Impacts of Emerging Information Technology on Data Collection and Availability

KENNETH J. DUEKER

New data collection and information systems technologies are fundamentally changing the type, quantity, and quality of data available for planning, managing, and operating transportation systems. The challenge is to integrate information technologies to create synergistic effects so that the individual parts are made more useful by contributing to the whole. The infrastructure life cycle management concept illustrates the importance of a shared spatial data base. The combination of planning, management, and operations requirements dictates that geographic information systems function in real time.

Key concerns in formulating a national transportation policy are

- Will today's transportation policies and infrastructure meet tomorrow's needs?
- How well are we meeting today's needs (*1*)?

Studies that have tried to answer these questions (*2*) "have encountered gaps in available information on the condition, performance, and use of the transportation system. As a consequence, . . . TRB is evaluating the current and anticipated state of transportation data, and . . . will recommend improvements to the information resources that are essential to support informed national decisionmaking in transportation. . . . The type of information that are [*sic*] needed . . . and our ability to obtain and use the requisite data are being affected by major technological and institutional forces (*2*)."

This paper focuses on the technological forces, which include the development of automatic vehicle location and identification (AVL/AVI) systems, weigh-in-motion (WIM), motorist information systems, vehicle navigational and route guidance systems, global positioning system (GPS), remote sensing, geographic information system (GIS), electronic exchange of shipping documents, and microcomputer-based data collection and analysis systems. Individually, a number of these technologies have proven useful. Together, how well do they contribute to gathering improved information on the condition, performance, and use of the transportation system to support informed national decision making in transportation?

New data collection and information systems technologies are fundamentally changing the type, quantity, and quality of data for planning, managing, and operating transportation systems. The challenge is to integrate data for maximum effectiveness. Data integration involves the establishment of linkages and standards among data sets. Foremost among linkage mechanisms is location, which is central to the management of spatially distributed infrastructures such as transportation systems. Consequently, GISs, which organize data around location, are essential in integrating data about transportation systems. The "horizontal" integration of data and systems across different data bases involves file transfers and raises a variety of issues concerning data exchange and standards.

The "vertical" integration of data and systems involves the aggregating and abstracting of data as one moves up the levels of an organizational hierarchy. Thus, data integration may mean aggregating locally useful data for strategic transportation policy planning at the national level. At the local level, it also means aggregating or selecting operations data for management and planning. This requires a real-time, transaction-based system for operations and administration, from which management and planning data can be derived as a byproduct, thereby avoiding expensive or duplicative data systems.

This paper describes data collection strategies, transportation data requirements, the role of GISs, and the importance of data integration. A conceptual model of the infrastructure life cycle provides a framework for fleet management and infrastructure management to illustrate the integration of transportation data. A truck port-of-entry example, a methodology for information integration, and technology determinants provide additional rationale for an integrated approach. Innovations in data collection technologies are examined, and their potential for use in integrated systems for policy planning is assessed. The paper concludes with recommendations for more effective use of new transportation data and information technologies.

DATA COLLECTION STRATEGIES

Data for national-level policy and planning can be collected in two ways. One is by means of special and separate national data programs (such as a census) or a mandate (such as the Highway Performance Monitoring System [HPMS]). The second approach is to rely on data being derived as a byproduct of decentralized administrative or operating systems of private firms or state and local governments.

The advantages of national-level special or mandated data programs include control, uniformity, and relevance to the policy being addressed. However, a uniform national-level data program may be costly and time consuming to install. A

Center for Urban Studies, Portland State University, P.O. Box 751, Portland, Oreg. 97207.

more significant problem may be quality, particularly if firms or state and local governments are not committed. For example, the quality of the data that went into the 1972 and 1974 national transportation studies is suspect. This is a particular problem when the data being collected are subjective rather than objective. Transportation "needs" are included in this subjective category.

A decentralized system where planning data are derived from, or are byproducts of, administrative and operating systems is more likely to produce reliable and accurate data, if appropriate standards are identified and applied by cooperating agencies. For example, uniform transit operating data is collected through UMTA's Section 15. Through aggregation and selection, UMTA is able to generate meaningful reports for federal policy use. From a local perspective, this is not entirely a separate data program but is, for the most part, a logical output of an administrative system. However, it is very difficult and time consuming to establish such a program. A major difficulty is to aggregate, sample, or generalize detailed data to discern underlying trends. It is difficult "to see the forest for the trees."

