U.S. Department of Transportation

Federal Highway Administration

PB98-160583

ITS AS A DATA RESOURCE

Preliminary Requirements for a User Service

Federal Highway Administration
Office of Highway Information Management

ITS AS A DATA RESOURCE

Preliminary Requirements for a User Service

Prepared by
Richard Margiotta
Science Applications International Corporation

for

Federal Highway Administration
Office of Highway Information Management

April, 1998

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF TABLES AND FIGURES

Tables

Figures

EXECUTIVE SUMMARY

Introduction

Much of the data generated by Intelligent Transportation Systems (ITS) can be of great value beyond their immediate use in real-time control strategies. However, unless ITS operators have made special provisions, data from system surveillance equipment are typically not stored for future use. Because the amount of data is so enormous, it is doubtful that simply saving the raw data would be of use to other stakeholders; some level of aggregation or sampling is required to make the data more meaningful to stakeholders. Further, the National ITS Architecture currently has no specification for a data archival process.

Data needs of many stakeholder groups have been identified in several past studies. In particular, the *ITS As A Data Resource Workshop* held in January, 1998 substantiated stakeholder needs and began the process of matching ITS-generated data with those needs.

Because of the wide range of support among stakeholders represented at the *ITS As a Data Resource Workshop*, it has been determined that there is a need for a new User Service to be included in the National ITS Architecture: the Archived Data User Service. The recommendations herein are the starting point for this process. Specifically, this document has four objectives:

(1) definition of the Archived Data User Service;
(2) development of preliminary requirements for revising the National ITS Architecture;
(3) identification of steps for implementation beyond inclusion in the Architecture; and
(4) fostering communication between stakeholder groups.

Stakeholder Groups

The number of stakeholder groups with an interest in the Archived Data User Service is much larger than for any other of the 30 currently specified User Services. Stakeholders include:

- MPO and state transportation planners;
- ITS operators and transportation engineers;
- Transit operators;
- Air quality analysts;
- MPO/state freight and intermodal planners;
- Safety planners and administrators;
- Maintenance personnel;
- Commercial vehicle enforcement personnel;
- Emergency management services (local police, fire, and emergency medical);
- Transportation researchers; and
- Private sector users.

It is noteworthy that a significant beneficiary of archived data is the ITS operator community, whose systems collect the data in the first place.

Uses and Benefits of Archived Data Generated By ITS

For the most part, data generated by ITS is similar to data collected by traditional means (e.g., traffic counts) but ITS-related data are collected continuously and at a very detailed level. Accordingly, a wide range of stakeholder functions can be supported with data from ITS. For example, roadway surveillance data can be used in many stakeholder applications, including development and calibration of travel demand forecasting and simulation models; congestion monitoring; transit route and schedule planning; intermodal facilities planning; and air quality modeling. In addition to identifying specific applications, several general observations on the uses and benefits of ITS-generated data can be made.

- The continuous nature of most data generated by ITS removes sampling bias from estimates and allows the study of variability.

- The detailed data needed to meet emerging requirements and for input to new modeling procedures can be provided by ITS.

- Use of data generated by ITS for multiple purposes is a way to stimulate the support of other stakeholders for ITS initiatives.

- Promoting the use of archived data for multiple purposes complements the initiative for integrating ITS in general.

- Because the data are already being collected for ITS control, other uses provide a value-added component to ITS.

- ITS is a rich data source for multiple uses, but not a panacea; traditional sources of data will continue to be important.

- As the focus of transportation policy shifts away from large-scale, long-range capital improvements and toward better management of existing facilities, ITS-generated data can support the creation and use of the system performance measures that are required to meet this new paradigm.

Basic Requirements of the Archived Data User Service

Based on the data needs of stakeholders and the nature of data in ITS, three types of recommendations are made. The recommendations are not exhaustive, merely a point of departure for future discussions with stakeholders.

First, general requirements for systems to support the User Service are specified. These relate to:

- support for both centralized and decentralized systems;
- data processing, storage, and retrieval;
- data aggregation and sampling;
- metadata specifications;
- data ownership;
- coordination with all relevant standards and data dictionary efforts; and

- data privacy.

The second type of recommendation is the specification of the types of ITS-generated data that should be considered, the data's attributes, internal storage structure, data reduction cycles, and level of accuracy. A detailed table of data elements is given in the report which includes, for example, traffic volume, vehicle speed, classification, as well as transit vehicle boardings, incident data, and environmental data. As with the general requirements, these data elements are the starting point for stakeholder discussions. They also can form the basis of a data dictionary and for the development of formal Architecture requirements, if required.

The third type of recommendation is aimed at the next steps in the process. These are offered for consideration in promoting development and use of the Archived Data User Service in the field.

- "Best Practice" procedures for performing quality control and editing on Archived Data User Service data should be developed.

- "Best Practice" procedures for analysis of archived data generated by ITS would provide guidance to stakeholders and would promote the use of Archived Data User Service.

- Coordination with ongoing data dictionary efforts is crucial to the future development of the Archived Data User Service.

- The needs of stakeholder groups should be represented in the development of standards that may have an impact on the relevant data.

- The Archived Data User Service should be integrated into other Federal, state, and local data collection programs.

- Training the various stakeholders in each others' needs is seen as an ongoing requirement.

- A concentrated field effort, similar to the ITS Field Operational Tests, to demonstrate the implementation and use of the Archived Data User Service should be undertaken.

- Funding for the Archived Data User Service should be implicit in future ITS grant programs and national legislation.

Successful implementation will require resolution of many difficult institutional and technical issues. A total of sixteen such issues are identified, including: development, operation, and maintenance costs; system access; ownership; data quality; data management; data and communications standards; privacy concerns; data analysis; coordination with other data collection efforts; liability; confidentiality of privately collected data; incremental and uncoordinated ITS deployments; retrofitting vs. new development of systems; data flows not defined by the National ITS Architecture; conformance with metric conversion standards; and training and outreach.

1. INTRODUCTION

1.1 Purpose of This Document

A key feature of Intelligent Transportation Systems (ITS) is the use of information about transportation system conditions and travelers to improve overall system performance. ITS can generate massive amounts of data that are used primarily in real-time to effect control strategies. Examples include the adjustment of ramp meter timing based on freeway flow conditions and the use of variable message signs (VMSs) to communicate traffic incidents to travelers.

Increasing deployment of ITS throughout the Nation has brought an awareness that ITS-generated data offer great promise for uses beyond the execution of ITS control strategies. Potential applications include transportation planning, administration, operations, and research. In most cases, ITS-generated data are similar to data traditionally collected for these applications, but are much more voluminous in quantity and temporal coverage. However, up to now there has been no organized effort to identify data generated by ITS for other uses and to specify methods for their storage and analysis. This document will serve as the starting point for the process of getting a new user service -- the **Archived Data User Service (ADUS)** -- into transportation practice. Specifically, this document has four main objectives in order to bring the issue into sharper focus:

1. <u>Define the Archived Data User Service.</u> The potential of data generated by ITS for multiple uses is so great that it warrants the definition of its own ITS User Service. The creation of a new User Service requires that the National ITS Architecture (see Section 1.2.4 for a description) be amended to include it. The first step in this process is the definition of the User Service, including its stakeholders, main functions, costs and benefits, and technical and institutional issues. This document provides the information necessary to specify the Archived Data User Service.

2. <u>Serve as a Preliminary Requirements Document for the National ITS Architecture.</u> In the National ITS Architecture revision process, once a User Service is defined a detailed set of requirements is generated. These requirements are then used to formally amend the National ITS Architecture. Development of the requirements is based on detailed discussions between stakeholders and the Architecture developers. Because an ample history already exists on the needs of stakeholders, there is enough information to develop *a preliminary list of general requirements* for the Archived Data User Service. These requirements must be refined through more intense discussions with stakeholders.

3. <u>Identify Steps For Implementation Beyond Inclusion in the National ITS Architecture.</u> As presented in Chapter 3, many technical and institutional issues are associated with implementing the Archived Data User Service. In many ways, these are larger barriers to implementation than adding the User Service to the National ITS Architecture. Therefore, this document will identify these issues and will offer preliminary recommendations for dealing with them (Chapter 4). Clearly, much more remains to be done in developing effective procedures for implementing the Archived Data User Service in the field, but this document is the first step in that direction.

4. Foster Communication Between Stakeholder Groups and the ITS Community. A need exists to bridge the communications gap between personnel who design, construct, and operate ITS and other transportation personnel. Therefore, this document will serve as an educational tool to foster communication between these diverse groups. The report is of interest to all potential stakeholders in that it identifies ITS-generated data that have application for multiple uses. It also provides a background on many of the technical chores (and their associated data needs) undertaken by the various stakeholder groups.

1.2 Background

1.2.1 Terminology

- *Multiple Uses (of ITS-Generated Data)* - Many transportation functions and applications that are not time sensitive can take advantage of ITS-generated data for use in nonreal-time. Generally, these functions are not related to the implementation of ITS control strategies, although certain types of ITS control strategies may use nonreal-time data to improve performance. *Multiple uses* refers to both the primary use for which the data were collected as well as additional uses of those data.

- *Metropolitan Planning Organization (MPO)* - The officially-designated agency in urban areas over 50,000 in population that is responsible for fulfilling the Federally-mandated transportation planning process.

- *Average Annual Daily Traffic (AADT)* - The total yearly traffic volume of a roadway divided by the number of days in the year. AADT is commonly estimated by factoring short-duration counts (1-2 days) to account for daily and seasonal variation.

- *K-factor*- the percent of daily traffic (AADT) that is present on a roadway during the peak hour of the day, both directions combined.

- *D-factor*- the percent of peak hour traffic that moves in the peak direction. Both K- and D-factors are developed from analyzing the 30th highest hour of traffic for a given year.

- *Vehicle-miles of travel (VMT).* The total number of miles traveled by vehicles over a given network and unit of time.

- *Travel Demand Forecasting (TDF) Models* - models used by transportation planners to forecast future demand for transportation facilities. The output from typical TDF models includes forecasted vehicle flows on a schematic network of the area's highway system and the demand for transit ridership between origin and destination points.

- *Highway Capacity Manual (HCM)* - a documented set of procedures used by transportation analysts to determine maximum traffic flow rates and vehicle speeds and delays for a given set of traffic and highway geometric conditions.

- *Intermodal Surface Transportation Efficiency Act (ISTEA).* Federal legislation passed in 1991 that outlines surface transportation policy and programs through 1997.

- *Transformed Data.* Data that are calculated from directly-measured data using other data or assumptions based on research findings. Examples include the conversion of spot speed data to segment travel time data and the calculation of equivalent single axle loads (ESALs) from truck weight and pavement data.

1.2.2 Multiple Uses of ITS-Generated Data: A Brief History

The idea that real-time data from traffic and transit operations could be archived and used for planning purposes has been expressed many times, but institutional and technical barriers have worked against it. As ITS has grown, the potential for ITS to provide such data has been voiced in various contexts, for example, the Highway Performance Monitoring System (HPMS) Steering Committee on August 7, 1996. This was also one of the findings of a conference March 2-5, 1997, in Irvine, CA, on "Information Needs to Support State and Local Transportation Decision Making into the 21st Century." The mid-year meeting of the TRB planning committees on July 21, 1997, in Portland, ME, had a session on this topic.

Meanwhile within the USDOT, on March 27, 1997, the FHWA Office of Highway Information Management, FHWA Office of Environment and Planning, and the Joint Program Office for ITS sponsored a meeting to discuss the use of ITS for collecting data for planning, research, performance monitoring, and policy purposes. The participants discussed the potential for ITS in providing the data needed for planning, performance monitoring, and other transportation activities. They were concerned that currently there are no guidelines or specifications that cover the collection, manipulation, and retention of data generated by ITS for use in other transportation activities. One of the action items that resulted was to investigate the feasibility of establishing a new "planning" ITS user service and standards for the flow of data into the planning process. In subsequent meetings within USDOT, other modes indicated their interest and a workshop was proposed in order to assemble the stakeholders in ITS-generated data.

1.2.3 Stakeholders for ITS-Generated Data

Several stakeholder groups have been identified as having an interest in the use of data generated by ITS. Table 1.1 introduces these stakeholder groups along with their primary functions. (Chapter 2 expands the discussion of how data generated by ITS can be used to support specific stakeholder functions.) The number of stakeholder groups with an interest in archived data is much larger than for any other User Service, and is an indication of the complexity of the problem. The data needs of these stakeholders -- and how they can be met with data generated by ITS -- are the basis for the recommendations put forth in this document. It is noteworthy that significant beneficiaries of archived data are Traffic Management operators, whose systems collect the data in the first place. In addition to aiding current operations by establishing pre-determined operation plans (e.g., ramp metering rates), use of archived data will allow Traffic Management operators to move to the next level of control strategies: proactive plans that intervene prior to conditions worsening. This next level of control is sometimes referred to as modeling support for traveler information and traffic management, and it is expected to take on greater significance to the transportation profession in the near future.

Table 1.1. Stakeholders for Data Generated by ITS

Stakeholder Group	Primary Transportation-Related Functions	Example Applications
MPO and state transportation planners	Identifying multimodal passenger transportation improvements (long- and short-range); congestion management; air quality planning; develop and maintain forecasting and simulation models	• congestion monitoring • link speeds for TDF and air quality models • AADT, K- and D-factor estimation • temporal traffic distributions • truck travel estimation by time of day • macroscopic traffic simulation • parking utilization and facility planning • HOV, paratransit, and multimodal demand estimation • congestion pricing policy
Traffic management operators	Day-to-day operations of deployed ITS (e.g., Traffic Management Centers, Incident Management Programs)	• pre-planned control strategies (ramp metering and signal timing) • highway capacity analysis • saturation flow rate determination • microscopic traffic simulation -- historical -- short-term prediction of traffic conditions • dynamic traffic assignment • incident management • congestion pricing operations
Transit operators	Day-to-day transit operations: scheduling, route delineation, fare pricing, vehicle maintenance; transit management systems; evaluation and planning	• capital planning and budgeting • corridor analysis planning • financial planning • maintenance planning • market research • operations/service planning • performance analysis planning • strategic/business planning
Air quality analysts	Regional air quality monitoring; transportation plan conformity with air quality standards and goals	• emission rate modeling • urban airshed modeling
MPO/state freight and intermodal planners	Planning for intermodal freight transfer and port facilities	• truck flow patterns (demand by origins and destinations) • HazMat and other commodity flow patterns
Safety planners and administrators	Identifying countermeasures for general safety problems or hotspots	• safety reviews of proposed projects • high crash location analysis • generalized safety relationships for vehicle and highway design • countermeasure effectiveness (specific geometric and vehicle strategies) • safety policy effectiveness
Maintenance personnel	Planning for the rehabilitation and replacement of pavements, bridges, and roadside appurtenances; scheduling of maintenance activities	• pavement design (loadings based on ESALs) • bridge design (loadings from the "bridge formula") • pavement and bridge performance models
Commercial vehicle enforcement personnel	Accident investigations; enforcement of commercial vehicle regulations	• HazMat response and enforcement • congestion management • intermodal access • truck route designation and maintenance • truck safety mitigation • economic development
Emergency management services (local police, fire, and emergency medical)	Response to transportation incidents; accident investigations	• labor and patrol planning • route planning for emergency response • emergency response time planning • crash data collection

Stakeholder Group	Primary Transportation-Related Functions	Example Applications
Transportation researchers	Development of forecasting and simulation models and other analytic methods; improvements in data collection practices	• car-following and traffic flow theory development • urban travel activity analysis
Private sector users	Provision of traffic condition data and route guidance (Information Service Providers); commercial trip planning to avoid congestion (carriers)	

1.2.4 Overview of ITS Data Flows

The foundation for identifying ITS-generated data relevant for multiple purposes is the National ITS Architecture prepared for FHWA by several contractors. It has been produced as a series of documents totaling more than 5,000 pages. As stated in the Executive Summary Document:

> ITS technologies have been encapsulated in a collection of interrelated user services for application to the nation's surface transportation problems. To date, 30 user services have been identified, the most recent being the Highway Rail Intersection. This list of user services is neither exhaustive nor final. The user services have been bundled into six categories as shown (...in Table 1.2).
>
> ITS presents stakeholders with a variety of options to address their transportation needs. Left without adequate guidance, stakeholders could easily develop systems solutions to their needs which were incompatible with their regional neighbors. Put another way, if City A chooses to implement User Services one way, and a neighboring City B another, then it is a real possibility that a motorist/traveler would find that none of the ITS equipment or services purchased for use in City A would work in City B. To fully maximize the potential of ITS technologies, system design solutions must be compatible at the system interface level in order to share data, provide coordinated, area-wide integrated operations, and support interoperable equipment and services where appropriate. The National ITS Architecture provides overall this guidance to ensure system, product, and service compatibility and interoperability, without limiting the design options of the stakeholder.
>
> The National ITS Architecture provides a common structure for the design of intelligent transportation systems. It is not a system design nor is it a design concept. What it does is define the framework around which multiple design approaches can be developed, each one specifically tailored to meet the individual needs of the user, while maintaining the benefits of a common architecture noted above. The architecture defines the functions (e.g., gather traffic information or request a route) that must be performed to implement a given user service, the physical entities or subsystems where these functions reside (e.g., the roadside or the vehicle), the interfaces/information flows between the physical subsystems, and the communication requirements for the information flows (e.g., wireline or wireless). In addition, it identifies and specifies the requirements for the standards needed to support national and regional interoperability, as well as product standards needed to support economy of scale considerations in deployment.

