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.

! <u>The continuous nature of most data generated by ITS removes temporal sampling bias from estimates and allows the study of variability.</u> 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).

! <u>Data to meet emerging requirements and for input to new modeling procedures will have to be more detailed than what is now collected.</u> 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.

! <u>Use of data generated by ITS for multiple purposes is a way to stimulate the support of other stakeholders for ITS initiatives.</u> 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.

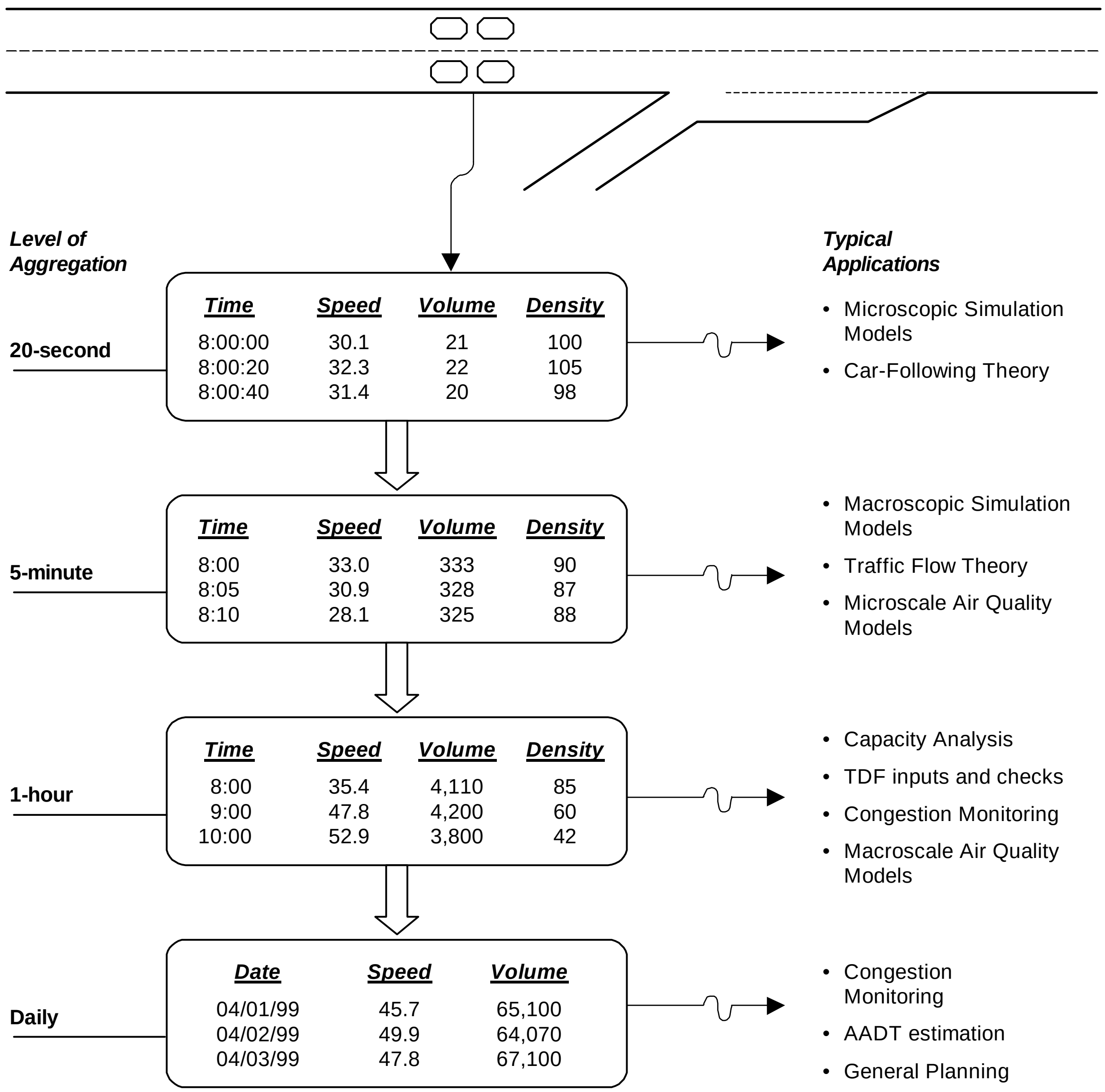
Table 2.1: Use of ITS Data for Stakeholder Applications

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.
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.

Table 2.1: Use 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/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: Use of ITS Data for Stakeholder Applications



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. Facilityspecific 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).