Although decentralized byproduct-type data programs are conceptually superior, they cannot be implemented in time for immediate policy analysis needs. Consequently, interim strategies must be found. Nevertheless, a long-term transportation data program is needed to reduce reliance on special ad hoc federal data programs, which usually yield suspect data. Routine, ongoing data programs that make better use of new technologies and tap into state and local administrative and operating systems may generate more reliable data for national policy use.

This analysis pursues the longer-term strategy of a decentralized system relying on data derived from administrative and operating systems. In either case, however, data integration is a problem. Any single data program or technology is insufficient, and the data must be placed in context for use in policy. This form of data integration often uses a denominator or conversion to a certain rate, such as per capita, per vehicle mile, or per passenger. It may use a time comparison—to last year or an average year. It may use a spatial comparison—to another state or county or to a median state or county. Finally, it may use subject comparison over a time period, such as comparing investment in transportation with investment in education over a 20-year period.

The challenge is to use the information technologies to integrate data and thereby make the data more useful than the individual data elements themselves. This integration needs to occur at appropriate levels; otherwise, misleading averages or rates may be generated. For example, transit cost and patronage need to be integrated at the agency level to be of most use. Then, comparisons with state and national data regarding costs per passenger are meaningful.

A national-level transportation data program is essential for policy analysis, formulation, and implementation. It will have two components: one orchestrates data derived from state and local administrative and operating systems, while the other collects data directly. New information technologies have application in both instances.

National-level data programs generally rely on sampling or mandated reporting. Information technologies can aid these kind of programs. Generally, computer-aided methods increase productivity. For example, the 1990 National Personal Transportation Study will be conducted by telephone using random digit dialing, with scanner input of forms to the computer. More powerful and versatile statistical packages will aid in analysis and reporting. Similarly, organizations such as Oak Ridge National Laboratory are employing computer-based systems to analyze large, national socio-demographic data sets in conjunction with transportation networks to address policy and planning issues at the national level with greater ease than previously possible. This kind of capability is essential to examine such issues as the contribution of transportation to the greenhouse effect or global warming. In these instances, aggregation of state and local data may not be particularly suitable.

It is difficult to project whether use of new information technologies, such as tagging and tracing a sample of persons, shipments, or vehicles, will become widespread. The technologies may be available, but the design of a national data program of this type may not be feasible or socially acceptable. A more likely data collection or integration program would yield aggregate data, such as traffic flows, in a more timely and spatially representative manner.

TRANSPORTATION DATA REQUIREMENTS

In addressing transportation data requirements, Schmitt (*2*) classifies transportation data as follows:

- *Facility Inventory, Condition, and Performance.* Data on the extent, ownership, physical condition, operating costs, speed, capacity, and other characteristics of rights-of-way, terminal and network facilities, and related transportation infrastructure.
- *Equipment Inventory, Condition, and Use.* Data on the number, miles of travel, ownership, physical condition, operating costs, speed, capacity, and other characteristics of vehicles, rolling stock, aircraft, and vessels that operate on transportation facilities.
- *Carrier Performance and Condition.* Data on the expenses, revenues, ownership, market coverage, labor force, and service characteristics of public and private for-hire carriers, shipper-owned transportation services, transportation services provided by social service and other organizations for their own account, and arrangers of transportation service.
- *Passenger and Freight Flows.* Data on the volume, geography, value, and other characteristics of passenger and freight flows.
- *Demographics and General Economic Activity.* Data on the number, geographic distribution, economic health, output or propensity to travel, vehicle availability, and other characteristics of households, businesses, and users of the transportation system.
- *Safety and Security.* Data on accidents, near-misses, personal injuries, emergency medical services, cargo damage, passenger and cargo restraints, hours of operation, drug and alcohol use, and terrorist incidents and countermeasures.
- *Finance and Program Administration.* Data on public agency cash flow, personnel, tax burden, bonding authority and other revenue sources, trust fund balances, the distribution

of obligations by contractor characteristics and geography, and other characteristics of public finance and administration.

This taxonomy of transportation data types is a useful way to describe the condition, use, and performance of transportation systems. However, all data are not equally important. The AASHTO Standing Committee on Planning (SCOP) provides focus to identify data requirements for national transportation strategic planning (*3*).

The SCOP posed three questions:

1. Are the data adequate in the particular modal area?
2. What information should be collected on a continuous basis?
3. What types of questions should be answered from a strategic planning process?

Responses indicated that the following types of data should be available for all modes:

- Facility inventories,
- Usage data,
- Financial data,
- Quality of service data, and
- Population and economic data.

