



Author: Shawn M. Turner, P.E.

ITS Data Archiving: Summary Report

As intelligent transportation system (ITS) applications continue to be deployed in Texas, large quantities of data are becoming available from ITS sensor and detector systems. The data, which typically include traffic conditions, incident/accident data, ITS control responses, and other roadway or environmental characteristics, are collected by

traffic management centers and used in real-time by the Texas Department of Transportation (TxDOT) and partner agencies to manage the transportation system. When these data are archived, they can be used for numerous transportation analyses, such as estimation of ITS impacts, computer model calibration, congestion monitoring, transportation

planning, or even pavement design. However, little guidance currently exists for data archiving.

The objective of this research project was to assemble guidance that could help TxDOT in further developing data archiving systems across the state. This guidance includes information concerning:

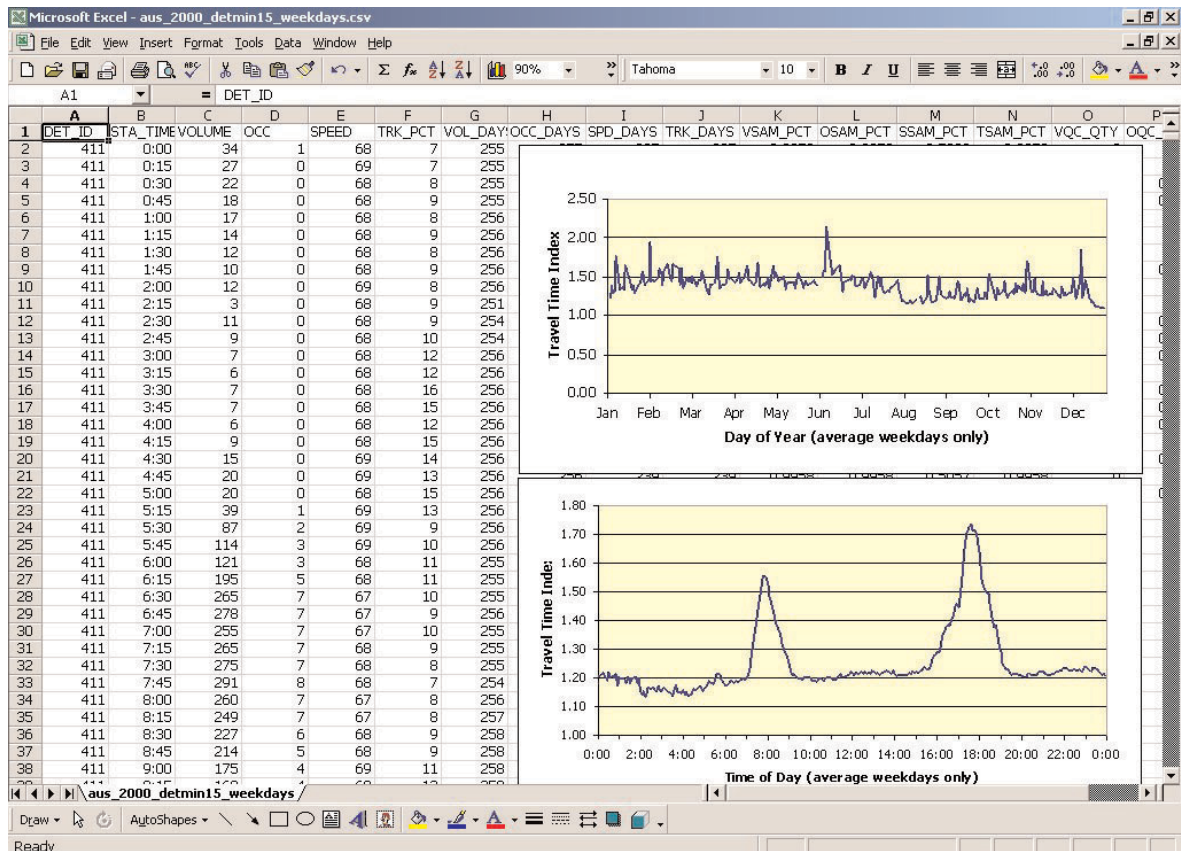


Figure 1. Archiving data from ITS saves money and provides better information for decisions



- basic principles of data archiving,
- determining what data and how much to archive,
- performing quality control to ensure that data meet certain quality requirements, and
- developing a data archiving system using the National ITS Architecture and relevant standards.

What We Did . . .

The research team conducted several activities to develop the project guidance and recommendations. We met with numerous stakeholders in Austin, Dallas, Ft. Worth, Houston, and San Antonio to discuss issues and ideas about data archiving. As a follow-up, Texas Transportation Institute researchers conducted a user requirements survey in the Dallas-Ft. Worth area to gather more specific information about what data are being collected and what data are most important to archived data users. We also participated in a national committee that is currently developing data archiving standards.

We talked with transportation agencies in other states about data archiving in an attempt to determine what works best. Several other areas have data archiving systems that have been in place for several years; thus, TxDOT could learn best practices from the collective experience of numerous other states.

We also analyzed archived ITS data from Austin, Houston, and San

Antonio to ascertain quality and completeness. This gave us a better sense of what procedures are necessary to ensure that data archives have relatively accurate, complete data.

What We Found . . .

The research team developed key findings for current practices and issues in ITS data archiving. This report presents these key findings first for data archiving in Texas, then for data archiving systems in other states and areas.

