Draft Guidelines on System Design and Data Issues*. Technical Paper, The World Bank (1990).

Stallings, W., *Handbook of Computer Communication Standards*. Volume 1, Second Edition, McMillan Publishing Co., Ind. (1990).

Tosta, N., *SDTS: Setting the Standard*. Geo Info Systems (July/August 1991).

Tsichritzis, D. C. and F. H. Lochovsky, *Data Base Management Systems*. Academic Press, Inc., New York, NY (1977).

ACKNOWLEDGMENTS

This research was jointly performed under NCHRP Project 20-7, Task 46 by The Urban Institute and Cambridge Systematics, Inc. William A. Hyman, Director of the Transportation Studies Program, The Urban Institute, was the Principal Investigator, and

Paul D. Thompson, Cambridge Systematics, Inc., was the Co-Principal Investigator. Roemer Alfelor of The Urban Institute and Michael J. Markow and Lance A. Neumann of Cambridge Systematics, Inc. provided technical input.

TRANSPORTATION RESEARCH BOARD

National Research Council
2101 Constitution Avenue, N.W.
Washington, D.C. 20418

NON-PROFIT ORG.
U.S. POSTAGE
PAID
WASHINGTON, D.C.
PERMIT NO. 8970

000015M006
ROBERT M SMITH
RESEARCH & ASST. MATLS SUPVR
IDAHO DEPT OF TRANSPORTATION
P O BOX 7129
BOISE ID 83707