In addition, the SCOP recommended that policy models be available to test the consequences of

- Various funding scenarios,
- Major changes in policy direction,
- Major changes in any of the above data categories, and
- Impact of external policies (air quality, energy, etc.).

Information regarding the condition, use, and performance of transportation systems is an important input to models that test changes in trends or policies. Consequently, this paper is attentive to data and information technologies that monitor transportation systems. Data on transportation system performance, in terms of mobility and congestion measures, and models of the impact of investment policies on performance are in great demand. Yet, data to develop models are in short supply.

New technologies to support transportation data and information programs are identified in Table 1. To identify technologies for data collection, data integration, and data use in trend and policy models, the transportation system is characterized by the following elements: driver/user, vehicle, fleet, traffic, roadway, network, and the demand side (users and nonusers).

The technologies identified in the table constitute new ways of measuring and understanding human factors, vehicle location and identification, fleet management, data about roadway traffic and the network within which it operates, the flow of persons and goods in networks, and of the flow of the users and nonusers of transportation that make up the socioeconomic group served by the transportation system. Individually, these technologies are being implemented, but most of the effort is toward making the technologies work correctly. Once the data are flowing, attention will shift to integrating and using the data.

TABLE 1 TECHNOLOGIES TO SUPPORT TRANSPORTATION DATA AND INFORMATION PROGRAMS

Transportation System Elements	Data Collection Technologies
Driver/user	•Human Factors
Vehicle	•Driver Simulators Crash Tests Accident Data Analysis
Fleet	•AVL Scheduling/Dispatching
Traffic	•ATR/AVC, WIM/AVI Real-Time Traffic Control
Roadway	•HPMS,LTPP, WIM Infrastructure Management
Network	•Hazmat Routing Transportation Models
Demand Side	•Socio-economic Data

The important point is that the data and information technologies must function together. Although the technologies are useful individually, data integration is required to assess changes in trends and policies. The challenge is to integrate the information technologies to create synergistic effects so that the individual parts are made more useful by contributing to the whole (*4*).

In this way, technologies can function together. For example, AVI and WIM can help enforce truck weight restrictions and provide planning data for roadway design. Similarly, ATR (automatic traffic recorder) technology can serve several functions. Traffic data can be aggregated upward to the system level and provide valuable VMT (vehicle miles traveled) data for national planning. These data can also be used to infer seasonal, growth, and truck factors to other locations on the highway system to serve the needs of highway design at the project level and traffic management at construction sites. A logical next step is to use data collected at a denser set of ATRs, in real time, for motorist information systems and traffic management, with planning data as a byproduct.

OTHER INNOVATIONS IN DATA COLLECTION

In addition to the hardware technologies of data collection, such as WIM, AVL, and AVI, there are other innovations in data collections. Examples of these are

- Administration records exchange/access. This is another example of using operational data for management and planning. Again, selective aggregation is the key ingredient.
- Combining survey data with aggregate data for origin-destination (O-D) estimation. Ben-Akiva and Morikawa have developed a data combination and updating method that corrects survey data for nonresponse biases and reduces sampling errors by statistically combining survey data with aggregate data (*5*).

These examples are mentioned to complete the picture; innovation in data collection is more than hardware improvements. A number of statistical issues must be addressed. For example, statistical techniques to infer appropriate seasonal factors for short-count traffic data locations from permanent traffic counting locations require the integration of data by location. This can best be accomplished by use of a GIS.

GEOGRAPHIC INFORMATION SYSTEMS FOR DATA INTEGRATION

Because transportation infrastructure is locationally distributed, geographic location is the key to integrating data about infrastructure. The integration of data for infrastructure systems is accomplished by the application of GIS concepts and technology. A consensus definition of GIS was developed for the American Society of Photogrammetry and Remote Sensing and the American Congress on Surveying and Mapping (*6*):

> A system of hardware, software, data people, organizations, and institutional arrangements for collecting, storing, analyzing, and disseminating information about areas of the earth.

A GIS is more than hardware and software to produce maps. It includes a spatial database consisting of structured data that

1. Enables linking of spatial and attribute data for geographic features,
2. Relates data across map layers, and
3. Supports routing, adjacency, and inclusion applications.

The linking of spatial and attribute data (item 1 above) enables spatial access to a data base containing attributes of the selected objects. It also enables map display or highlighting of spatial objects having specific characteristics or attributes. In the first instance, the spatial search identifies objects for which the attribute data are reported. In the latter case, the data base is searched first, and the objects having the requisite attributes are displayed.