Table 1.2. User Services in the National ITS Architecture

User Services Bundle	User Services
Travel and Transportation Management	• En-Route Driver Information • Route Guidance • Traveler Services Information • Traffic Control • Incident Management • Emissions Testing and Mitigation • Demand Management and Operations • Pre-trip Travel Information • Ride Matching and Reservation • Highway Rail Intersection
Public Transportation Operations	• Public Transportation Management • En-Route Transit Information • Personalized Public Transit • Public Travel Security
Electronic Payment	• Electronic Payment Services
Commercial Vehicle Operations	• Commercial Vehicle Electronic Clearance • Automated Roadside Safety Inspection • On-board Safety Monitoring • Commercial Vehicle Administration Processes • Hazardous Materials Incident Response • Freight Mobility
Emergency Management	• Emergency Notification and Personal Security • Emergency Vehicle Management
Advanced Vehicle Control and Safety Systems	• Longitudinal Collision Avoidance • Lateral Collision Avoidance • Intersection Collision Avoidance • Vision Enhancement for Crash Avoidance • Safety Readiness • Pre-Crash Restraint Deployment • Automated Highway System

Specification of User Services is the method by which stakeholder needs are addressed by the National ITS Architecture. To develop functions that meet these needs, the National ITS Architecture converts the User Services into a large series of **subsystems** that are connected by real-time **data flows**. Each subsystem performs a unique set of functions. Data are collected by the subsystems and then shared through the data flows with other subsystems. Thus, each data flow in the National ITS Architecture has been created to assist the control functions of the ITS subsystems. An example of a data flow diagram is given as Figure 1.1 for Emergency Management. It shows the various subsystems that are involved with Emergency Management as well as the flow of data between the subsystems. Note that the data flows in this diagram represent general types of data; these are further broken down into specific data elements elsewhere in the National ITS Architecture. The general nature of the data flows and their associated data dictionary maintains the purpose of the National ITS Architecture, namely, that it serve as a blueprint for guidance in developing ITS; it is not meant to imply any specific system design.

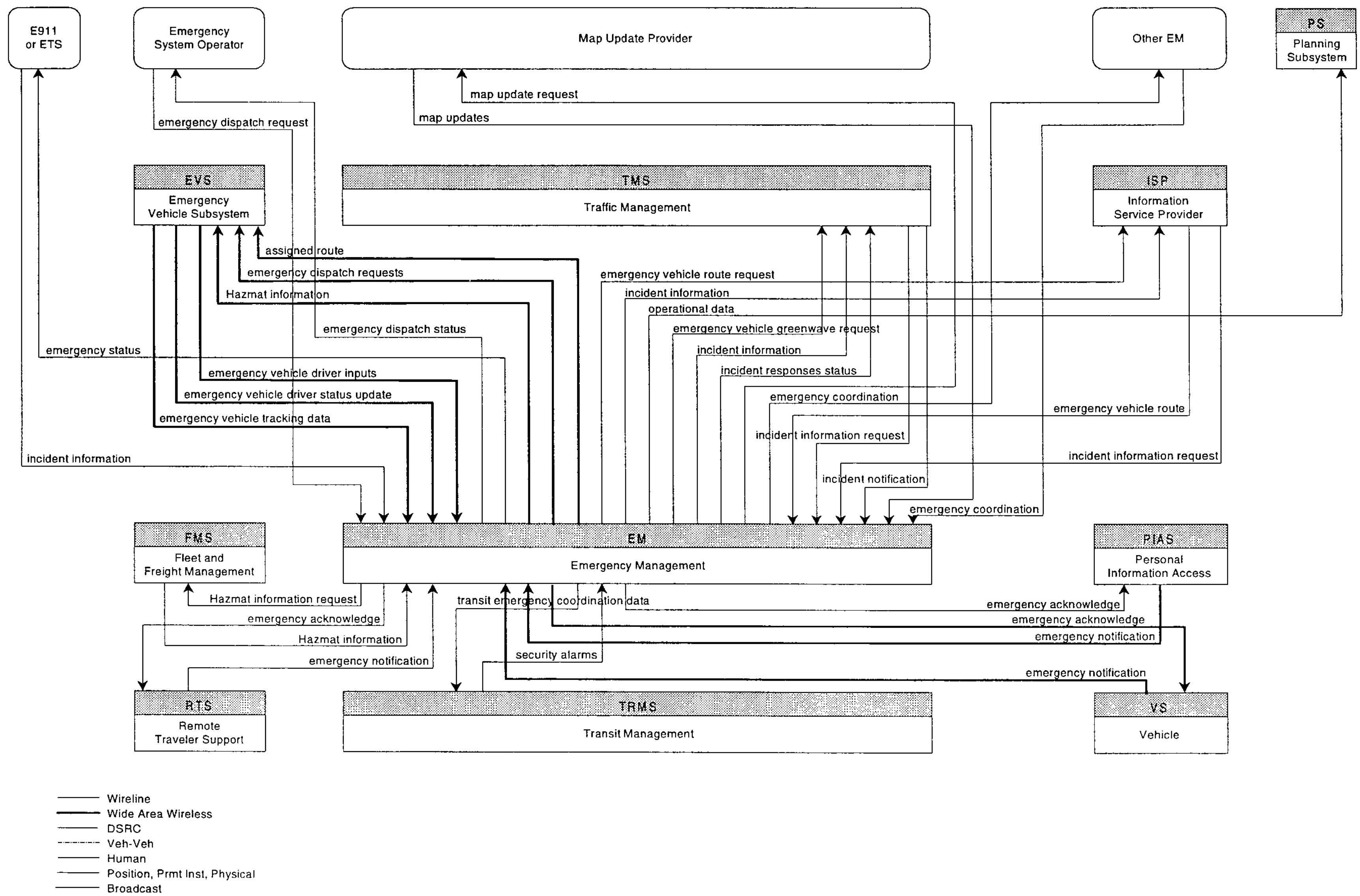


Figure 1.1. Architecture Flow Diagram for Emergency Management

A Planning Subsystem has also been defined in the National ITS Architecture (Figure 1.2). In its present form, the Planning Subsystem focuses on data for evaluating the performance of ITS and planning for future ITS. It is clear that the current Planning Subsystem is not as robust as required to meet the myriad needs of transportation planners. In addition, other stakeholders (see above) could benefit from use of the data and their needs should be considered. Therefore, it is anticipated that the formal Archived Data User Service for the National ITS Architecture being specified here would be used to modify or replace the current Planning Subsystem to account for the needs of other stakeholders.

The data flow diagrams in the National ITS Architecture serve as the initial basis for identifying ITS-generated data for multiple uses. For the most part, the data flows are based in real-time (or very close to real-time) and are used by the subsystems very shortly after they are collected. Multiple uses of the data for many forms of transportation planning, operations, and research will likely not need the data in any form approaching real-time and may not need the level of temporal detail represented in the raw data flows. Therefore, a key component of the Archived Data User Service is the structure for processing and storing data, both of which are discussed in detail in Chapter 4.

Table 1.3 presents a brief description of the selected data sources and representative data flows (i.e., data elements). In constructing Table 1.3, an attempt was made to match data generated by ITS with the needs of stakeholder groups. Thus, it is neither an exhaustive list of available data generated by ITS nor stakeholder needs. (For a complete discussion of data generated by ITS, the reader is referred to the National ITS Architecture documents. Appendix A provides recent efforts to document stakeholder needs.) The data sources correspond to subsystems in the National ITS Architecture but could be expanded to account for data sources not previously identified by that effort. (Some ITS deployments preceded the release of the National ITS Architecture.) The data elements shown in Table 1.3 roughly correspond to the those in the National ITS Architecture's data dictionary and can be used as a starting point for verifying that the National ITS Architecture accounts for all relevant data flows. These data elements are the basis for specific recommendations on the User Service structure presented in Chapter 4.

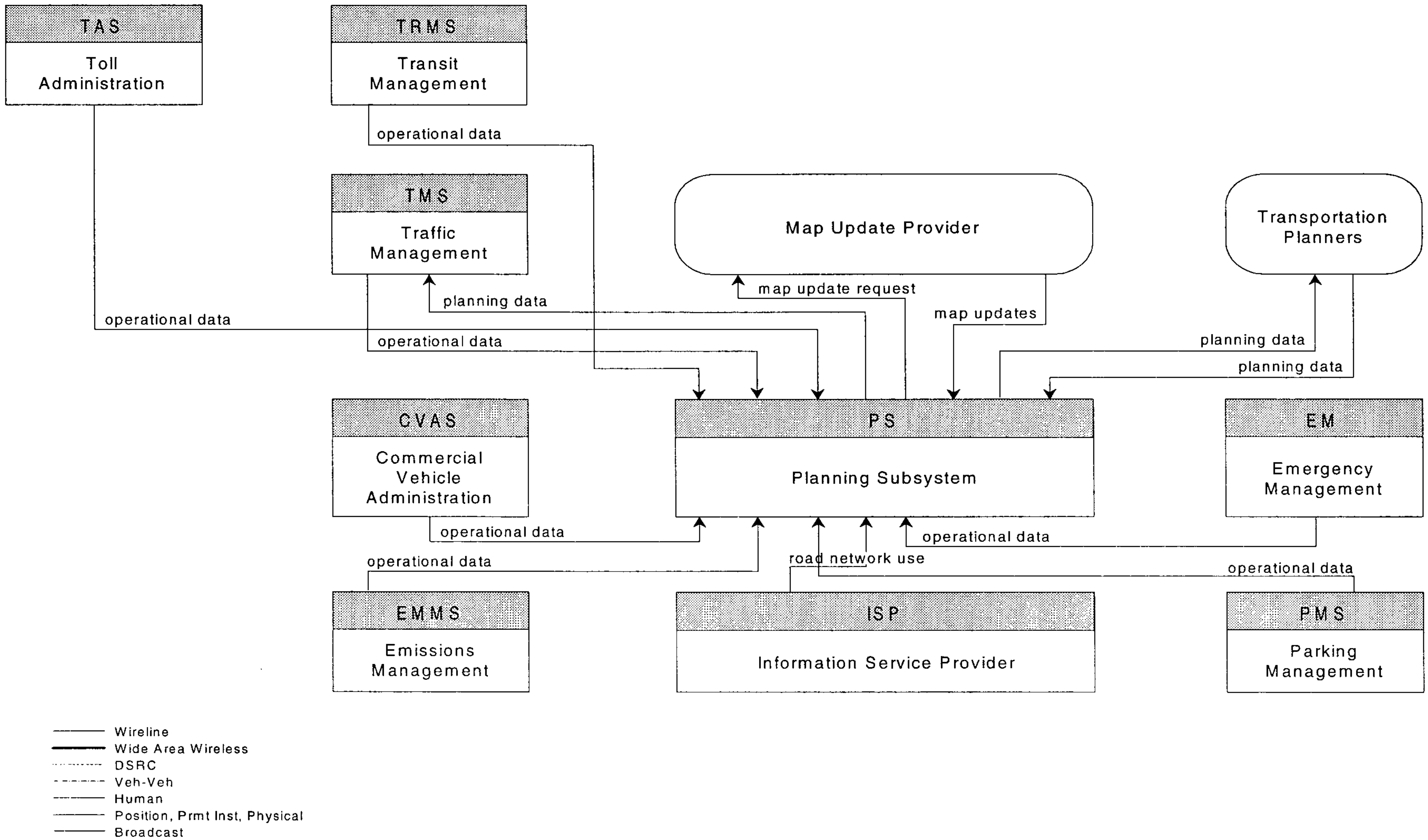


Figure 1.2. Architecture Flow Diagram for the Planning Subsystem

Table 1.3. ITS Data Relevant for Multiple (Nonreal-Time) Uses

Ref. No.	ITS data source	Primary data elements	Features of the Data Source: Typical collection equipment	Spatial coverage	Temporal coverage	Real-time uses	Possible multiple uses of ITS-generated data
FREEWAY AND TOLL COLLECTION							
1	Freeway traffic flow surveillance data	• volume • speed • occupancy	• loop detectors • video imaging • acoustic • radar/ microwave	Usually spaced at <= 1 mile; by lane	Sensors report at 20-60 second intervals	• ramp meter timing • incident detection • congestion/queue identification	• Congestion monitoring • Link speeds for TDF and air quality models • AADT, K- and D-factors • Saturation flow rates • Pre-planned TMC operations
		• vehicle classification • vehicle weight	• loop detectors • WIM equipment • video imaging • acoustic	Usually 50-100 per state; by lane	Usually hourly	Pre-screening for weight enforcement	• Truck percents by time of day for TDF and air quality models • Truck flow patterns • Pavement loadings
2	Ramp meter and traffic signal preemptions	• time of preemption • location	Field controllers	At traffic control devices only	Usually full-time	Priority to transit, HOV, and EMS vehicles	Network details for microscopic traffic simulation models (e.g., TRAF, TRANSIMS)
3	Ramp meter and traffic signal cycle lengths	• begin time • end time • location • cycle length	Field controllers	At traffic control devices only	Usually full-time	Adapt traffic control response to actual traffic conditions	• Network details for microscopic traffic simulation models (e.g., TRAF,TRANSIMS) • Pre-planned TMC operations
4	Visual and video surveillance data	• time • location • queue length • vehicle trajectories • vehicle classification • vehicle occupancy	• CCTV • aerial videos • image processing technology	Selected locations	Usually full-time	• coordinate traffic control response • congestion/queue identification • incident verification	• Congestion monitoring • Car-following and traffic flow theory
5	Vehicle counts from electronic toll collection	• time • location • vehicle counts	Electronic toll collections equipment	At instrumented toll lanes	Usually full-time	Automatic toll collection	Traffic counts by time of day
6	TMC generated traffic flow metrics (forecasted or transformed data)	• link congestion indices • stops/delay estimates	TMC software	Selected roadway segments	Hours of TMC operation	• incident detection • traveler information • preemptive control strategies	• congestion monitoring • effectiveness of prediction methods
ARTERIAL AND PARKING MANAGEMENT							
7	Arterial traffic flow surveillance data	• volume • speed • occupancy	• loop detectors • video imaging • acoustic • radar/ microwave	Usually midblock at selected locations only ("system detectors")	Sensors report at 20-60 second intervals	• progression setting • congestion/queue identification	• congestion monitoring • link speeds for travel forecasting models (free flow only) • AADT, K- and D-factors