The geometric relation of data layers (item 2 above) is often called "polygon overlay" and is analogous to the physical overlaying of mylar sheets. Overlaying reveals the number of point, line, or area objects that are contained in another set of areas (for example, the area of roadway clearing through forested land).

Applications such as routing in transportation networks rely on topology relations consisting of explicit knowledge of the mutual connectivity and relative spatial positions of the points, lines, and areas that make up map space. This maintenance of topology relations is central to a GIS and is sometimes referred to as "map intelligence." The relationships among mapped objects are explicitly maintained.

In addition to the three data structure issues, there are two other important issues related to spatial data bases. The layers of data in a spatial data base must be in spatial registration, either by means of registration to the same coordinate base or by means of common spatial referencing systems, such as route/milepoint or node/link identifiers. Also, the data types or layers in the spatial data base may be institutionally independent or may be managed or controlled by the unit or organization responsible for maintenance. By means of spatial registration, the data layers can be shared as needed, but control and maintenance can be decentralized. This serves to avoid organizational conflict.

The data model used in spatial data bases can be concisely expressed as follows (*7*):

> A data model is the whole of concepts and expression tools that enables the description of a complex set of data items. The main concepts in the . . . data model are features, attributes and relationships. A feature is a formalized entity that is used to represent a topographical object. The properties and particularities of the objects are represented by means of attributes. Properties involving more than one feature are described by means of relationships. A group of features that is strongly related is called a layer.

GISs are important in addressing data integration needs of transportation applications. Examples of partial GISs include network modeling data bases, which are made up of load nodes and links and nodes describing the transportation system, and roadway inventory and mapping systems. These are viewed as partial GISs because they do not fully integrate the roadway inventory information, the roadway cartography, and the roadway network in a flexible systems environment to input, store and retrieve, model, and display the results.

The use of GIS technology for urban applications provides an example of an integrated data base. In the 1970s, the U.S. Bureau of the Census developed a system called GBF/DIME (Geographic Base Files/Dual Independent Map Encoding) for U.S. metropolitan areas. The extent of coverage has been expanded to include the continental United States, and the system is now called TIGER (Topologically Integrated Geographic Encoding and Referencing System) (*8*).

The concepts underlying TIGER are important in understanding the power of data integration. TIGER maintains the relationship of points, lines, and areas that make up the street/road and jurisdictional/statistical systems of the nation. In simple terms, this enables the assignment of people, located by street address, to their correct census geographical units for tabulation. In broader terms, it demonstrates how to integrate data, such as relating socioeconomic data for users of transportation systems to the facilities that supply transportation service. It also demonstrates the power of topologically relating the individual street facilities to enable analysis of flows.

Of course, these concepts are not new, and the process of relating supply and demand for transportation has been incorporated into transportation planning models. However, the challenge is to place the network assignment model data bases and the TIGER files into roadway inventory and real-time operational environments. This requires hardware and software systems having GIS functionality. Unfortunately, current GISs do not support many of the needs of transportation systems, particularly the incorporation of real-time transactional data flows that may come from ATR or AVL systems.

TenEyck et al. (*9*) provide an example of using a GIS to integrate data. Data from the accident record system are integrated with data from the roadway management system to show that "run-off-road" accidents occur on stretches of highway having shoulder drop-offs of more than 2 in. This integration of data at the time of analysis is made possible by relating data via common keys, in this case control sections or, more generally, route and milepoint. The Highway Safety Strategic Transportation Research Study (STRS) program builds on this type of research by targeting the three principal elements affecting highway safety: the vehicle, the roadway environment, and the driver. This allows an examination of "cross-cutting issues, such as the impact of new technology on highway and vehicle safety, the development and use of better data on injuries sustained in accidents, and the role and control of congestion in highway safety" (*10*, p. 7). This is in contrast with the Long-Term Pavement Performance (LTPP) program of the Strategic Highway Research Program (SHRP). The LTPP program involves monitoring more than a thousand 500-ft sections of in-service highways (*11*, p. 5) and requires data integration by advance planning. It will streamline those analyses that are anticipated but will make difficult those that are not.

FRAMEWORKS FOR TRANSPORTATION DATA INTEGRATION

Integrating transportation data, particularly those coming from new technologies such as ATR, WIM, and AVL, poses difficult problems in systems design. Five system design and data integration approaches to these problems are examined below. One is a conceptual framework called infrastructure life cycle management as described by McDonnell-Douglas (*12*). The second is an example of fleet management in transit organizations (*13*). The third example deals with truck port-of-entry automation (*14*). The fourth example involves information integration methodology (*4*), and the fifth deals with technological determinants (*15*).

Infrastructure Life Cycle Management

Life cycle management is described as a continuous process of planning, designing, constructing, and operating, as illustrated in Figure 1. When applying this concept to infrastructure systems, the life cycle takes on the components of strategic planning, engineering design and construction, and facilities management (see Figure 2). Supporting infrastructure management with new information technologies results in modification of the life cycle concept to include GISs, computer-aided drafting (CAD), computer-aided engineering (CAE), and facilities management (FM). This is also illustrated in Figure 2. The key feature of an integrated system to support operations, management, and planning is a common spatial data base. As a result, changes made by one group

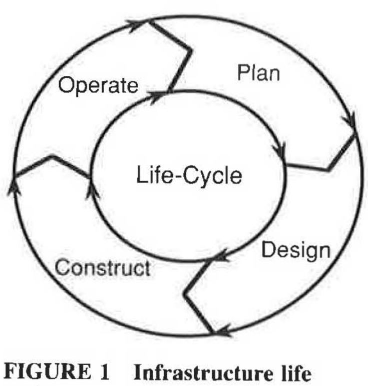

FIGURE 1 Infrastructure life cycle.

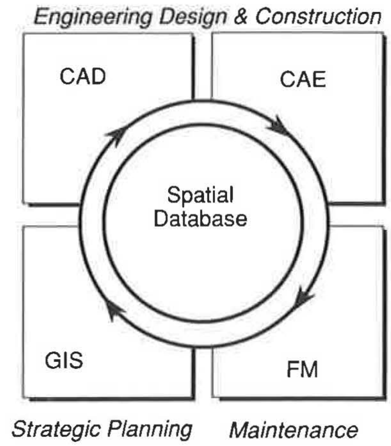

FIGURE 2 Infrastructure management system.

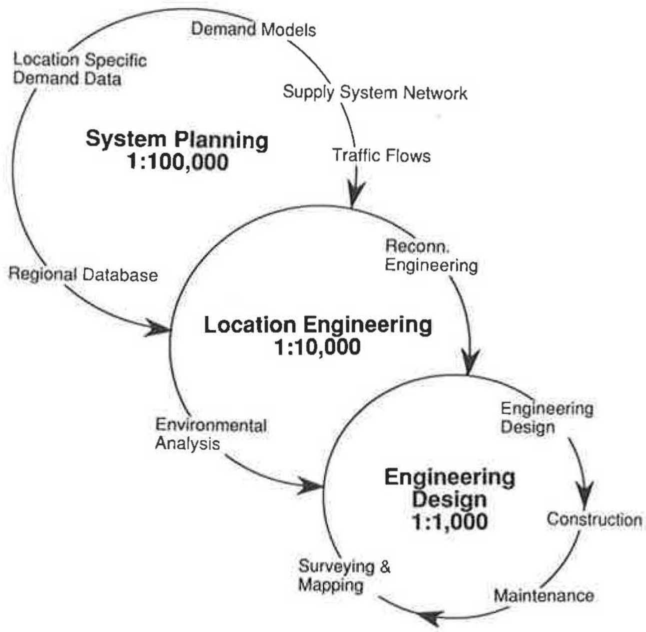

FIGURE 3 Infrastructure life-cycle management: map scale difference.

are available to all. However, the nature of this data base is not easily defined. For example, is it a single data base? Or is it several data bases with file transfers from a GIS structure to a CAD data structure? A number of data standards and file transfer issues must be addressed.

Applying this infrastructure life cycle management concept to a highway information system involves both potentials and problems. The chief problem is one of scale differences (see Figure 3). Operations, management, and planning often require different time and spatial scales, which makes it easier to develop and support separate systems rather than integrated ones (*16*).

The scale at which to support engineering design and construction involves CAD and CAE systems that deal with highway plan and profile information. The highway maintenance function would use and keep current the as-built plans. The highway planning function related to this detail would include accident and traffic analysis and site planning of adjacent land uses that might require curb cuts and traffic control devices. Again, the key feature of this integrated system to support engineering design and construction, facilities management, and detailed planning is a common spatial data base. Changes made by one group are available to all, and everyone can count on up-to-date data.