Table 1.3. ITS Data Relevant for Multiple (Nonreal-Time) Uses

Ref. No.	ITS data source	Primary data elements	Features of the Data Source: Typical collection equipment	Spatial coverage	Temporal coverage	Real-time uses	Possible multiple uses of ITS-generated data
8	Traffic signal phasing and offsets	• begin time • end time • location • up/down-stream offsets	Field controllers	At traffic control devices only	Usually full-time	Adapt traffic control response to actual traffic conditions	Network details for microscopic traffic simulation models (e.g., TRAF, TRANSIMS)
9	Parking management	• time • lot location • available spaces	Field controllers	Selected parking facilities	Usually day time or special events	Real-time information to travelers on parking availability	Parking utilization and needs studies
TRANSIT AND RIDESHARING							
10	Transit usage	• vehicle boardings (by time and location) • station origin and destination (O/D) • paratransit O/D	Electronic fare payment systems	Transit routes	Usually full-time	Used for electronic payment of transit fares	• route planning/ run-cutting • ridership reporting (e.g., Section 15)
11	Transit route deviations and advisories	• route number • time of advisory • route segments taken	TMC software	Transit routes	Usually full-time	Transit route revisions	Transit route and schedule planning
12	Rideshare requests	• time of day • O/D	CAD	Usually areawide	Day time, usually peak periods	Dynamic rideshare matching	• travel demand estimation • transit route and service planning
INCIDENT MANAGEMENT AND SAFETY							
13	Incident logs	• location • begin, notification, dispatch, arrive, clear, depart times • type • extent (blockage) • HazMat • Police accident rpt. reference • cause	• CAD • computer-driven logs	Extent of Incident Manage-ment Program	Extent of Incident Management Program	Incident response and clearance	• incident response evaluations (program effectiveness) • congestion monitoring (e.g., % recurring vs nonrecurring) • safety reviews (change in incident rates
14	Train arrivals at Highway Rail Intersections	• location • begin time • end time	Field controllers	At instru-mented HRIs	Usually full-time	• coordination with nearby traffic signals • notification to travelers	Grade crossing safety and operational studies

Table 1.3. ITS Data Relevant for Multiple (Nonreal-Time) Uses

Ref. No.	ITS data source	Primary data elements	Features of the Data Source: Typical collection equipment	Spatial coverage	Temporal coverage	Real-time uses	Possible multiple uses of ITS-generated data
15	Emergency vehicle dispatch records	• time • O/D • route • notification, arrive, scene, leave times	CAD	Usually areawide	Usually full-time	Coordination of Emergency Management response	• Emergency management labor and patrol studies • Emergency management route planning
16	Emergency vehicle locations	• vehicle type • time • location • response type	Automatic Vehicle Identification (AVI) or GPS equipment	Usually areawide	Usually full-time	• tracking vehicle progress • green wave and signal preemption initiation	• Emergency management route planning • Emergency management response time studies
17	Construction and work zone identification	• location • date • time • lanes/ shoulders blocked	TMC software			Traveler information	Congestion monitoring
COMMERCIAL VEHICLE OPERATIONS							
18	HazMat cargo identifiers	• type • container/ package • route • time	CVO systems	At reader and sensor locations	Usually full-time	• Identifying HazMat in specific incidents • routes for specific shipments	• HazMat flows • HazMat incident studies
19	Fleet Activity Reports	• carrier • citations • accidents • inspection results	CVO inspections	N/A	Usually summarized annually	May overlap with SAFETYNET functions	
20	Cargo identification	• cargo type • O/D	CVO systems	At reader and sensor locations	Usually full-time	Clearance activities	Freight movement patterns
21	Border crossings	• counts by vehicle type • cargo type • O/D	CVO systems	At reader and sensor locations	Usually full-time	Enforcement	Freight movement patterns
22	On-board safety data	• vehicle type • cumulative mileage • driver log (hrs. of service) • subsystem status (e.g., brakes)	CVO systems	At reader and sensor locations	Usually full-time	Enforcement and inspection	Special safety studies (e.g., driver fatigue, vehicle components)
ENVIRONMENTAL AND WEATHER							
23	Emissions Management System	• time • location • pollutant concentrations • wind conditions	Specialized sensors	Sensor locations	Usually full-time	Identification of hotspots and subsequent control strategies	• trends in emissions • special Air Quality studies

Table 1.3. ITS Data Relevant for Multiple (Nonreal-Time) Uses

Ref. No.	ITS data source	Primary data elements	Features of the Data Source: Typical collection equipment	Features of the Data Source: Spatial coverage	Features of the Data Source: Temporal coverage	Real-time uses	Possible multiple uses of ITS-generated data
24	Weather data	• location • time • precipitation • temperature • wind conditions	Environmental sensors	At sensor locations	Usually full-time	Traveler information	• congestion monitoring (capacity reductions) • freeze/thaw cycles for pavement models
VEHICLE AND PASSENGER INFORMATION							
25	Location referencing data	Special case; pertains to all location references in ITS and planning					Need conversion from lat/long to highway distance and location (e.g., milepost references for queue lengths)
26	Probe data	• vehicle ID • segment location • travel time	• probe readers and vehicle tags • GPS on vehicles	GPS is areawide; readers restricted to highway locations	Usually full-time	• coordinate traffic control response • congestion/queue identification • incident detection • real-time transit vehicle schedule adherence • electronic toll collection	• congestion monitoring • link speeds for travel forecasting models • historic transit schedule adherence • traveler response to incidents or traveler information • O/D patterns
27	VMS messages	• VMS location • time of msg • msg content	TMC software	VMS locations	Hours of TMC operation	Traveler information	Effects of VMS message content on traveler response
28	Vehicle trajectories	• location (route) • time • speed • acceleration • headway	• AVI or GPS equipment • advanced video image processing	AVI restricted to reader locations; GPS is areawide	1-10 second intervals	Collected as part of surveillance function	• Traffic simulation model calibration for local conditions (driver type distributions) • Modal emission model calibration • Traffic flow research
29	TMC and Information Service Provider generated route guidance	• time/date • O/D • route segments • estimated travel time	TMC/Informatio n Service Provider software	Usually areawide	Hours of TMC operation	Traveler information	• O/Ds for travel demand forecasting model inputs • Interzonal travel times for travel demand forecasting model calibration
30	Parking and roadway (congestion) pricing changes	• time/date • rte. segment/ lot ID • new price	TMC software	Facilities subject to variable pricing	Hours of TMC operation	Demand management	• Special studies of traveler response to pricing • Establishment of pricing policies

2. THE NEED FOR AN ARCHIVED DATA USER SERVICE

2.1 Overview

The planning, operation, and evaluation of transportation systems have always necessitated data in many forms. In recent years, formal requirements (e.g., air quality conformity) and a shift toward a more short-term planning and management perspective (e.g., congestion management) have increased the data needs of transportation personnel. At the same time, resources for the maintenance of data programs have been limited. The need to do "more with less" is apparent in most transportation agencies. However, the advent of ITS offers a remedy to some of the data problems of transportation personnel. In most cases, ITS-generated data are fundamentally the same type of data traditionally collected by planners, operators, and researchers but are much more detailed in their temporal and spatial coverage. In addition, ITS-generated data are collected automatically without the need for manual data collection efforts. The increased detail allows more precise and wide-ranging analyses to be performed and may obviate the need for other special data collection efforts to be funded.

Prior to examining how data generated by ITS can be used by stakeholders in the performance of their functions, a review of data needs and opportunities is in order. Several recent efforts have focused on this topic and are reviewed in more detail in Appendix A. These efforts include:

- *Information Needs to Support State and Local Transportation Decision Making Into the 21st Century.*

- *Guidance Manual for Managing Transportation Planning Data (NCHRP Report 401).*

- *Highway Performance Monitoring System (HPMS) Reassessment Workshop.*

Most recently (January 1998), the *ITS As A Data Resource Workshop*, a joint effort between ITS America and USDOT, brought many of the stakeholders together to discuss how data generated by ITS could be used to address many of their data needs. (See summary in Appendix A.) This Workshop was instrumental in providing guidance for this document. Moreover, a major outcome of the Workshop was the need to define a formal User Service for the National ITS Architecture as a means to get the issues formally defined and highlighted. It must be pointed out that the Archived Data User Service is fundamentally different from other User Services in the National ITS Architecture in that its "control" function lies outside of ITS. While multiple uses of data are within the already established functions of the stakeholders (e.g., transportation planning and traditional transit operations), they are outside the scope of the current National ITS Architecture.

All of these previous efforts identify how the data needs of planners, operators, and researchers have expanded due to increased demands placed on them. Appendix A provides a list of stakeholder data needs; these were used to develop the recommendations in this document. They also note the shrinking budgets that stakeholders must manage for data programs, and also cite ITS as a potential source for many of the data required to meet needs. The starting point in this process is the specification for a formal User Service (ADUS) to be incorporated into the National ITS

Architecture. However, as is discussed in the next Chapter, this is a necessary but not sufficient condition for successful implementation.

2.2 Uses and Benefits of Archived Data Generated By ITS

Table 2.1 presents a listing of how data generated by ITS can be used to supplement or replace existing data for selected stakeholder applications. The point that data generated by ITS and "traditional" data are similar is highlighted in this Table (e.g., traffic counts) but data generated by ITS are collected continuously. Figure 2.1 reverses the emphasis to show how one type of data (freeway surveillance data) can be used for a variety of purposes depending on the level of aggregation used. In addition to the aggregation levels shown, 15-minute and peak period (multiple hour) summaries are also used. Accordingly, several general observations can be made concerning the use of data generated by ITS by stakeholders.

- The continuous nature of most data generated by ITS removes temporal sampling bias from estimates and allows the study of variability. Nearly all of the data currently collected for planning, operations, administration, and research applications are through the use of sample surveys (e.g., household travel surveys, short-duration traffic counts). Although attempts are made to adjust or expand the sample, the procedures are imperfect. With continuous data, there is no need to perform adjustments to control sample bias. (Equipment or nonresponse errors are still present, though). Further, continuous data allows the direct study of variability, which is becoming an important factor in the study of personal travel habits and the effect of extreme events (e.g., days with very high volumes).

- Data to meet emerging requirements and for input to new modeling procedures will have to be more detailed than what is now collected. The next generation of Travel Demand Forecasting (TDF) models (e.g., TRANSIMS) and air quality models (modal emission models) will operate at a much higher level of granularity than existing models. Traditional data sources are barely adequate for existing models -- there is little doubt that they will be incapable of supporting the next generation of models. Much data generated by ITS are collected at the levels of detail necessary to support these models. For example, roadway surveillance data (volumes, speeds, and occupancies) are typically reported every 20 seconds and GPS-instrumented vehicles can report positions and activity at time intervals as small as one second. Also, GPS-derived locations can pinpoint incident locations to within a few meters. This level of detail will be required for the input and calibration data used by the new models. Finally, as data generated by ITS are used more frequently for nonreal-time purposes, it is likely that additional uses not currently foreseen will emerge. In addition, data on activity patterns and how travelers respond to system condition will be important for the next generation of models.

- Use of data generated by ITS for multiple purposes is a way to stimulate the support of other stakeholders for ITS initiatives. If groups besides those involved in ITS development see value in data generated by ITS, they will be inclined to learn more about ITS and to support deployment. Mutual interest in data generated by ITS will stimulate cooperation among stakeholders. This could prove to be extremely valuable in the "main streaming" of ITS into standard transportation practice, particularly among transportation planners.

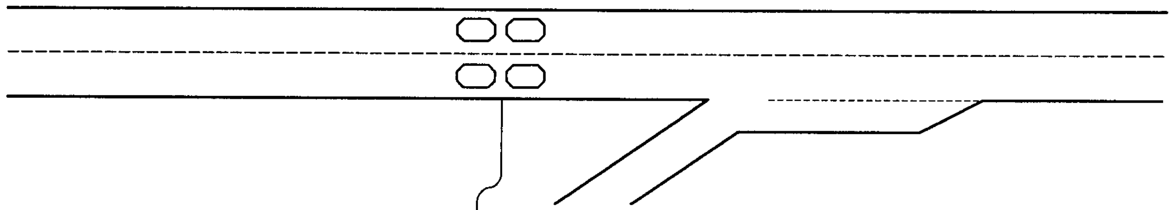
Table 2.1. Uses of ITS Data for Stakeholder Applications

Stakeholder Group	Application	Method or Function	Collection and Use of: Current Data	Collection and Use of: ITS-Generated Data
MPO and State Transportation Planners	Congestion Management Systems	Congestion Monitoring	Travel times collected by "floating cars": usually only a few runs (small samples) on selected routes. Speeds and travel times synthesized with analytic methods (e.g., *HCM*, simulation) using limited traffic data (short counts). Effect of incidents missed completely with synthetic methods and minimally covered by floating cars.	Roadway surveillance data (e.g., loop detectors) provide continuous volume counts and speeds. Variability can be directly assessed. Probe vehicles provide same travel times as "floating cars" but greatly increase sample size and areawide coverage. The effect of incidents is imbedded in surveillance data and Incident Management Systems provide details on incident conditions.
	Long-Range Plan Development	Travel Demand Forecasting Models	Short-duration traffic counts used for model validation. O/D patterns from infrequent travel surveys used to calibrate trip distribution. Link speeds based on speed limits or functional class. Link capacities usually based on functional class.	Roadway surveillance data provide continuous volume counts, truck percents, and speeds. Probe vehicles can be used to estimate O/D patterns without the need for a survey. The emerging TDF models (e.g., TRANSIMS) will require detailed data on network (e.g., signal timing) that can be collected automatically via ITS. Other TDF formulations that account for variability in travel conditions can be calibrated against the continuous volume and speed data.
	Corridor Analysis	Traffic Simulation Models	Short-duration traffic counts and turning movements used as model inputs. Other input data to run the models collected through special efforts (signal timing). Very little performance data available for model calibration (e.g., incidents, speeds, delay).	Most input data can be collected automatically and models can be directly calibrated to actual conditions.
Traffic Management Operators	ITS Technology	Program and Technology Evaluations	Extremely limited; special data collection efforts required.	Data from ITS provide the ability to evaluate the effectiveness of both ITS and non-ITS programs. For example, data from an Incident Management System can be used to determine changes in verification, response, and clearance times due to new technologies or institutional arrangements. Freeway surveillance data can be used to evaluate the effectiveness of ramp meters or HOV restrictions.
		Pre-Determined Control Strategies	Short-duration traffic counts and "floating car" travel time runs. A limited set of pre-determined control plans is usually developed mostly due to the lack of data.	Continuous roadway surveillance data makes it possible to develop any number of pre-determined control strategies.
		Predictive Traffic Flow Algorithms	Extremely limited.	Analysis of historical data form the basis of predictive algorithms: "What will traffic conditions be in the next 15 minutes?" (Bayesian approach).
Transit Operators	Operations Planning	Routing and Scheduling	Manual travel demand and ridership surveys; special studies.	Electronic Fare Payment System and Automatic Passenger Counters allow continuous boardings to be collected. Computer-aided dispatch systems allow O/D patterns to be tracked. AVI on buses allows monitoring of schedule adherence and permits the accurate setting of schedules without field review.

Table 2.1. Uses of ITS Data for Stakeholder Applications

Stakeholder Group	Application	Method or Function	Collection and Use of: Current Data	Collection and Use of: ITS-Generated Data
Air Quality Analysts	Conformity Determinations	Analysis with the MOBILE Model	Areawide speed data taken from TDFs. VMT and vehicle classifications derived from short counts.	Roadway surveillance provides actual speeds, volumes, and truck mix by time of day. Modal emission models will require these data in even greater detail and ITS is the only practical source.
MPO/State Freight and Intermodal Planners	Port and Intermodal Facilities Planning	Freight Demand Models	Data collected through rare special surveys or implied from national data (e.g., Commodity Flow Survey).	Electronic credentialing and AVI allows tracking of truck travel patterns, sometimes including cargo. Improved tracking of congestion through the use of roadway surveillance data leads to improved assessments of intermodal access.
Safety Planners and Administrators	Safety Management Systems	Areawide Safety Monitoring; Studies of Highway and Vehicle Safety Relationships	Exposure (typically VMT) derived from short-duration traffic and vehicle classification counts; traffic conditions under which crashes occurred must be inferred. Police investigations, the basis for most crash data sets, performed manually.	Roadway surveillance data provide continuous volume counts, truck percents, and speeds, leading to improved exposure estimation and measurement of the actual traffic conditions for crash studies. ITS technologies also offer the possibility of automating field collection of crash data by police officers (e.g., GPS for location).
Maintenance Personnel	Pavement and Bridge Management	Historical and Forecasted Loadings	Volumes, vehicle classifications, and vehicle weights derived from short-duration counts (limited number of continuously operating sites).	Roadway surveillance data provide continuous volume counts, vehicle classifications, and vehicle weights, making more accurate loading data and growth forecasts available.
Commercial vehicle enforcement personnel	Enforcement of Commercial Vehicle Regulations	Hazardous Material Inspections and Emergency Response	Extremely limited.	Electronic credentialing and AVI allows tracking of hazardous material flows, allowing better deployment of inspection and response personnel.
Emergency Management Services (local police, fire, and emergency medical)	Incident Management	Emergency Response	Extremely limited.	Electronic credentialing and AVI allows tracking of truck flows and high incident locations, allowing better deployment of response personnel.
Transportation Researchers	Model Development	Travel Behavior Models	Mostly rely on infrequent and costly surveys: stated preference and some travel diary efforts (revealed preference).	Traveler response to system conditions can be measured through system detectors, probe vehicles, or monitoring in-vehicle and personal device use. Travel diaries can be imbedded in these technologies as well.
		Traffic Flow Models	Detailed traffic data for model development must be collected through special efforts.	Roadway surveillance data provide continuous volume counts, densities, truck percents, and speeds at very small time increments. GPS-instrumented vehicles can provide second-by second performance characteristics for microscopic model development and validation.
Private Sector Users	Truck Routing and Dispatching	Congestion Monitoring	Current information on real-time or near real-time congestion is extremely limited.	Roadway surveillance data and probe vehicles can identify existing congestion and can be used to show historical patterns of congestion by time-of-day. Incident location and status can be directly relayed.
	Information Service Providers	Trip Planning	Information on historical congestion patterns is extremely limited. This information could be used in developing pre-trip route and mode choices, either alone or in combination with real-time data.	

Table 2.1. Uses of ITS Data for Stakeholder Applications



Level of Aggregation → *Typical Applications*

20-second

Time	Speed	Volume	Density
8:00:00	30.1	21	100
8:00:20	32.3	22	105
8:00:40	31.4	20	98

- Microscopic Simulation Models
- Car-Following Theory

5-minute

Time	Speed	Volume	Density
8:00	33.0	333	90
8:05	30.9	328	87
8:10	28.1	325	88

- Macroscopic Simulation Models
- Traffic Flow Theory
- Microscale Air Quality Models

1-hour

Time	Speed	Volume	Density
8:00	35.4	4,110	85
9:00	47.8	4,200	60
10:00	52.9	3,800	42

- Capacity Analysis
- TDF inputs and checks
- Congestion Monitoring
- Macroscale Air Quality Models

Daily

Date	Speed	Volume
04/01/99	45.7	65,100
04/02/99	49.9	64,070
04/03/99	47.8	67,100

- Congestion Monitoring
- AADT estimation
- General Planning

Note: Density usually computed from loop occupancy.

Figure 2.1. ITS Traffic Surveillance Data is useful to Stakeholders at many levels of Aggregation

- Promoting the use of archived data for multiple purposes complements the initiative for integrating ITS in general. To a very large degree, integration of ITS components can be viewed as the sharing and use of data between individual ITS components, usually in real- or near real-time. (For example, the transfer of freeway surveillance data for purposes beyond control of the freeway such as for traffic signal control and traveler information is a form of integration.) For integration to occur, system linkages must be established. It is precisely these linkages that can be tapped to archive data under the proposed User Service. Therefore, the Archived Data User Service can be thought of as another form of ITS integration -- the linking of ITS components with the rest of the transportation world.

- Because the data are already being collected for ITS control, other uses provide a value-added component to ITS. In an era of shrinking budgets and close public scrutiny, additional ways in which investments can be justified improve the chances for successful implementation. Put another way, marginal investments in a data archiving function can have substantial payoffs in terms of enhanced and more cost-effective data programs. The net effect is to reduce the burden on and/or augment the existing data collection programs of stakeholders.

- ITS is a rich data source for multiple uses, but not a panacea. In general, ITS should be seen as way to supplement existing data sources and will not be capable of meeting **all** stakeholder data needs. In some cases, ITS can replace existing programs. (For example, roadway surveillance equipment will eliminate the need for collecting traffic volume information on instrumented highway segments.) However, even in these cases, ITS will probably be focused on selected routes, leaving the need for data collection on the remaining system. As system coverage of ITS technologies and the market penetration of personal traveler information grows, the coverage issue will lessen but will not completely disappear.

- As the focus of transportation policy shifts away from large-scale, long-range capital improvements and toward better management of existing facilities, the creation and use of **system performance measures** is taking on greater significance. The specification of transportation management systems in ISTEA has created the need for more intense system performance monitoring than current data can adequately support; this is a common thread running through all of the stakeholder uses discussed in Table 2.1. Further, planners, operators, and administrators are increasingly being required to shorten their planning horizons. System performance measures provide objective feedback to transportation professionals on the effectiveness of programs and improvements, and also provide a common basis for comparing different jurisdictions. This kind of feedback is extremely important as the focus shifts to short-term management strategies. However, it is clear that data with higher resolution and accuracy than have been traditionally collected are required to support the use of system performance measures.

An example from transportation planning will illustrate these points. AADTs (i.e., daily traffic count estimates for a highway) are one of the most essential data types used by planners and engineers. Nearly all AADTs used by planners are estimates based on 24- or 48-hour short counts that have been adjusted using areawide factors for daily and seasonal variability. Facility-specific data on the temporal distribution of traffic and its variability are extremely limited. ITS roadway surveillance equipment can provide detailed data on the *actual* average daily traffic and its variability. This would improve the accuracy and usefulness of one of the core performance measures (AADT) used by transportation planners.

Further, more detailed data will be required as the management paradigm becomes more widespread. TDF models for predicting long-term demand characteristics for 20-years into the future work adequately with average values -- basically, one wants to make decisions about adding capacity to the nearest additional lane of accuracy (i.e., 2,200-2,300 vehicles per hour). Average peak hour traffic counts are precise enough for this purpose. However, for meeting the newer planning requirements which tend to be more short-range in nature -- such as congestion monitoring and microscale air quality modeling -- information on extreme events are important. For example, consider a freeway section where the only traffic data available are an average AADT developed from a factored 48-hour short count and K- and D-factors borrowed from other urban sites. A capacity analysis on this section using these data would not only ignore days where volumes were higher than average, but is prone to the sampling bias inherent in using factored and borrowed data. Since delay is a nonlinear function of volume as volumes approach and exceed capacity, these rare but highly influential events would be missed if the short-term and borrowed data were the basis for congestion monitoring. Moreover, the impact of incidents on delay would be completely ignored in the current approach. On the other hand, ITS roadway surveillance data would directly measure congestion, including days with abnormally high volumes and incidents.

It is also possible to extend this example into the realm of traffic operations. Although ITS generally uses real-time data to implement control patterns, nonreal-time data can also be of use. Consider that ramp metering is also present for the hypothetical freeway segment mentioned above. The metering rates are generally pre-timed, actuated by mainline traffic flow, or a combination. In the pre-timed case, data on historical volume and congestion patterns can be used to set metering rates by time of day. In advanced systems that are proactive (i.e., they predict traffic conditions in the very near future), historical patterns can be used in predictive algorithms. Finally, historical ramp metering rates and freeway traffic conditions are valuable to operators of traffic signal control systems in that pre-timed or proactive arterial timing plans can be developed with that data. From an archival viewpoint, the needs of operators would tend to be more short-term (what happened yesterday or last week) than those of transportation planners.

The examples cited above portray situations now faced by local transportation planners and operators. However, as system performance measures expand to include multimodal considerations, high resolution data for a wider variety of stakeholders will be in great demand. Because data generated by ITS are collected both at the system level and the level of the individual traveler, they will be extremely valuable in addressing the emerging need for multimodal performance measures (e.g., data on person movements, not just vehicles).

3. TECHNICAL AND INSTITUTIONAL ISSUES FOR IMPLEMENTATION

3.1 Introduction

Beyond defining the Archived Data User Service for inclusion in the National ITS Architecture, it is critical to outline the serious technical and institutional issues affecting its implementation. In fact, since most of the relevant data flows have already been defined in the National ITS Architecture, overcoming the technical and institutional issues is the major challenge for implementation. The institutional issues are particularly onerous since the beneficiaries of the data are primarily personnel not directly involved in ITS operations, leading to such questions as who bears the cost, who maintains the system, and who owns the data. This Chapter identifies these issues and the recommendations in Chapter 4 have been made with them in mind. However, complete guidance for resolving these issues will require case studies of current efforts as well as future efforts specifically targeted to the use of data generated by ITS for multiple purposes.

3.2 Technical and Institutional Issues

3.2.1 Development, Operation, and Maintenance Costs

Operating agencies such as state DOTs and local traffic engineering agencies are the purveyors of ITS. This means that funding for ITS comes from their budgets and unless they are convinced of the value of archiving the data, they may balk at including the archiving function in their systems. As an information management system, an implemented archived data system requires data base administration, backup procedures, routine operation of quality control and summarization programs, responding to special users, maintaining existing code, and developing new code for new applications. Even if the data "owners" are convinced of the data's value, staff resources for operation and maintenance may be slim. Moreover, because there is little precedent in the field, the costs of building, operating, and maintaining an archival system are largely unknown. Finally, the sharing of costs among stakeholder groups, while in theory could help defray costs to any one group, is a problem that needs to be worked out locally. It is therefore crucial that these costs be explicitly addressed in future funding of ITS deployments. Local agencies should also be free to pursue innovative approaches to paying for archived data systems, including the use of non-ITS funds (e.g., state and local planning allocations).

3.2.2 System Access

For the data to be of use to stakeholders, they must be easily accessible. This is largely a technical issue of how to design the communications and interconnects in the final system. However, access to the data by remote users must be accounted for in the system design. Mechanisms for data distribution depend on the arrangements made for access; they can range from free and open access to the data via direct communications or the Internet to periodic release of the data on CD-ROMs or other media. A related issue is the degree to which unverified data or data that have failed quality control checks will be released to certain users.

3.2.3 Ownership

If ITS operating agencies are responsible for the systems that archive data, and if those data are not really used or are not seen as having value by them, there is a concern that data quality and completeness will suffer. Ownership may also affect access to the data if the "owners" are sensitive to releasing data to others (even if the data are seen as high quality). In some cases, data in a central repository may have multiple owners, further complicating the institutional arrangements for data sharing. For example, freeway traffic surveillance data may be "owned" by the state DOT and transit ridership data may be "owned" by the local transit authority. Therefore, issues of ownership need to be addressed prior to field implementation.

3.2.4 Data Quality

Even if the issue of ownership does not affect data quality, development of quality control (QC) procedures for data generated by ITS has been limited. Data generated by ITS collected by field surveillance equipment are prone to errors due to equipment malfunctions, calibration "drift", and communication disruptions. Therefore, it is clear that raw data from the field must be subjected to quality control and editing procedures at some point. Two options exist: (1) perform QC and editing before they are sampled, summarized, and stored for later use, and (2) simply store all data as they are received from the field and leave QC up to the individual end users. The first option is preferred because "owners" of the data understand their data and are in the best position to perform QC. At a minimum, the conditions under which the data are collected and possible sources of error must be communicated to end users so they can make their own judgments; this dictates that full disclosure of equipment status and environmental conditions be documented.

An even more complex problem with data quality is the mechanism for correcting problems with data collection equipment, given that QC procedures have identified the problems. In some cases, data may be ancillary for control purposes. In others, the level of accuracy required for control may not be as high as for other applications. For example, for detecting freeway traffic breakdowns for VMS control, operators need to know if speeds are either free flow (around 50 mph) or forced flow (less than 30 mph). However, many models used by transportation planners require greater resolution of speeds, and researchers want speeds at even higher accuracy. The care and maintenance of field data collection equipment is therefore seen as an integral part of archived data systems process and must be explicitly addressed.

Other issues surrounding data quality include: detection procedures for questionable data; categorizing and flagging errors in the data; the use of missing or questionable data in the data aggregation process; geographic coverage of the data; full understanding of what the data truly represent; and the imputation and replacement of questionable data using *post hoc* quality control procedures. The stakeholder groups identified here are not used to reviewing, cleaning, and editing data and they will require guidance on this matter. Because little experience exists for multiple uses of data generated by ITS, quality control is a new facet for data owners as well.

3.2.5 Data Management

Data management issues relate to how much data should be kept and at what level of summarization or sampling should be present. The needs of various stakeholder groups are quite different in this respect. A specific example involving freeway traffic surveillance data was provided as Figure 2.1. It shows how these data, which are reported from field controllers at 20-second intervals, can be used for different purposes depending on the level of aggregation. For example, transportation planners may require roadway surveillance data (volumes and speeds) at 1-hour intervals summarized over all lanes of the facility. On the other hand, researchers investigating microscopic traffic flow may need surveillance data at the smallest possible intervals for individual traffic lanes. The temptation is to save the data at its lowest level of aggregation so that all stakeholders can be accommodated, but that may not be cost-effective and may discourage use for some applications. Alternative solutions to this problem include specifying different levels of aggregation in the archive system and storing sampled data at the lowest level of aggregation rather than saving it all.

A related data management issue is the degree to which transformed data should reside within the archived data system structure. (See Section 1.2.1 for a definition.) The example shown in Figure 2.1 is not transformed data -- it merely shows how data may be summarized to different levels of aggregation. Because transformed data are usually used for highly specific purposes, specification in the Archived Data User Service may not be cost-effective -- it is more properly a responsibility of the end user. However, some local circumstances may require that certain transformed data be permanently archived.

Another data management issue is how to link data from disparate sources. To some extent, the National ITS Architecture and the various standards development efforts (e.g., NTCIP) will help to provide conventional data definitions. However, a major problem still remains with the development of a common location referencing system. The fundamental problem is how to link the geo-coordinates (latitude/longitude or other x/y system) to points on the transportation network. For systems where the location of the data collection is fixed -- as with loop detectors or AVI (probe) readers -- location referencing is less of a problem than with systems that rely on GPS or other geo-coordinates. Fixed data collection equipment can be easily mapped one time with existing referencing systems (e.g., mileposts, distance to nearest intersection). However, as the use of advanced technologies such as in-vehicle navigation systems become more widespread, the location referencing problem takes on a greater significance. Even though many nonreal-time applications of the data do not require a great level of precision (e.g., locations of traffic counts for planning uses could be accurate to the nearest major intersection), the translation of geo-coordinates to points on the highway network is still highly problematic.

Finally, the mechanism for how the system for maintaining archived data generated by ITS will interface with legacy systems must be resolved. Standards for data definitions and communications can foster the interface but there are still large technical problems to be addressed, especially since many legacy systems were developed using dated data base management systems.

3.2.6 Data and Communications Standards

Standards and protocols are highly important because they foster interoperability of field equipment, aid retrieval of data, and allow data comparisons across areas. Wherever possible, standards for archived data should conform to existing standards for transportation data (e.g., the *Traffic Monitoring Guide*). Standards should include the specification of metadata which not only describe how the data are represented but identify the conditions under which they were collected. Also, many standard development efforts are currently underway that could affect archived data systems. Coordination with these efforts is essential to successful implementation.

3.2.7 Privacy Concerns

Some forms of data generated by ITS have the potential for violating individuals' privacy. This becomes a more sensitive issue if the data, which now may disappear after immediate use by Traffic

Management operators, are permanently archived. Examples include video surveillance and probe vehicle data. In the case of video surveillance data, image processing can be used to extract key information (e.g., volumes, speeds, vehicle classifications) without concern for the identify of individual travelers. Probe data can either be summarized over time without regard to individual vehicles or a "pseudo" identification tag can be assigned prior to permanent storage. In no case should a cross-link between the "pseudo" tag and the actual identity be maintained.

3.2.8 Data Analysis

Although not strictly related to implementing the Archived Data User Service, data analysis is a major issue affecting its success. Operations, planning, maintenance, and administrative staffs may not have the resources to manipulate the data. Even if they did, there are no guidelines for how the data could be used (e.g., spot speeds from loop detectors). Lack of a common geographical referencing system also hinders analysis of the data. Data analysis can be facilitated to some degree by effective data management and access. For example, if the data are pre-screened for accuracy and summarized to levels commonly used for traditional analyses, these steps can be avoided by end users. Specifically, less sophisticated software can be used (e.g., conventional spreadsheets rather than expensive statistical packages) and working storage requirements for end users can be greatly reduced.

3.2.9 Coordination With Other Data Collection Efforts

Data generated by ITS can be used to supplement or replace existing data collection programs, such as statewide traffic monitoring, ISTEA-related management systems, and HPMS. The most effective structure for an archival system would include methods of integrating the various data collection efforts in use throughout stakeholder agencies. In this way, duplication of effort can be avoided and data programs will become more cost-effective. Careful review of what ITS can provide to existing programs is in order -- stakeholders must be aware that ITS may not have the level of geographic coverage required to meet all of their needs. For example, only instrumented highway segments can report traffic surveillance data; there is still a need to continue traffic monitoring on other segments.