However, most planning is done not at an engineering scale but at a more generalized corridor or system scale. At a smaller spatial scale, there is another highway information life cycle. Maintaining information about traffic flows, signs and signals, and accidents does not warrant the spatial detail needs of highway design and construction. Consequently, a cycle whose spatial data base contains more generalized data is needed. This more generalized spatial data base would be statewide rather than project specific and would represent highway elements as single line widths rather than rights-of-way containing facilities. This spatial data base would have the following characteristics:

- A highway system consisting of links and nodes for use in minimum path routing and traffic assignment models,
- An accurate cartographic representation of highway system links for mapping purposes, and
- Use of the route and milepoint spatial referencing system by recording the range of milepoints for each highway link (*17*).

As shown in Figure 4, this method of organizing the spatial data base for the highway system allows data integration. Accident, traffic, and roadway inventory data that are normally recorded by route and milepoint can be related, displayed, and incorporated into models. This is analogous to the TIGER data structure with the route and milepoint replacing or augmenting the address ranges.

Fleet Management

Application of the life cycle management concept to transit fleet management takes on a different form. The temporal dimension becomes more critical, and the detailed spatial scale associated with engineering design and construction is of less concern, unless the transit property is involved with fixed-rail systems. Unfortunately, the use of temporal data in a GIS is not well developed. A recent article by Langran (*18*) provides a review of temporal data base research and its use in GIS applications.

The importance of the temporal scale is illustrated in Figure 5 with different cycles of operations, management, and planning. The real-time operations cycle supports the needs of dispatching and monitoring schedule adherence. The management cycle deals with seasonal or periodic scheduling at the times drivers sign up for routes and shifts, and the planning cycle occurs on a more seasonal or annual time frequency when routing changes are made.

Fleet management is concerned with the flow of data from buses to dispatcher, then archiving and integrating the data for management and planning uses. Transit organizations are adopting new technologies for communications, automatic passenger counting, automatic fare counting, automatic vehicle location, etc. It is the integration of these technologies that is necessary to generate data for achieving the objectives of life cycle management.

Modern radio and computer technology enables a polling of bus fleets (with digital communication) to the dispatcher's computer for identification of the bus, driver, route/run, and operating data, including the odometer reading, door opening, passengers on and off, and location. Immediately, this information is used for exception reports on schedule adherence and for schedule adherence feedback to drivers. For management and planning, selected polling records are aggregated for analysis. Data can be summarized by route, time, driver, bus, and timepoint for analysis.

Currently, the focus of attention in AVL for transit fleets is on implementation of the technology. Early adopters are now focusing on extracting from the flow of data used for on-street service operation to create a historical data base for planning and management using the GIS concept (*19*).

Truck Port-of-Entry Automation

Information technologies, in the form of AVI and WIM systems, are being integrated with existing truck scales and a

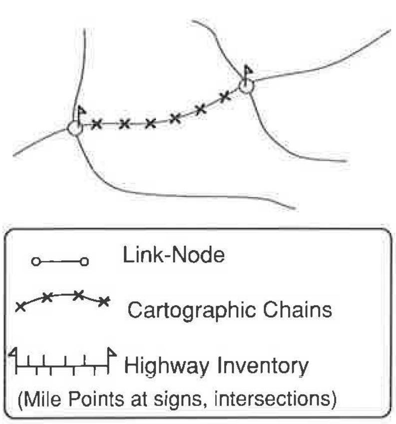

FIGURE 4 Highway system integrated data base.

FIGURE 5 Fleet management vehicular transportation systems.

supervising system computer (SSC) at the Woodburn Southbound Port-of-Entry in Oregon (*14*):

> The purpose of this integration was to allow trucks with transponders, meeting legal weight as shown by the WIM, and Public Utility Commission (PUC) criteria from the PUC vehicle database, to bypass the static scales. Data on these trucks is stored by the SSC. Trucks, without transponders, meeting the legal weight as shown by WIM, are directed to the static scales (where their PUC identification number is keyed into the computer). The SSC brings the PUC information onto the screen. If the trucks meet the PUC criteria with respect to taxes and safety, only its weight from the static scales is recorded. It then is allowed to pass through. Those trucks that do not meet weight or PUC criteria are stopped and issued citations and/or go to the PUC location to obtain permits.
>
> The data is downloaded to the mainframe and changes in the PUC database are uploaded on a regular basis. Daily statistical records and tables are produced by the SSC.
>
> The automation system has allowed the weighmasters to weigh vehicles more quickly, reducing congestion. The system has also revealed outstanding weight-distance tax payments owed by some firms. A chronic offender list is being developed through the use of the SSC. Such offenders typically are found to have a history of overweight or permit violations.
>
> If all trucks had transponders, the automated system would dramatically reduce the number of vehicles weighed at . . . (perhaps by 50 percent), thus greatly reducing the weighmaster workload. This would allow for rescheduling personnel to other duties.