3.2.10 Liability

Litigation has been a large issue to the transportation profession for many years. Archiving some forms of data generated by ITS may actually create problems for transportation agencies involved in lawsuits (e.g., video surveillance data). Public records statutes may force release of archived data.

3.2.11 Confidentiality of Privately Collected Data

Particularly within CVO applications, some data generated by ITS are collected by private companies for internal use. There may be reluctance to share these data on the grounds it might give a company's competitors confidential information on operations or that the government will take advantage of increased opportunities for regulation and taxation (e.g., weight-distance taxes). Therefore, any specification of data for the Archived Data User Service from private sources should conform to the "Fair Information Principles for ITS/CVO" now being developed by ITS America.

3.2.12 Incremental and Uncoordinated ITS Deployments

ITS deployment in metropolitan areas evolves as part of many separate projects over time, rather than being developed as part of a single, coordinated project. Multiple agencies are usually involved in ITS deployment and coordination with other existing or ongoing ITS projects is not guaranteed.

Therefore, the nature of the ITS deployment process will make it difficult to coordinate the archival of data for later use. For this reason, inclusion of the Archived Data User Service in the National ITS Architecture is crucial because of its importance in guiding local ITS development.

3.2.13 Retrofitting vs. New Development of Systems

Once ITS systems are deployed, it is more arduous to retrofit an archiving function rather than building it directly as part of initial system development. The fact that many ITS projects have and are being deployed without having the advantage of the National ITS Architecture means that retrofitting will have to be accommodated if the Archived Data User Service is to be realized. The process may be aided as ITS integration is promoted -- the infrastructure for achieving ITS integration (e.g., central servers, communication "backbones") can serve as a basis for implementing the Archived Data User Service.

3.2.14 Data Not Defined by the National ITS Architecture

Although the National ITS Architecture is the basis for the Archived Data User Service, ITS have evolved without the benefit of its guidance. Further, future ITS may include features not foreseen by the Architecture. Examples of data not currently defined by the National ITS Architecture that may be used for multiple applications include: door-to-door origin/destination patterns and tracking intermodal freight containers. Resolution of these matters is vital: either the National ITS Architecture must be amended to include these new data flows or it must be flexible enough to allow inclusion of them in field-implemented systems.

3.2.15 Conformance With Metric Conversion Standards

The National ITS Architecture does not now address metrication; units are specified in the English system. The degree to which metrication will influence the construction of data elements in the National ITS Architecture is still an open issue. Similarly, metrication must be resolved for the Archived Data User Service as well.

3.2.16 Training and Outreach

There will be a continuing need to train personnel not familiar with ITS in the fundamental concepts and methods used in ITS technologies. This will foster not only the development of the Archived Data User Service but the "main streaming" of ITS in general. In a similar fashion, training and outreach activities for ITS operators should be geared to understanding the needs and duties of other transportation agencies.

4. BASIC REQUIREMENTS

4.1 Introduction

The basic requirements for the Archived Data User Service are presented in this Chapter. These requirements are based on integrating the needs of the stakeholders with data that are either now available from ITS or could be easily achieved through ITS. A review of the general data types given in the National ITS Architecture -- as previously presented in Table 1.3 -- is the basis for identifying available data. This information, along with the background and issues presented in the earlier Chapters of this document, can be used to open detailed discussions with the stakeholder groups for the purpose of specifying the Archived Data User Service within the National ITS Architecture. The information is also of sufficient detail that it can also be used as a foundation for implementing the Archived Data User Service in real-world ITS applications. It must be noted, however, that *the recommendations regarding specific data elements identified in Section 4.2 are not meant to be exhaustive, merely a starting point for opening discussions with stakeholders.* In addition, several recommendations regarding the implementation and use of the Archived Data User Service by stakeholders are made in Section 4.3.

As previously stated, this User Service is fundamentally different from the 30 User Services currently defined by the National ITS Architecture in two key aspects. First, the number of stakeholder groups is very large, necessitating a high degree of coordination. Second, many stakeholder functions lie outside of the realm of ITS. For example, MPO transportation planners are concerned with a wide array of issues, with ITS only being a subset of those issues. It is not the intent of the National ITS Architecture to re-engineer these functions. Further, it is clear that many stakeholder needs can never be met with the existing structure of the National ITS Architecture or even with minor modifications to it. Rather, given that the functions already exist and that data from ITS can be used to assist those functions, the central question then becomes: *how can data from ITS be used to facilitate stakeholder functions.* The data needs of stakeholders have been well documented in past studies (as summarized in Appendix A). Stakeholder functional requirements, as they relate to available data from ITS, were given in Tables 1.3 and 2.1, although they have not been stated in formal system requirements terms. The reason for this is that the stakeholder requirements lead to specific recommendations about what data elements should be considered and what the structure of those elements should be (Table 4.1). In essence, many of the formal system requirements are imbedded in the recommendations concerning data elements. The next steps in the process from the point of view of the National ITS Architecture are to: (1) ensure that recommended data elements meet stakeholder needs by opening direct discussions with stakeholders and (2) work out the details of how the data specified in Table 4.1 can be accommodated with the National ITS Architecture.

4.2 Guiding Principles for the Requirements Definition

The following principles are general in nature and should serve as the basis for developing a more detailed set of requirements for the National ITS Architecture.

1. The systems to support the Archived Data User Service should take advantage of the data

flows inherent in ITS to provide information in usable form for stakeholder applications. In doing so, the systems should consider the needs of all the stakeholder groups identified in Table 1.1 and should provide for data generated by ITS to meet those needs wherever possible.

2. The systems to support the Archived Data User Service should be based on, but not limited to, existing data flows within the National ITS Architecture. As new data flows are added to the National ITS Architecture -- and as additional uses of existing data flows are identified -- they should be examined for their inclusion within the systems to support the Archived Data User Service. The systems should also be flexible enough to accommodate data flows unique to individual ITS deployments that may not warrant a change to the National ITS Architecture. They should also be capable of handling data from existing data collection programs that may not be deemed as being in the ITS realm.

3. To accommodate both existing ITS and the incremental deployment of new ITS, the systems to support the Archived Data User Service should encompass two types of development:

   a. Type 1: Decentralized Structure - Each subsystem possesses its own archiving function with a minimum of interconnects with other subsystems but with compatible data definitions.

   b. Type 2: Centralized Structure - Relevant data flows from each subsystem should be captured in a central repository either directly or "virtually" through the use of appropriate distributed technologies and standards.

   Type 2 should be the preferred level of development and should be consistent with all definitions and standards within the National ITS Architecture. Type 1 should be defined in such a manner that it can be applied to ITS deployments that are not based on the National ITS Architecture (e.g., pre-existing systems). There should be a clear migration path from Type 1 to Type 2.

4. Three distinct subsystems are defined to support Type 2 systems to support the Archived Data User Service development. These subsystems should be implemented as part of any new deployment that incorporates the principles of the Archived Data User Service as part of its system design:

   a. Data Processing - the receipt and processing of incoming data from other ITS subsystems. The processing includes data quality control and editing as well as aggregation.

   b. Data Storage - both online and offline storage of raw and processed data.

   c. Data Retrieval - the interface between the data repository and stakeholders.

   These functions should either be developed as separate National ITS Architecture subsystems or should be incorporated into a single subsystem. The functions are enumerated as part of other requirements, listed below.

5. The functions or subsystem(s) specified in {4} should be linked through data flows with the appropriate subsystems in the National ITS Architecture. As new subsystems are added to the National ITS Architecture, new data flows generated by them should be considered for inclusion in the systems to support the Archived Data User Service.

6. The functions of the existing Planning Subsystem should be integrated into the systems to support the Archived Data User Service. It is recommended that the Planning Subsystem be renamed so that it is more clear that its focus is the support of multiple uses of data.

7. The systems to support the Archived Data User Service should accept user-specified data quality control and editing procedures. Whenever these procedures are applied to specific data, a permanent record should be made of the results in the metadata portion of the data dictionary (see {8} below).

8. The systems to support the Archived Data User Service should have the ability to perform the following data processing functions:

   a. Store data in the same structure as received from the field.

   b. Accommodate levels of aggregation of the data flows, depending on the type of data represented.

   c. Sample raw data flows for permanent storage in accordance with user specifications. Permanent storage of the sampled data should be either online, offline, or both.

9. To facilitate use and access by stakeholders, the data dictionary and schema of the systems to support the Archived Data User Service should contain **metadata** on each data element including: basic data dictionary attributes, the source of the data (including make and model of equipment), the conditions under which the data were collected (e.g., weather), status of the data collection equipment and/or tests of sensor output (where appropriate), error flags assigned by field equipment, the type of editing/quality control used to process the data, any imputation used to fill in missing or erroneous data (for aggregated data items), and the results of the data editing/quality control (including error flags).

10. The raw data from incoming flows should be capable of being stored online for a period of time specified by individual system designs.

11. Data owners (i.e., personnel responsible for the equipment that collect data generated by ITS) should have the ability to review, edit, and flag data prior to them being permanently archived.

12. The systems to support the Archived Data User Service should be consistent and should work with all applicable standards and protocols for both ITS and non-ITS programs as specified by ITS America, American Association of State Highway and Transportation

Officials, Institute of Transportation Engineers (ITE), Society of Automotive Engineers (SAE), and USDOT, including:

a. Location Referencing Specification
b. NTCIP and TCIP
c. applicable electronic data interchange standards (e.g., ANSI's ASC X.12)
d. *Traffic Monitoring Guide*
e. Highway Performance Monitoring System
f. Fair Information and Privacy Principles for ITS/CVO

13. Permanent or temporary storage of data within the systems to support the Archived Data User Service should preclude the possibility of identifying or tracking individual citizens and should follow the ITS Privacy Principles developed by ITS America. Unique system-developed identifiers may be assigned to stored data that do not distinguish individuals. Public domain identifiers -- such as Social Security Numbers and license plate numbers -- should not be tagged or cross-linked with the stored data.

14. The systems to support the Archived Data User Service should accommodate, at a minimum, the data and associated structures shown in Table 4.1. The primary data elements shown in Table 4.1 should be matched to specific data flows in the National ITS Architecture. If the data flows do not currently exist in the National ITS Architecture, the National ITS Architecture should be modified (i.e., new data flows should be defined) to account for them.

15. The systems to support the Archived Data User Service should be capable of archiving *transformed* data as defined by local option. (See Section 1.2.1 for a definition.) Where transformed data are archived, the source data should also be archived for a period of time and keyed to the transformed data, along with metadata describing the calculation methods, assumptions, and external data used to perform the transformations.

16. The systems to support the Archived Data User Service should be compatible with and should take advantage of data specified the various ITS data dictionaries now being developed. These include: the Advanced **T**raffic **M**anagement Systems **D**ata **D**ictionary (TMDD) being produced by ITE, the Advanced Travelers Information Data Dictionary being prepared by SAE, and efforts by the Transportation Research Board Committee on Traffic Flow Theory and Characteristics to support data needs for advanced ITS modeling.

4.3 Specific Recommendations Regarding the Data Structure for the Archived Data User Service

To promote the adoption of the Archived Data User Service, a preliminary definition of the data structure is provided in Table 4.1. Either the primary data element or record type is given, often with a description of supplemental data fields that should be considered in the data structure. Time and location of the data collection are crucial companions to all the primary data elements shown and must be accounted for when systems are designed. (The issue of location referencing is beyond the scope of this document. However, there are several ongoing studies dealing with this issue that should be incorporated into final system designs.)

Table 4.1 Requirements for Archived Data from ITS for Multiple (Nonreal-Time) Uses

<table>
<tr><th>Primary Data Element or Record Type</th><th>Definition</th><th>Units</th><th>Internal Data Structure and Data Reduction Cycle</th><th>Level of Accuracy</th></tr>
<tr><td>Mainline traffic volume</td><td>Count of vehicles, during a given time period, past a point on the highway that is not influenced by a traffic control device or intersection. Includes volume counts from Electronic Toll Collection equipment.</td><td rowspan="3">Vehicles per unit time</td><td rowspan="10">Raw data from field sensors should be stored online for at least one day in the form in which they are received from the field. Data reduction should allow for permanent offline (or online) storage of the raw data. Data reduction should store the data online in multiple levels for each highway location where the data are collected:

1) 5-minute summaries by lane and direction -- online storage for at least one month
2) 1-hour summaries by direction -- online storage for at least one year
3) 24-hour summaries by direction -- permanent online storage

In performing the aggregation, locally-specified rules for handling missing or erroneous data should be applied. (The rules should be part of the metadata for these data elements.) Volumes in the aggregation are the simple totals for the time period; loop occupancy and density are simple averages; and speed should be the volume-weighted average.

For vehicle classification and weights, categories should conform to those in the most recent Traffic Monitoring Guide. Both classification and weight data will be stored for each highway location where the data are collected. Classification data should be stored as 1-hour summaries for each direction and lane and should be kept online for at least one year. Vehicle weight data should be accumulated online 24-hour intervals, at which time they should be summarized to the same structure as Level 3 of the Long-Term Pavement Performance Traffic Data Base (Appendix B). Permanent offline storage of both classification and weight data received directly from the field should be provided.

For vehicle headways, the type of lead and following vehicle should also be indicated.</td><td>+/-5%</td></tr>
<tr><td>Traffic control device approach volumes</td><td>Count of vehicles during a given time period on the approach to signalized intersections or at ramp meter controls.</td><td>+/-5%</td></tr>
<tr><td>Signalized intersection turning movements</td><td>Count of vehicles during a given time period for each turning movement at a signalized intersection.</td><td>+/-5%</td></tr>
<tr><td>Vehicle speed</td><td>Average speed of vehicles past a point on the highway during a given time period</td><td>Kilometers per hour</td><td>+/-5%; low speeds more important than high speeds</td></tr>
<tr><td>Loop occupancy</td><td>Average percent time that inductance loops sense vehicles during a given time period</td><td>Percent</td><td>+/-5%</td></tr>
<tr><td>Density</td><td>Average density of vehicles occupying a segment of highway during a given time period</td><td>Vehicles per lane-kilometer</td><td>+/-5%</td></tr>
<tr><td>Vehicle headway</td><td>The distance between two vehicles in the traffic stream, measured from the front bumper of the lead vehicle to the front bumper of the following vehicle.</td><td>Meters</td><td>+/-5%</td></tr>
<tr><td>Vehicle classification</td><td>Count of vehicles in pre-defined categories during a given time period</td><td>Vehicles per unit time by category</td><td>+/-10%</td></tr>
<tr><td>Vehicle weight</td><td>Weight of individual vehicles, axle groups for individual vehicles, or axles for individual vehicles</td><td>Kilograms</td><td>+/-10% for total vehicle weight</td></tr>
<tr><td>Traffic control device queue detection</td><td>Presence or absence of a queue located at significant setbacks from traffic control devices.</td><td>Yes/No</td><td>Time and setback from traffic control device should be stored on this record. Data should be stored at the level received from field controllers.</td><td>Accurate reading 90% of time</td></tr>
<tr><td>Traffic control device preemptions</td><td>Number of times traffic control devices (ramp meters and signals) have their timing preempted by transit, HOV, or emergency vehicles</td><td>Number</td><td>No special summarization required; save all data in raw form.</td><td>95-100% accuracy</td></tr>
</table>

Table 4.1 Requirements for Archived Data from ITS for Multiple (Nonreal-Time) Uses

Primary Data Element or Record Type	Definition	Units	Internal Data Structure and Data Reduction Cycle	Level of Accuracy
Traffic control device cycle lengths, phasing, and offsets	Time allocated for each phase (traffic signals only) and cycle (ramp meters and traffic signals). Offsets for traffic signals immediately upstream and downstream of the signal being inventoried.	Seconds	No special summarization required; save all data in raw form.	95-100% accuracy
Visual-based queue length	**Freeways:** Length of a platoon of vehicles where front-to-rear headways between vehicles are less than 25 feet, measured over a given time period. **Nonfreeways:** Length of a platoon of vehicles where front-to-rear headways between vehicles are less than 15 feet, measured over a given time period.	Kilometers	These data are determined from video or still-photography by either image processing or manual coding. Data should be permanently stored for each highway location at time intervals no smaller than 1-minute for signalized intersections and freeway ramps, and no smaller then 5-minutes for freeway mainline segments (requires computing average queue length for each time interval). The data should indicate the downstream point (beginning) and upstream point (end) of the queue using the local standard for location referencing.	+/-500 feet
Locally-derived traffic flow metrics generated by TMCs	Indices and other measurements used by local agencies to define congestion at either points on the highway (volume-to-capacity ratio) or highway segments (travel rate, accessibility index). These types of metrics are calculated from measured data (e.g., spot speeds used to calculate travel times).	Locally determined	Storage of speed and travel time data should follow the recommendations listed under the appropriate entries in this table. Locally-derived metrics should use these as a guide for their storage structure. Metadata must contain definitions and a thorough description of methods used in the calculations.	Unknown
Parking lot utilization	Proportion of parking spaces in use at a given time for a given locations	Percent	Stored data should contain not only the percent utilization but the total number of spaces available at the parking location being inventoried. Stored data should be summarized by 15-minute intervals for each parking location.	+/-10%
Transit vehicle boardings	Number of individuals paying transit fares upon entering a transit vehicle at specific times and locations (applies to both fixed-route and paratransit vehicles)	Number	Data should be permanently stored at the level that they are collected by electronic fare payment systems. Fields for identifying vehicle and route should be included in the structure.	+/-5%
Transit vehicle locations and times	The time that fixed-route transit vehicles arrive at scheduled stops and transfer points	Time	Data should be permanently stored at the level that they are collected by automatic vehicle location technologies. Fields for identifying vehicle and route should be included in the structure. Supplemental data should also include if an advisory for a route deviation was issued.	Unknown
Rideshare requests	The origin and destination of rideshare patrons by time of the request	Prevailing location referencing system	Data should be permanently stored by individual request.	95-100% accuracy

Table 4.1 Requirements for Archived Data from ITS for Multiple (Nonreal-Time) Uses

Primary Data Element or Record Type	Definition	Units	Internal Data Structure and Data Reduction Cycle	Level of Accuracy
Key times for incident specification	<u>Incident start</u> - time the incident occurred <u>Incident notification</u> - time the incident was reported from the field to a central operator <u>Incident verification</u> - time the incident was verified and recorded by a central operator <u>Incident dispatch</u> - time an EV was dispatched to the scene (multiples allowed) <u>Incident scene arrival</u> - time an EV arrived at the incident scene (multiples allowed) <u>Incident lane clearance</u> - time when each lane blocked was re-opened to traffic (for lane blockage incidents; multiples allowed) <u>Incident shoulder clear</u> - time when all shoulders were cleared of the blockage **and** EV <u>Incident return</u> - time each EV left the incident scene (multiples allowed)	Time	Data should be permanently stored for each incident reported from the field, whether they are verified or not. Times should be uniquely keyed to individual incidents. Supplemental data must include whether the times were actually measured/reported or estimated by system operators.	+/-5 minutes
Incident type	Category of incident	Formatted codes	Incident types should include at a minimum: (1) traffic crash, (2) debris (not water), (3) disabled/stalled vehicle (not crash-related), (4) fire on or adjacent to roadway (not related to a crash), (5) flooding or excess water on roadway (6) other weather-related (dust storm, tornado).	95-100% accuracy
Incident extent	Extent of traffic lane and shoulder blockage	Number	Number of lanes and/or shoulders blocked by the incident, including the times the blockage started and ended (to allow for multiple phase incidents).	
Incident hazardous material category	Hazard class and U.N. numbers (where appropriate) from the placard (multiple entries allowed)		Supplemental data on the incident; data should be permanently stored at the level that they are collected from the field.	
Incident hazardous material release	Amount of material released (multiple entries allowed)	Must be specified depending on container type		
Police accident report (PAR) reference	If the incident is a crash, the PAR reference number	Number		
Construction and work zone extent	Number of lanes and shoulders blocked by the construction or work zone activity	Number	Beginning and ending times **and** locations of each activity also need to be specified. Supplemental data could include a description of the activity.	95-100% accuracy
Train arrivals at HRIs	The beginning and ending times that HRIs are blocked by trains	Time	Data should be permanently stored at the level that they are collected from the field.	95-100% accuracy

Table 4.1 Requirements for Archived Data from ITS for Multiple (Nonreal-Time) Uses

Primary Data Element or Record Type	Definition	Units	Internal Data Structure and Data Reduction Cycle	Level of Accuracy
Emergency vehicle dispatch times	EV dispatch - time an EV was dispatched to the scene EV scene arrival - time an EV arrived at the scene EV clear - time emergency personnel reported the case "cleared" EV leave - time an EV left the scene EV destination arrival - time an EV arrived at its return location (e.g., hospital for medical service)	Time	These data are relevant for **all** emergencies, not just incidents. Data are usually collected by individual agencies through computer-aided dispatch systems.	+/-5 minutes
Emergency vehicle locations during response	EV origin - location of the EV when it was dispatched EV destination - location of the case or incident EV intermediate location - location of the EV at selected time intervals between origin and destination	Prevailing location referencing system	These data are used to track the routes taken by EVs in responding to cases or incidents. Intermediate locations should be recorded at 1-minute intervals between the times the vehicle was dispatched and it arrives at the scene. Because of the volume of the data generated, permanent online storage is optional if the data are stored offline.	Unknown
Commercial vehicle cargo type	The SIC code for the type of cargo being transported	SIC code	These data are collected by CVO systems, usually field sensors that detect the passage of individual trucks. The data should include time, location, and a vehicle identification code. Archiving data from every truck would probably not be cost-effective; however, provision to permanently store a sample of data should be made.	90-95% accuracy
Commercial vehicle origin and destination	For the shipment being made by this vehicle, the first point of origin and last destination	Prevailing location referencing system		Unknown
Intermodal container cargo type	The SIC code for the type of cargo being transported and the type of container.	SIC code		90-95% accuracy
Commercial vehicle origin and destination	The first point of origin and last destination for the container.	Prevailing location referencing system	Same as for commercial vehicle cargo and O/D.	Unknown
Hazardous material cargo type	Hazard class and U.N. numbers (where appropriate) from the placard (multiple entries allowed)			95-100% accuracy
Hazardous material pre-planned shipment route	The specified route to be taken for hazardous material shipments that require such treatment	Highway routes (as determined by the issuing agency)		Unknown

Table 4.1 Requirements for Archived Data from ITS for Multiple (Nonreal-Time) Uses

Primary Data Element or Record Type	Definition	Units	Internal Data Structure and Data Reduction Cycle	Level of Accuracy
Commercial vehicle driver log	Selected locations and dates/times to determine hours of service for drivers	Prevailing location referencing system	These data are collected from on-board safety systems that are downloaded to field sensors. Archiving data from every truck would probably not be cost-effective; however, provision to permanently store a sample of data should be made. Supplemental data include vehicle identification and cumulative vehicle mileage. Privacy concerns may preclude the collection of these data.	95-100% accuracy
Commercial vehicle subsystem status	Type of subsystem and status of its operation	(N/A)		Unknown
Roadside emission concentration	Volumetric concentration of pollutants measured by roadside sensors	Grams per unit volume (HC, CO, NO_x, SO_x)	Data should be saved for a minimum of one day online in the form that they are received from the field. These data should be aggregated and permanently stored for 15-minute time intervals (average concentrations for 15-minutes).	Unknown
Roadside temperature	Air temperature	Degrees Celsius	These data are collected by roadside weather sensing equipment. Data should be aggregated to no longer than 15-minute summaries (total precipitation, average temperature, predominant wind direction, average speed of wind in predominant direction) for permanent storage.	Unknown
Roadside precipitation	Type and amount of precipitation	Cubic centimeters (liquid volume)		Unknown
Roadside light conditions	Light level at roadside	At a minimum, should follow the codes specified in the Fatal Accident Reporting System		Unknown
Roadway surface condition	The surface condition of the roadway in terms of amount and type of moisture			Unknown
Roadside wind conditions	Direction and speed of wind	Kilometers per hour (speed)		Unknown
Segment travel times from probe vehicles	The time for a probe vehicle to traverse a given roadway segment	Seconds	For permanent storage, probe information (times at given points on the highway system) should be converted to total seconds. The data should be permanently stored online as 5-minute summaries (total probes counted, average travel time). A supplemental data item for permanent storage is the segment length. The raw probe data may be stored offline if actual vehicle identification is not included.	+/-10%
Transit vehicle times and locations	Data from AVI- or GPS-equipped transit vehicles	Seconds; prevailing location referencing system	For permanent storage, arrival times at pre-determined stops should be recorded along with vehicle and route identification. If transit vehicles are used as general travel time probes, they should be included as a special category under "Segment travel times from probe vehicles".	+/-10%
Traveler message content	Actual text of message displayed on a VMS or to travelers via personal devices.	(N/A)	All VMS messages (along with time and location) should be permanently stored online. Messages to personal devices (e.g., in-vehicle signing) may be sampled prior to permanent storage.	Close to 100% accuracy

Primary Data Element or Record Type	Definition	Units	Internal Data Structure and Data Reduction Cycle	Level of Accuracy
Vehicle trajectories	Time and location of individual vehicles; measured for very short time intervals (1-10 seconds)	Prevailing location referencing system	Vehicle trajectory data can be collected through either GPS or advanced video image processing. The type of vehicle should also be indicated (see scheme under "vehicle classification"). Permanent online storage of both vehicle trajectory and origin/destination data are not recommended because of the sheer volume of data. Therefore, data should be aggregated to locally-defined geographic zones (e.g., traffic analysis zones, Census block groups) by activities. However, data may be accumulated online for short periods of time and stored offline for future use. Origin and destination activities (GPS-collected data) may be either collected directly or inferred from GIS base information (in advanced deployments.)	Unknown
Traveler origins and destinations	Origin - the point at which the trip began Destination - the point at which the trip ended Origin/Destination Activity - type of activity engaged in by the traveler at the origin and destination	Prevailing location referencing system		Unknown
Route guidance	Starting Point - the location of the trip at the time the guidance was given Ending Point - the desired destination Recommended Route - the route segments recommended	Prevailing location referencing system		Unknown
Variable facility pricing	Amount charged for a parking facility or toll for a highway segment where congestion pricing is in effect	Dollars	All recorded changes to pricing should be permanently stored.	Unknown

Notes: *This is a general indication of the desired accuracy.
(1) In addition, metadata in accordance with the principles put forth in Section 4.2 must be specified for each data element.
(2) time and location referencing must be considered in constructing systems around these data elements.

ABBREVIATIONS
HRI: Highway Rail Intersection
EV: Emergency Vehicle
TMC: Traffic Management Center
VMS: Variable Message Sign

Unless otherwise specified, the data specified in Table 4.1 should be saved online in their recommended forms for a minimum of one year, at which time they should be copied to offline storage media. The data structure is based on the notion that some data will be more useful to a wide variety of users if it is aggregated (summarized) and/or sampled prior to permanent storage. The aggregations described are simple totals or averages of lower data levels. Transformations of the data -- such as converting speeds to travel times -- are not specified for the basic structure of the Archived Data User Service. However, this does not preclude state and local agencies from incorporating data transformations in their implementations of the Archived Data User Service.

The information in Table 4.1 can be easily converted to User Service Requirements as developed for the other User Services in the National ITS Architecture; an example is provided in Table 4.2.

Before this is done, however, it is important to reach consensus on the full list of data elements and their attributes through stakeholder discussions. The structure shown in Table 4.1 is thought to be more effective for stakeholder interaction than the traditional requirements format.

Table 4.2. Example Derivation of User Service Requirements

1.0	ROADWAY SURVEILLANCE DATA
1.1	MAINLINE TRAFFIC VOLUME
1.1.0	ADUS shall provide data on the count of vehicles, during a given time period, past a point on the highway that is not influenced by a traffic control device or intersection. This shall include volume counts from Electronic Toll Collection equipment.
1.1.1	Raw data from field sensors shall be stored online for at least one day in the form in which they are received from the field.
1.1.2	Data reduction shall allow for permanent offline (or online) storage of the raw data.
1.1.3	Data reduction should store the data online in multiple levels for each highway location where the data are collected.
1.1.3.1	5-minute summaries of total volume shall be provided by ADUS
1.1.3.1.1	5-minute summaries shall be for each lane and direction of traffic
1.1.3.1.2	5-minute summaries shall be stored online for at least one month

The structure shown in Table 4.1 can also be used as a basis for constructing a detailed data dictionary and schema for the National ITS Architecture. It is recommended that the entire structure be incorporated into the systems developed around the Archived Data User Service. The structure can also be used as guide for system developers who wish to retrofit data archiving on top of legacy systems, but development cost may prohibit full implementation. If this is the case, developers may wish to prioritize the retrofitting in accordance with the guidance provided below. (These general priorities are offered for guidance only. Local conditions may dictate otherwise.)

High Priority
- Traffic surveillance data (volumes, speeds, densities, and loop occupancies)
- Truck monitoring data (classifications and weights)
- Probe vehicle data
- Incident-related data
- Transit vehicle boardings

Medium Priority
- Transit vehicle locations and times
- Rideshare requests
- Construction and work zone extent
- Commercial vehicle cargo type and origin/destination data

Low Priority

- Traffic control device cycles, phasing, offsets, and preemptions
- Queue length
- Locally-derived traffic flow metrics
- Parking lot utilization
- Train arrivals at HRIs
- Emergency vehicle dispatch times and locations
- Pre-planned routes for hazardous material shipments
- Commercial vehicle driver log
- Commercial vehicle subsystem status
- Roadside emission concentration
- Roadside weather conditions
- VMS content
- Vehicle trajectories
- Route guidance
- Variable pricing

It is important to recognize that stakeholders must be free to develop their own priorities depending on local conditions. Specification of the system architecture is for guidance only.

4.4 Applicable Technologies

The Archived Data User Service is essentially an information management system, therefore, several information technologies (IT) apply to its development. System designers should make full use of these technologies when implementing the Archived Data User Service. These include (1) relational and distributed data base design; (2) data warehousing; (3) data mining; (4) expert systems, and (5) geographic information systems (GIS). Relational and distributed design and data warehousing are IT concepts that aid in the management and retrieval of data. Data mining can be applied to the analysis of the data to identify trends that otherwise might be buried in the huge volume of information that exist. Expert systems can be used as an aid in data quality control. GIS is an obvious technology for managing and displaying data generated by ITS.

4.5 Other Recommendations for Implementation

As stated many times in this document, simply defining the data that should be maintained in the Archived Data User Service is insufficient to achieve successful implementation. Therefore, several other recommendations are offered that will promote development and use of the Archived Data User Service in the field.

- "Best Practice" procedures for performing quality control and editing on Archived Data User Service data should be developed. Because ITS operations are a new phenomenon, little is known about how to identify and adjust questionable data received from field equipment. Data quality is a particular concern for those data elements that are aggregations of raw data: what should be the "rules" for handling not only questionable but missing data in the aggregation process. An ongoing research effort jointly funded by several states is examining QC and editing procedures for vehicle classification and weight data.

Consideration should be given to performing similar studies for not only other forms of traffic surveillance data (e.g., volumes, speeds, densities) but for other types of Archived Data User Service data as well. Once the procedures are established, there is an additional need to develop *automated tools* to facilitate quality control and editing.

- Similarly, "Best Practice" procedures for analysis of archived data generated by ITS would provide guidance to stakeholders and would promote the use of Archived Data User Service. Demonstration of analytic methods (including graphical displays) would be extremely valuable to stakeholder groups. This is especially important because the sheer volume of data may be daunting to some groups not accustomed to working with large data sets. Assistance could take many forms: providing software and sample data, case studies of how other stakeholders use data, or documenting analysis techniques for specific applications (e.g., congestion monitoring, bus route planning, TDF model input preparation).

- Coordination with ongoing data dictionary efforts is crucial to the future development of the Archived Data User Service. There are several ongoing efforts to develop data dictionaries for various subcomponents of ITS (e.g., TMDD). Because these efforts are specifying data structures, they are highly relevant for the Archived Data User Service. Therefore, at a minimum, the stakeholders identified here should have input to these efforts. Further, consideration should be given to developing an Archived ITS Data Dictionary in accordance with the guidance provided by this document and the input of stakeholders. The endeavor of creating a data dictionary will force stakeholders and system architects to resolve many of the technical issues raised here.

- The needs of stakeholder groups should be represented in all forms of standards development that may have an impact on the relevant data. Other efforts have the potential for influencing the nature of data used in ITS technologies and are relevant to the Archived Data User Service. Therefore, input from Archived Data User Service stakeholders to these efforts will ensure that data are compatible.