Information Integration Methodology

Achieving the integration of data and systems is a complex matter. Nyerges (*4*) describes a methodology for information integration:

> Information integration is defined in this context as the bringing together of information parts into a working whole, controlling redundancy where appropriate. A synergistic effect is anticipated such that the individual parts are made more useful by contributing to the whole. This bringing together does not mean that the whole is one "physical" whole, however it does mean that the parts cooperate. The term "information integration" is derived from the concatenation of "information sharing" and "systems integration." In this regard, information integration could involve one or more of the four components of a geographic information system: data, software, hardware, and/or personnel. Consequently, information integration could be accomplished through various strategies involving the integration of data, the integration of software/hardware functionality and/or the integration of personnel.
>
> Data integration requires at least two steps: integration of data descriptions for a database, and integration of the data itself. Two types of data descriptions are important for data integration. These are descriptions contained in a schema, and descriptions (definitions) contained in a data dictionary. Structural descriptions in the schema describe how data are represented and stored in a database.
>
> Functional integration involves bringing together separate software/hardware components to enhance or extend the analyses in a system; or reducing/eliminating database management system duplication. This process involves the integration of dissimilar software and/or hardware. Software integration usually involves the merging of data-structure constructs. Hardware integration involves the physical linkage of computers or the linkage of peripherals with computers to enhance data manipulation, e.g., graphics plotters or special high-speed processors.
>
> Personnel integration involves the merging or reorganization of staff. . . .
>
> An information integration methodology . . . consists of four stages. The stages are: (1) integration strategy planning, (2) integration analysis, (3) integration design, and (4) integration implementation. To better understand the nature of the activity at each stage in the process, each stage can be described at three levels of abstraction. The levels are: (a) conceptual understanding, (b) techniques that can be used for expressing the concepts in terms of data constructs, and (c) software/hardware tools available for implementation of data constructs.

Complex problems, such as building a vehicle navigation system, require integration of data and subsystems for address conversion, route planning, route guidance, and position display. New tools are needed to build these complex systems. Information integration tools are important in achieving the needed integration of information technologies, particularly as the trend in computing technology evolves to workstations in a server-net environment. This technological trend makes data integration increasingly feasible.

Technological Determinants

Opportunities for data integration and GIS applications are shaped by technological developments. In a paper by Travis et al. (*15*), three trends are identified that will influence developments:

1. The server-net model of computing environments,
2. Rapid technological change, and
3. The emergence of open systems.

Server-Net Model

The server-net model conceives of computing as specialized services emanating from different nodes of a centerless network. A GIS design involves three kinds of server nodes: GIS servers, data base servers, and user workstations.

Rapid Technological Change

Technological change is occurring more rapidly in computing than ever before. Thus, it is difficult to avoid design constraints and limitations that soon become unnecessary because of technological progress. This difficulty is a major challenge to the GIS concept and data integration designers.

Emergence of Open Systems

Systems constructed according to vendor-independent standards are referred to as open systems and have the distinct advantage of freeing users from dependence on particular vendors and their proprietary standards. The server-net model in an open system computing environment is incrementally modifiable with new server nodes added and old server nodes upgraded or replaced without major impact on the other nodes in the net. This is possible only if the nodes connect to and interact with each other according to standardized protocols.

These rapid technological developments have the potential for further relaxation of economic and institutional constraints to allow adoption of new methods and approaches (*20*). The server-net model and the single data base ought to both reduce cost and reduce the struggle over the control of computing and information. These are democratic forces in contrast to existing centralized mainframe computing (*21*). However, in a decentralized computing environment, a considerable investment of time and energy is needed to work out responsibilities for data layers and standards for data exchange.

The application of technology to transportation data requires frameworks for integrating various data. Without integration, data are of little use. Context is needed. The approaches described above illustrate ways in which integration can be achieved and data made more useful.

CONCLUSIONS AND RECOMMENDATIONS

This analysis shows that effective use of new data and information technologies for the operations, management, and planning of transportation systems requires integration of information resources. Both horizontal integration of data and systems across organizational units and vertical integration across levels are needed. Horizontal integration is facilitated by GISs and data exchange standards, while vertical integration is keyed by systems for data aggregation and abstracting.