- The Archived Data User Service should be integrated into other Federal, state, and local data collection programs. Although the Archived Data User Service is only one source of data for stakeholders, it can be used to supplement or replace many existing data programs. Examples include submittals of certain data to HPMS and statewide traffic monitoring. Full consideration should be given to how the Archived Data User Service fits into a comprehensive data collection program, including data sharing and standards.

- Training the various stakeholders in each others' needs is seen as an ongoing requirement. Personnel not directly involved in ITS operations currently have a limited working knowledge of the National ITS Architecture and of ITS in general. Likewise, many personnel who come from a systems engineering background rather than a transportation background do not yet have an appreciation for the breadth of traditional transportation functions. Education and outreach activities need to be increased for all transportation professionals. Several immediate options are available including: additional training under the Professional Capacity Building effort, strong recruitment of all stakeholders for the upcoming Regional Architecture Workshops, and development of a special short course for transportation managers.

- A concentrated field effort to demonstrate the implementation and use of the Archived Data User Service should be undertaken. Similar to the concept of ITS Field Operational Tests, the Archived Data User Service demonstration would provide a model for how to perform system development as well as how the data may be put to use. A key part of the demonstration would be to document the value of the increased information provided by the Archived Data User Service to local decision making and operations. Full documentation on the institutional issues as well as the technical hurdles to developing such a system must also be addressed.

- Funding for the Archived Data User Service should be implicit in future ITS deployment. Once the Archived Data User Service is specified in the National ITS Architecture, special emphasis should be placed on it when Federal ITS funding is offered to state and local governments. The cost of implementing the Archived Data User Service should be included when budget levels are determined.

APPENDIX A

SUMMARY OF RECENT DATA NEEDS IDENTIFICATION EFFORTS

A1. IRVINE, CALIFORNIA CONFERENCE

A conference titled Information Needs to Support State and Local Transportation Decision Making into the 21st Century was held in Irvine, California, March 2-5, 1997. The conference was sponsored by several agencies — Transportation Research Board, Bureau of Transportation Statistics, Federal Highway Administration, Federal Transit Administration, American Association of State Highway and Transportation Officials, and Association of Metropolitan Planning Organizations. The proceedings of the conference was published by the Transportation Research Board in 1997.[1] The participants of the conference represented a variety of transportation-related organizations including state departments of transportation, metropolitan planning organizations, federal agencies, universities, and private consulting companies. The discussions and deliberations that took place at the conference covered a variety of issues related to transportation-related data including:

1. The types of data needed for planning and policy analysis;

2. Data collection requirements and methods;

3. Current data collection programs and institutional arrangements;

4. Improvements needed in data; and

5. Future trends.

The findings with regard to data needs are presented in the proceedings in an organized manner under different categories and subcategories. These findings are presented below in an itemized manner using the groupings and categories used in the report.

I. Socioeconomic Data
   1. Demographics
      - Emerging and critical population subgroups
      - Household characteristics of nonpermanent residents
      - Vehicle ownership and availability
   2. Economics
      - Changes in patterns of building and development
      - Tax data that reflect economic activity
      - Military base abandonments and conversions
      - Housing market data
      - Tourists and visitors
      - Business establishments

---

[1] *Information Needs to Support State and Local Transportation Decision Making into the 21st Century*, Proceedings of a Conference, Irvine, California, March 2-5, 1997. National Academy Press, Washington, D.C., 1997.

- Transportation investment to support business locations
- Employment and worker characteristics
- Geocoding of employment
- Characteristics of multi-job holders
- Labor force
- Transportation access to employment opportunities
- ES202 employment data
- Shifts in population and employment
- State-level employment and labor

3. Land Use
   - Tax assessment data
   - Local land use policies
   - Economic development plans
   - Land use ratios per capita
   - Suitability of vacant land
   - Values of land overlaid against transportation improvements

II. Financial Data

1. Revenue Forecasting
   - Truck registration
   - Federal and state sales taxes
   - Sales and fuel tax revenues
   - Expenditures of fuel taxes by businesses and households
   - Long-range substate-level financial projections
   - Financial impact of transportation system pricing
   - Finance streams to support maintenance
   - Cash-flow prediction at all levels of government
   - Returns on different investment strategies
   - Toll charges and revenues
   - Carrier revenues
   - Data for toll facility analysis
2. Alternative Financing
   - Statutory limits on creative financing
   - Public-private partnerships
   - Administrative costs associated with privatization
3. Cost and assets
   - Capital and operating costs of transportation systems
   - Costs of programs and services
   - Infrastructure costs
   - Life-cycle costs
   - Costs of rights-of-way, construction, and raw materials
   - Capital costs
   - Expenditures by levels of government for each highway functional class
   - Expenditures by revenue source
   - Letting costs
   - Highway and other transportation cost allocation

III. Supply and System Characteristics Data
   1. Networks and Facilities
      - Geographic detail about the location and connectivity of transportation infrastructure
      - Capacity and speed of network links and other facilities
      - Operating restrictions by time of day
      - Tolls and other facility-specific charges
      - Functional class of each highway segment
      - Condition measures for pavement, bridge, and other physical infrastructure
      - Materials used in construction and maintenance
      - Jurisdiction of each agency responsible for operation and maintenance of facilities
   2. Transportation Service
      - Routes and schedules of fixed-route bus and rail services
      - Area and time coverage of paratransit services
      - Service operated in compliance with the Americans with Disabilities Act
   3. Intelligent Transportation System Infrastructure
      - Facility-specific control systems
      - Communication networks
   4. Linkages Between Transportation Systems and Areal Data
      - Data on other systems in surrounding area
   5. Vehicle Fleets — Private and Public
      - Characteristics and size of vehicle fleets of different modes
      - Geographic distribution of vehicle fleets

IV. Demand and Use Data
   1. User Behavior and Characteristics
      - Activity by location
      - Trip generation by age
      - Journey to work by mode and trip length
      - Latent demand and induced demand
      - Time of day of travel
      - Use of time and deferral of trips
      - Trip chaining
      - Work schedule changes, telecommuting and teleshoping
      - Actual versus theoretical trip routing
      - External travel distribution
      - Local travel of out-of-town visitors
      - Travel of people who move to warm climates in winter
      - Customer satisfaction
      - Network trips
      - Transportation disadvantaged
      - Changing trip generation characteristics
      - Social characteristics of users by mode
      - Demand characteristics of the elderly
      - Usage characteristics by mode

2. Freight
    - Freight movement by mode: Inter- and intracity
    - Commodity flow
    - Freight demand by time
    - Origin-destination of international freight
    - Border crossing vehicles and their travel pattern
    - Port data
    - Cargo in transit through the United States
3. System Use
    - Vehicle types on different classes of roads
    - Temporal and origin-destination patterns of passengers and goods movement
    - Highway traffic counts and transit ridership
    - Personal trips by alternative modes
    - Occupancy of different types of vehicles
    - Vehicle miles traveled (VMT)
    - Travel time
    - Trip productions and attractions
    - Airport and ground access demand
    - Impact of urban design on travel behavior
    - Traffic impact of special events

V. System Operation Data
1. System performance
    - Volume versus capacity by time of day
    - Travelers violating transit fare policies
    - Violations of high-occupancy vehicle restrictions
2. Reliability and Congestion
    - Travel time variations by mode and time
    - Characteristics of recurring and nonrecurring congestion
    - Economic impacts of congestion
    - Effects of human factors on traffic flow
    - Systemwide versus localized congestion management measures
    - Vehicle occupancy by hour
    - Responses to scheduled disruptions, special events, and incidents
3. Freight Operations
    - Turn around times at marine and air terminals
    - Factors affecting route choice
    - Delivery by trucks
    - Weight-in-motion systems and truck safety
    - Effect of just-in-time delivery requirements

VI. Impact and Performance Data
1. Performance Measures
    - Travel times and speeds that the customer considers effective for all modes compared with perceived and actual times and speeds
    - Acceptable delay (for users) by time of day

- User costs
- Value of user's time
- User sensitivity to toll charges
- Quality of trips from customer's perception
- User mode preferences
- User versus nonuser benefits
- Customers' views about traffic calming strategies

2. Communities, Safety, and the Environment
   - Impact of different socioeconomic groups
   - Quality of life
   - Sustainability
   - Social impact of bicycling and walking
   - Negative impacts of transportation improvements
   - Accident rates by mode
   - Accident rates for auto passengers
   - Perceptions of safety and mode choice
   - Correlation of delay and accidents
   - System elements not harmful to the environment
   - Emissions data by mode
   - Acceleration/deceleration of modes
   - Transit ridership
   - User responses to Travel Demand Management and congestion pricing programs
   - System conduciveness to use of alternatives modes — bicycling, walking, ferries
   - Environmental justice
   - Important habitats
3. Economic Impact
   - Economic impacts of congestion, especially on freight movement
   - Impact on personal and business incomes
   - Economic impacts of transportation programs for reduction of energy consumption and environmental degradation
4. Freight Movements
   - Access to ports
   - Processing and permitting
   - Cross-border inspections

A2. NCHRP REPORT 401

A research project sponsored by the National Cooperative Highway Research Program (NCHRP) titled Multimodal Transportation Planning Data, examined the availability of data needed to support statewide and metropolitan multimodal transportation planning along with a few other data related issues. One of the products of this report is NCHRP Report 401, *Guidance Manual for Managing Transportation Planning Data.*[2]

One of the contributions of the report is the presentation of a hierarchical classification system of various data items. The four major components of data identified in the report are:

1. Supply attributes;
2. Demand attributes;
3. System performance; and
4. System impacts.

Further stratifications of data and examples of individual data items for each of these components are presented below:

I. Supply Attributes of Each Mode — Highway, Rail, Transit Systems, Ports and Inland Waterways, and Airports
   1. Systems Data
      - Guideways, routes, terminals, runways, etc.
      - Capacity
      - Land use for facility expansion
   2. Service Data
      - Access — connections to other modes
      - Intermodal access
      - Areas or cities served
      - Frequency of service
      - Service providers
      - Fare of fee structure
      - Drayage services
   3. Facilities Data
      - Inventory of stops/stations, rest areas, garages, etc.
      - Loading/unloading facilities
      - Vehicles
      - Berth facilities for boats and ships
      - Cargo storage facilities

---

[2]Jack Faucett Associates, COMSIS Corp, and Mid-Ohio Regional planning Commission, *Guidance Manual for Managing Transportation Planning Data*, National Cooperative Highway Research Program Report 401. National Academy Press, Washington, D.C., 1997.

4. Condition Data
    - Pavement ratings
    - Bridge structures
    - Tunnel clearance
    - Vehicle age
    - Navigation aids
    - Channel depth and width
    - Lock condition
5. Project Data
    - State projects (proposed)
    - MPO projects (proposed)
    - Major investment data
    - Project evaluation data
    - Project maintenance data
    - Planned expansions and modifications
    - Project history data

II. Demand Attributes

1. Economic Data
    - Income
    - Employment
    - Vehicle ownership
    - Travel cost
    - Locations of industries, wholesalers, etc.
    - Commodity production and consumption data
    - Export/import data by pint of exit/entry
2. Demographic Data
    - Population and labor force
    - Housing characteristics
3. Land Use Data
    - Acreage
    - Housing
    - Employment
    - Zoning
    - Access data
4. Commodity Flow Data
    - Origin-destination
    - Modal split
    - Factors affecting modal split — modal rates, delivery time, etc.
5. Travel Data
    - Trip generation
    - Trip distribution
    - Special generators
    - Traffic volumes by time of day
    - VMT data by vehicle type, road type, time of day, etc.
    - Shipper modal selection factors

6. Travel Behavior Data
    - Mode choice
    - Route choice
    - User preference
    - Time of day of pickup and delivery
    - Carrier behavior

III. System Performance Attributes
1. Safety Data
    - Incidents and accidents
    - Security
    - Medical services
2. Performance Measures
    - Performance of each mode — highway, transit, etc.
    - Intermodal system performance
    - Efficiency data by mode
    - User cost data
    - Delivery times by mode/intermodal
    - cargo damage
    - Terminal congestion
    - Shipment costs

IV. System Impact Attributes
1. Air Quality Data
    - Vehicle registration by vehicle class, fuel type, etc.
    - VMT data by road class, time of day, etc.
    - Speed data by road class, vehicle class, time of day, geographic area, etc.
    - Trip data — cold versus hot starts, etc.
    - Emissions contributions by vehicle class, etc.
    - TCM effectiveness estimates
2. Other Environmental Data
    - Visual and aesthetic impacts
    - Noise and vibration impacts
    - Ecosystems
    - Archaeological and cultural impacts
    - Parklands
3. Land Use Data
    - Socioeconomic impacts
    - Neighborhood impacts
4. Energy Data
    - Energy consumption impacts by mode
    - Energy efficiency impacts by mode
    - Energy price impacts
5. Economic Growth Data
    - Local employment impacts
    - Regional employment impacts
    - Access to natural resources

- Access to domestic markets
- Access to ports and foreign markets

A3. HPMS REASSESSMENT WORKSHOP

A national workshop was held in Minneapolis, Minnesota, in July 1997, to examine various aspects and issues involving the Highway performance Monitoring System (HPMS). The participants represented a variety organizations — transportation agencies at the federal, state, and local levels; universities; research organizations; transportation consulting firms; etc. The three major themes of the workshop were:

1. HPMS mission, goals, and objectives;

2. Improving intergovernmental partnerships and data sharing; and

3. Future scope and scale of HPMS.

The discussions and deliberations of the workshop were organized under these three themes. The first two themes did not examine individual data items in detail. The third theme's scope included some individual data items although it also included many other issues of HPMS. The report titled *HPMS Reassessment Workshop/Steering Committee Meeting — Summary*,[3] captures the highlights of discussions at various general sessions and breakout groups. The report is published by the Federal Highway Administration, of the U.S. Department of Transportation. The major categories of data and some of the individual data items mentioned in this report are presented below.

The major categories of data that were mentioned included:

1. Vehicle miles traveled (VMT);

2. Pavement condition;

3. Congestion data; and

4. Safety data.

Several issues involving the use of HPMS generated VMT data for meeting the requirements of the Clean Air act Amendments of 1990 were discussed at the workshop. The United States Environmental Policy Agency (USEPA) and the Federal Highway Administration (FHWA) have an agreement regarding the appropriate uses of HPMS VMT for air quality analyses. In some cases additional data, such as local traffic counts outside of HPMS sample counts may be used with the consent of FHWA and USEPA. VMT growth rates and the spatial and temporal allocation of VMT may be obtained from other methods such as travel demand models.

---

[3] *HPMS Reassessment Workshop/Steering Committee Meeting — Summary*, Publication No. FHWA-PL-98-012. Office of Highway information Management, Federal Highway Administration, U.S. Department of Transportation.

For pavement condition data, there are opportunities for improving current procedures used. The protocols develop by the American Association of State Highway and Transportation Officials may be useful for this purpose.

With regard to congestion data, it was mentioned that FHWA is moving away from using volume/capacity ratio as a measure of congestion. Alternative measures that are being considered include:

1. Travel time;

2. Travel speed; and

3. Delay.

The above listed data usually are available from an Intelligent Transportation System (ITS) on a real time basis. However, currently there are no protocols for retaining, managing, and sharing data generated by ITS. Procedures to incorporate ITS real time data into HPMS static data base should be examined/investigated.

There are different data bases for safety related data that are maintained by various agencies, including the National Highway Traffic Safety Administration. FHWA may be examining opportunities for reporting existing safety data through the HPMS.

A.4. "ITS AS A DATA RESOURCE" WORKSHOP

ITS America, in association with the Federal Highway Administration, Federal Transit Administration, and the Association of Metropolitan Planning Organizations, sponsored a Workshop entitled: *ITS As a Data Resource* on January 9 and 10, 1998. The Workshop identified opportunities for tapping data produced by Intelligent Transportation Systems for use in:

- Transportation Planning Applications (covering both highway and transit).
- Transportation Operations (including state and local traffic engineering).
- Commercial Vehicle and Intermodal Freight Planning.

The objectives of the Workshop were to:

- Bring transportation planners and operators together with representatives of the ITS community to discuss common data needs and concerns.
- Identify currently available ITS-generated data that can meet the data needs of transportation planners and operators.
- Identify opportunities for expanding ITS-generated data collection to meet additional data needs of planners and operators.

Workshop participants dealt with a variety of issues including:

- The nature of ITS-generated data that are now available, with particular emphasis on the data structures in the ITS National ITS Architecture.
- A discussion of the general data needs of transportation planners and operators.
- Matching available ITS-generated data with data needs.
- Discussing how additional data needs can be met by expanding ITS-generated data services.
- Levels of summarization that would have to be made to raw ITS-generated data to meet data needs.
- Computer resource requirements.

As preparatory information for the participants, three White Papers were developed: (1) *ITS Data for Freight Planning*, (2) *Using ITS-derived Data for Transportation Planning, Programming, and Operations*, and (3) *Use of ITS Data for Transit Planning*. In the first and third papers, a list of data needs were developed; these appear in the following tables.

Table A.1. Public Sector Freight Planning Data Needs

Function	Data Needs	Planning Application
Congestion management	Truck-hours of travel Average speed or travel rate (hours per kilometer) for truck Added truck-hours or truck-hours per kilometer due to congestion Truck transport cost (total, or per truck-kilometer, metric ton-kilometer, or dollar value of freight carried) Added cost due to congestion Transport time reliability Types of trucks and commodities caught in congestion Energy consumption for trucks: total or per truck-kilometer or metric ton-kilometer Emissions rates for trucks: total or per truck-kilometer or metric ton-kilometer	Understand impact of congestion on goods movement Understand contribution of trucks on urban congestion and air quality problems
Intermodal access	Volumes of trucks entering or exiting an intermodal facility Congestion-related delays on access roads to the facility Queuing counts related to the capacity of the facility Accident rates on access roads to the facility Travel time contours around the facility (e.g., driving distance within 30 minutes) Number of people living or working within x kilometers of the facility	Identify landside access improvement needs
Truck route designation and maintenance	Truck traffic volumes Origin/destination patterns Truck size and weight data	Identify high-volume truck routes and corridors Assess pavement damage and replacement needs
Safety mitigation	Accident rates Rail-grade crossings Low-clearance bridges Steep grades	Identify safety hazards and develop mitigation strategies
Economic development	Truck volumes Commodity movements Origin/destination patterns Shipping costs	Assess economic benefits and costs of freight transportation investment projects

Table A.2. Data Needs of Transit Operators

Transit Planning Function	Data Elements/Messages
Corridor Analysis planning	strategic transportation planning data (see Strategic/Business planning, below)
	travelers per vehicle data
	vehicle emissions data
	emissions impact data
	multimodal construction cost data
	multimodal construction cost projections
	multimodal operating cost data
	multimodal operating cost projections
	multimodal replacement/renewal cost data
	multimodal replacement/renewal cost projections
	multimodal performance data (reliability, performance, variation
	multimodal performance projections
	multimodal travel/shipping cost data
	multimodal travel/shipping cost projections
	multimodal mobility impact data
	multimodal mobility impact projections
	multimodal land development impact data
	multimodal land development impact projections
	multimodal economic impact data
	multimodal economic impact projections
	multimodal customer satisfaction projections
Maintenance planning	system reliability data
	time-to-repair data
Market Research	strategic transportation planning data
	Corridor analysis planning data
	demographic data
	customer expectation data
	customer expectation projections
	highway performance data (O-D travel time, reliability)
	highway performance projections
	highway travel cost data
	highway travel cost projections
	airline performance data
	airline performance projections
	airline travel cost data
	airline travel cost projections
	competing transit carrier performance data
	competing transit carrier performance projections
	competing transit carrier travel cost data
	competing transit carrier travel cost projections
	transit performance data
	transit performance projections
	transit travel cost data
	fare elasticity data
	fare elasticity projections
	transit travel cost projections

Table A.2. Data Needs of Transit Operators

Transit Planning Function	Data Elements/Messages
New Project planning (including Major Investment Studies and environmental impact reports)	zoning data
	hardware/software maintenance data
	hardware/software reliability data
	construction cost data
	construction cost projections
	operating cost data
	operating cost projections
	replacement/renewal cost data
	replacement/renewal cost projections
Operations planning (including day-to-day scheduling and headways)	Ridership
	Accident data
	Claims data
	Maps
	Weather
	Demographics
	Passenger Travel Times
	Running Time
	Signal Timing
	Roadway Conditions
	Service Disruptions
Research and development planning	Time-to-market data
	Implementation time variance data (planned versus actual)
	capital cost variance data (planned versus actual)
	operating cost variance data (planned versus actual)
	performance variance data (planned versus actual)
	reliability variance data (planned versus actual)
	usage/utility variance data (planned versus actual)
Service planning (including route locations, scheduling and runcutting)	strategic transportation planning data
	corridor analysis planning data (above)
	transit market research/business planning data (above)
	crime statistics
	modal performance data
	modal performance projections
	modal operating cost data
	modal operating cost projections
	highway performance data
	highway performance projections

Table A.2. Data Needs of Transit Operators

Transit Planning Function	**Data Elements/Messages**
Strategic/Business planning	map data
	current demographics
	journey-to-work data
	journey-to-school data
	movement-of-goods data
	recreational travel data
	current travel demand
	current travel constraints
	land development strategy
	mobility management strategy
	zoning data
	forecast demographics
	forecast travel constraints

APPENDIX B

SUPPLEMENTAL INFORMATION ON THE LONG-TERM PAVEMENT PERFORMANCE TRAFFIC DATA STRUCTURE

Overview of Data Types in the LTPP Central Traffic Data Base (CTDB)

The LTPP-CTDB data are classified into five **levels** as summarized in Table B.1. Note that several subtypes of data (volume, classification, and weight) are available for Levels 3 and 4. Therefore, the actual record layouts will vary depending on the subtype chosen. Level 5 (Supporting Data) is independent while the remaining levels represent various levels of traffic data aggregation.

Table B.1. LTPP-CTDB Data Types Available

DATA TYPE	DESCRIPTION OF SUBTYPES AVAILABLE
Level 1	Annual axle loads, ESALs for all vehicles, and other summary statistics
Level 2	Annual axle loads, ESALs by vehicle class, and other summary statistics
Level 3	Volume (daily traffic volumes by lane)
	Vehicle class (daily traffic volumes by vehicle class)
	Weight (daily ESALs and weight ranges by vehicle class)
Level 4	Volume (hourly traffic volumes by lane; "3-card")
	Vehicle class (hourly volumes by vehicle class; "4-card")
	Weight (individual truck weight records; "7-card")
Level 5	Vehicle class data submittal forms
	Weight data submittal forms
	Historical data sheets (Sheets 1-9) plus monitoring estimates (Sheet 10)

Note: Levels 3 and 4 have three distinct data types: volume, vehicle classification, and weight. Likewise, Level 5 is composed of three different types. Record layouts will therefore vary.

Level 3: Daily Traffic Summaries

The same three basic data types (volume, classification, and weight) that exist for Level 4 are also present in Level 3. Only traffic data physically collected and submitted by the SHAs will be included in this file. **Data at this level of the data base will not have been factored, or adjusted by either the submitting SHA or FHWA**. The data are aggregated in three dimensions. First, the data are summarized to represent individual days. Second, they are summed into three categories of direction and lane combinations: (1) the test section lane only, (2) the test section direction, and (3) the opposite direction. Third, the weight data are reported as a distribution: the number of axles weighed in specified weight ranges by axle grouping (single, tandem, tridem, and four or more) and vehicle type. ESALs per vehicle are calculated for each vehicle type. Key parameters for ESAL estimation (pavement type, depth or structural number, and terminal serviceability) are also reported.

This level of the database is intended to allow detailed analysis of the traffic data used to estimate the annual totals. It is specifically designed so that researchers will be able to determine which data are "real" and which data are "interpreted" and make their own assumptions about how limitations in the available traffic data should be overcome. This level of data will also be the starting point for research into different methods for producing annual traffic estimates from short duration count data. It will also provide estimates of seasonal loadings for LTPP researchers who need to separate loadings for particular time periods, as opposed to the annual conditions presented in database Levels 1 and 2.

In Level 3, up to 365 records for each type of data (volume, weight, or class) may appear in the database for each LTPP site for each year since the site was opened for traffic. Records will only occur in the database for days on which an SHA actually collected data, or where a record is necessary to inform a researcher that no data were collected at that site for an entire calendar year.

"Missing data" will not be inferred or entered into this level of the database. Space on the data records is available for an SHA to provide additional information pertaining to this count. Additional information includes factor(s) that a state might normally use to estimate annual totals based on that count, or comments about events that may affect how that particular count should be used by LTPP or a LTPP researcher. The database will store information for both the LTPP study lane and all other lanes for which an SHA submits information.

A formatted example of Level 3 data appears as Figure B.1.

Figure B.1. Example of Level 3 Data

```
Level 3 W-4 Table - Version 1
----------------------------------------
Region: North Central   State: Minnesota   State Id: 1019   SHRP Id: 1085
Route: ST 16   Milepost: 203.34
Location: 3.8 MI. E. OF I 90

Device Type: PERMANENT   Make: IRD   Model: 1060
SENSOR TYPE: BENDING PLATE

Software version: 530
PROCESSING DATE: January 02, 1994

DATE: January 03, 1993   ORIGIN DATE: January 03, 1993   BEGIN TIME: 00:00:00:00
LEVEL 4 SOURCE FILENAME: 7W271085.C13
GPS DIRECTION: 3  GPS LANE:  1
TOTAL CARDS READ: 19  TOTAL CARDS USED: 10  FLAGGED USED CARDS: 0
Classification By: FHWA 6 Digit
PAVEMENT TYPE: Flexible  Depth: 0  Structural Number: 4.26  Service Index: 2.5

Actual Vehicles Weighed - 10

Mean ESALs/Vehicle Calculated From Vehicles -  0.699
Mean ESALs/Vehicle Calculated from Table    -  0.704
```

(CONTINUED)

Figure B.1. (Continued)

----- VEHICLE CLASS 5 -----

Actual Vehicles Weighed - 4

Mean ESALs/Vehicle Calculated From Vehicles - 0.005
Mean ESALs/Vehicle Calculated from Table - 0.004

SINGLE		TANDEM		TRIDEM		QUAD +	
0- 2999	4	0- 5999	0	0-11999	0	0-11999	0
3000- 3999	2	6000- 7999	0	12000-14999	0	12000-14999	0
4000- 4999	1	8000- 9999	0	15000-17999	0	15000-17999	0
5000- 5999	1	10000-11999	0	18000-20999	0	18000-20999	0
6000- 6999	0	12000-13999	0	21000-23999	0	21000-23999	0
7000- 7999	0	14000-15999	0	24000-26999	0	24000-26999	0
8000- 8999	0	16000-17999	0	27000-29999	0	27000-29999	0
9000- 9999	0	18000-19999	0	30000-32999	0	30000-32999	0
10000-10999	0	20000-21999	0	33000-35999	0	33000-35999	0
11000-11999	0	22000-23999	0	36000-38999	0	36000-38999	0
12000-12999	0	24000-25999	0	39000-41999	0	39000-41999	0
13000-13999	0	26000-27999	0	42000-44999	0	42000-44999	0
14000-14999	0	28000-29999	0	45000-47999	0	45000-47999	0
15000-15999	0	30000-31999	0	48000-50999	0	48000-50999	0
16000-16999	0	32000-33999	0	51000-53999	0	51000-53999	0
17000-17999	0	34000-35999	0	54000-56999	0	54000-56999	0
18000-18999	0	36000-37999	0	57000-59999	0	57000-59999	0
19000-19999	0	38000-39999	0	60000-62999	0	60000-62999	0
20000-20999	0	40000-41999	0	63000-65999	0	63000-65999	0
21000-21999	0	42000-43999	0	66000-68999	0	66000-68999	0
22000-22999	0	44000-45999	0	69000-71999	0	69000-71999	0
23000-23999	0	46000-47999	0	72000-74999	0	72000-74999	0
24000-24999	0	48000-49999	0	75000-77999	0	75000-77999	0
25000-25999	0	50000-51999	0	78000-80999	0	78000-80999	0
26000-26999	0	52000-53999	0	81000-83999	0	81000-83999	0
27000-27999	0	54000-55999	0	84000-86999	0	84000-86999	0
28000-28999	0	56000-57999	0	87000-89999	0	87000-89999	0
29000-29999	0	58000-59999	0	90000-92999	0	90000-92999	0
30000-30999	0	60000-61999	0	93000-95999	0	93000-95999	0
31000-31999	0	62000-63999	0	96000-98999	0	96000-98999	0
32000-32999	0	64000-65999	0	99000-UP	0	99000-UP	0
33000-33999	0	66000-67999	0				
34000-34999	0	68000-69999	0				
35000-35999	0	70000-71999	0				
36000-36999	0	72000-73999	0				
37000-37999	0	74000-75999	0				
38000-38999	0	76000-77999	0				
39000-39999	0	78000-79999	0				
40000-UP	0	80000-UP	0				
TOTAL	8	TOTAL	0	TOTAL	0	TOTAL	0
VEHICLES	4	VEHICLES	0	VEHICLES	0	VEHICLES	0

Figure B.1. (Continued)

----- VEHICLE CLASS 6 -----

Actual Vehicles Weighed - 1

Mean ESALs/Vehicle Calculated From Vehicles - 3.022
Mean ESALs/Vehicle Calculated from Table - 2.834

SINGLE		TANDEM		TRIDEM		QUAD +	
0- 2999	0	0- 5999	0	0-11999	0	0-11999	0
3000- 3999	0	6000- 7999	0	12000-14999	0	12000-14999	0
4000- 4999	0	8000- 9999	0	15000-17999	0	15000-17999	0
5000- 5999	0	10000-11999	0	18000-20999	0	18000-20999	0
6000- 6999	0	12000-13999	0	21000-23999	0	21000-23999	0
7000- 7999	0	14000-15999	0	24000-26999	0	24000-26999	0
8000- 8999	0	16000-17999	0	27000-29999	0	27000-29999	0
9000- 9999	0	18000-19999	0	30000-32999	0	30000-32999	0
10000-10999	0	20000-21999	0	33000-35999	0	33000-35999	0
11000-11999	1	22000-23999	0	36000-38999	0	36000-38999	0
12000-12999	0	24000-25999	0	39000-41999	0	39000-41999	0
13000-13999	0	26000-27999	0	42000-44999	0	42000-44999	0
14000-14999	0	28000-29999	0	45000-47999	0	45000-47999	0
15000-15999	0	30000-31999	0	48000-50999	0	48000-50999	0
16000-16999	0	32000-33999	0	51000-53999	0	51000-53999	0
17000-17999	0	34000-35999	0	54000-56999	0	54000-56999	0
18000-18999	0	36000-37999	0	57000-59999	0	57000-59999	0
19000-19999	0	38000-39999	0	60000-62999	0	60000-62999	0
20000-20999	0	40000-41999	0	63000-65999	0	63000-65999	0
21000-21999	0	42000-43999	1	66000-68999	0	66000-68999	0
22000-22999	0	44000-45999	0	69000-71999	0	69000-71999	0
23000-23999	0	46000-47999	0	72000-74999	0	72000-74999	0
24000-24999	0	48000-49999	0	75000-77999	0	75000-77999	0
25000-25999	0	50000-51999	0	78000-80999	0	78000-80999	0
26000-26999	0	52000-53999	0	81000-83999	0	81000-83999	0
27000-27999	0	54000-55999	0	84000-86999	0	84000-86999	0
28000-28999	0	56000-57999	0	87000-89999	0	87000-89999	0
29000-29999	0	58000-59999	0	90000-92999	0	90000-92999	0
30000-30999	0	60000-61999	0	93000-95999	0	93000-95999	0
31000-31999	0	62000-63999	0	96000-98999	0	96000-98999	0
32000-32999	0	64000-65999	0	99000-UP	0	99000-UP	0
33000-33999	0	66000-67999	0				
34000-34999	0	68000-69999	0				
35000-35999	0	70000-71999	0				
36000-36999	0	72000-73999	0				
37000-37999	0	74000-75999	0				
38000-38999	0	76000-77999	0				
39000-39999	0	78000-79999	0				
40000-UP	0	80000-UP	0				
TOTAL	1	TOTAL	1	TOTAL	0	TOTAL	0
VEHICLES	1	VEHICLES	1	VEHICLES	0	VEHICLES	0

*U.S. GOVERNMENT PRINTING OFFICE 1998/441-498