GIS technology is a major tool in data integration. However, the state of the art of GIS does not handle well the real-time data base requirements of transportation system applications. Improvements are needed to handle the needs of planning, managing, and operating transportation systems.

A conceptual framework of infrastructure life cycle management is used to show why and how data integration, by means of a shared data base, becomes the new issue. Although data collection technologies are essential, the challenge is to integrate information technologies to create synergistic effects so that the individual parts are made more useful by contributing to the whole.

REFERENCES

1. *Moving America–New Directions, New Opportunities. Volume 1: Building the National Transportation Policy*. U.S. Department of Transportation, 1989.
2. R. Schmitt. *Information for National Transportation Decision Making: Needs, Resources, and Issues*. National Transportation Policy Team, U.S. Department of Transportation, 1989.
3. H. Peyrebrune. *Report on Data Requirements for National Transportation Strategic Planning—AASHTO's 20/20 Experience*. New York Department of Transportation, Albany, 1989.
4. T. Nyerges. Information Integration for Multipurpose Land Information Systems. *Journal of the Urban and Regional Information Systems Association*, Vol. 1, No. 1, 1989, pp. 27–38.
5. M. Ben-Akiva and T. Morikawa. Data Combination and Updating Methods for Travel Surveys. In *Transportation Research Record 1203*, TRB, National Research Council, Washington, D.C., 1988, pp. 40–47.
6. K. Dueker and D. Kjerne. *Multipurpose Cadastre: Terms and Definitions*. American Society for Photogrammetry and Remote Sensing; American Congress on Survey and Mapping, 1988.
7. H. Claussen, H. Lichtner, L. Heres, P. Lahaije, and J. Siebold. GDF: A Proposed Standard for Digital Road Maps To Be Used in Car Navigation Systems, Vehicle Navigation and Information Systems. *Proc., Vehicle Navigation and Information Systems Conference*, Piscataway, N.J., Institute of Electrical and Electronics Engineers, 1989, pp. 325–326.
8. R. Marx. The TIGER System: Automating the Geographic Structure of the U.S. Census. *Government Publications Review*, Vol. 13, 1986, pp. 181–201.
9. T. TenEyck et al. *The Development of a GIS in Pennsylvania*. Pennsylvania Department of Transportation, Harrisburg, Dec. 1989.
10. Highway Safety STRS: Reassessing the Direction of Current Research. *TR News*, July–Aug. 1989.
11. D. J. Kulash. SHRP Research Products Expected by 1993. *TR News*, July–Aug. 1989.
12. *Lifecycle Solutions for Infrastructure Management*. McDonnell-Douglas Information Systems Company, St. Louis, Mo., undated.
13. K. Dueker. *GIS Applications in Transit Organizations*. Center for Urban Studies, Portland State University, Portland, Oreg. (in preparation).
14. M. Krukar and K. Evert. The Metamorphosis of Weight Enforcement Strategies in Oregon. Presented at 3rd National Weight-in-Motion Conference, St. Paul, Minn., Oregon Department of Transportation, 1988.
15. L. Travis, R. Smith, and W. Berg. *Research on Adaption of GIS for Transportation: The Server-Net Model*. Department of Civil Engineering, University of Wisconsin-Madison, 1990.
16. K. Dueker. Urban Applications of Geographic Information Systems Technology: A Grouping into Three Levels of Resolution. *Proc., Urban and Regional Information Systems Association*, Vol. 2, 1988, pp. 104–109.
17. T. Nyerges and K. Dueker. Geographic Information Systems in Transportation. *Regional Computer-Assisted Cartography Conference: Summary*. Pennsylvania Department of Transportation, Harrisburg, 1988.
18. G. Langran. A Review of Temporal Database Research and Its Use in GIS Applications. *International Journal of Geographic Information Systems*, Vol. 3, No. 3, 1989, pp. 215–232.
19. J. Hoffman and H. Gault. AVLC Data for Planning and Management: Information Overload. *Proc., International Conference on Automatic Vehicle Location in Urban Transit Systems*, ed. B. Hemily. Canadian Urban Transit Association, 1988, pp. 201–220.
20. K. Dueker. Multipurpose Land Information Systems: Technical, Economic, and Institutional Issues. *Photogrammetric Engineering and Remote Sensing*, Vol. 53, No. 10, Oct. 1987, pp. 1361–1365.
21. H. Jacobs. Implementing Local Multipurpose Land Information Systems: Political-Economic Research Issues. *Computers, Environment and Urban Systems*, Vol. 13, 1989, pp. 3–13.