Several regions in Texas (i.e., Austin, Houston and San Antonio) currently archive some of the data that are collected by ITS. Ft. Worth has plans to archive some of its data in the near future, and Houston and Austin would like to further develop and extend their archiving system capabilities. For these three areas that do archive data, most archive freeway detector data to large or numerous text files in the original field-collected format (i.e., lane-by-lane, 20-second or 1-minute time interval). Other groups (such as planners, designers, and air quality analysts) would like to be able to use these archived data but are unable to because of size and format. To date, most users have been researchers or other “power users” who have specialized database or programming skills.

In stakeholder and other meetings, we heard concerns about the following issues: a) Who should take responsibility for owning and managing the data archives? Won't data archives be expensive because they require large amounts of computer storage space? b) What

ITS data are most important to archive and at what level? What are the archived data users' requirements? c) How “good” or accurate are the ITS data, and how much quality control is performed? d) Isn't some data collection being duplicated between the planning and operating divisions?

In surveying and talking with people outside of Texas, we found several areas that have developed more sophisticated data archiving systems that enable ordinary computer users to access large databases of archived ITS data. The majority of areas, however, were at a similar early stage of data archiving as TxDOT. That is, many other areas simply log original detector data to a text file and are just beginning a dialogue with the many potential users of the data. We conducted in-depth studies of those areas that have already developed effective data archiving systems, which have these characteristics in common:

- A workgroup or agency has taken or been assigned the responsibility of operating and maintaining the archive.
- Most systems have started as modest prototypes focused on a single source of data, with the most widely archived data being traffic data (i.e., vehicle volume, occupancy, and speed) from detector systems.
- The data archiving system has been developed in a way that permitted ordinary users with typical desktop computers to access and analyze the data.



- The most effective method of data access and distribution has been through the Internet or CD-ROM.
- Original data as collected from the field have been saved permanently in off-line storage, but data summaries were made available for most users.
- Quality control methods (most relatively simple, but some more complex) are used to flag or remove suspect or erroneous data from the data archive.
- Adequate documentation of the data archive and the corresponding data collection system is provided.

The Researchers Recommend . . .

Based upon the findings of the research team, we offer the following recommendations:

TxDOT should begin (or continue in some regions) the dialogue to determine who will be responsible for maintaining ITS data archives at the regional and statewide level. The responsibility for data archives should be considered part of a regional traffic management center’s “concept of operations.” At the statewide level, responsibility should be assumed for archived ITS data of statewide interest, such as traffic volumes and speeds on TxDOT-maintained roadways. In some districts such as Houston or Ft. Worth, planning has already started for district-level or regional data archives. In other regions and at the statewide level, no one has assumed this responsibility for further development of data archives.

The dialogue about data archiving at the regional level should include the TxDOT district, the metropolitan planning organization, city and county transportation agencies, transit agencies, and other interested stakeholders as identified in [Report 2127-3](#). Similarly at the statewide level, TxDOT’s Transportation Planning and Programming (TPP) Division, the Texas Natural Resource Conservation Commission (TNRCC), and other interested statewide groups or agencies should be engaged in conversations about maintaining archived ITS data in statewide information systems.

TxDOT should develop one or two modest prototypes of single-source data archives (freeway detector data being the most desirable) to demonstrate proof of concept to system developers as well as potential data users.

These prototypes will make the existing detector data files (which are large, multi-million record text files) easily accessible to typical computer users, thus enabling most users to better understand the type and quantity of data being collected in Texas’ traffic management centers. The prototypes will also help system developers and integrators to better understand the issues of larger data archiving efforts. This “start small but think big” approach comes from other industries, where large, complex data warehousing efforts have failed or struggled for years trying to “be all things to all people.”

We recommend that data archiving efforts in Texas focus on permanently storing freeway detector data (vehicle volume, occupancy, and speed) at a common

time interval, such as 5 minutes by lane, then making various data summaries and reports available through an easy-to-use interface (spreadsheet or web browser). Once experience has been gained, enhancements can include integration of other databases (such as roadway incidents and work zones) as well as more sophisticated analysis and reporting features.

TxDOT should consider the data archiving guidelines in [Report 2127-3](#), as well as the results from ongoing federal activities, when further developing data archiving systems.

The guidelines ([Report 2127-3](#)) address several relevant issues in the development of data archiving systems, such as: a) basic principles of data archiving; b) determining what data and how much to archive; c) performing quality control to ensure data meet certain quality requirements; and d) developing a data archiving system using the National ITS Architecture and relevant standards. Beyond the guidelines that were developed in this TxDOT project, however, the Federal Highway Administration (FHWA) has a data archiving program that is producing useful information for implementing the archived data user service (ADUS) as documented in the National ITS Architecture.



For More Details . . .

The research is documented in the following reports:

2127-2, TxDOT ITS Data Archiving Research Project Activities

2127-3, Guidelines for Developing ITS Data Archiving Systems

Research Supervisor: Shawn M. Turner, P.E., TTI, shawn-turner@tamu.edu, (979) 845-8829

Researchers: Robert Benz, P.E., TTI, r-benz@tamu.edu, (713) 686-2971
Scott Cooner, P.E., TTI, s-cooner@tamu.edu, (817) 462-0525
Jason Crawford, P.E., TTI, jcrowford@tamu.edu, (817) 462-0534
William Frawley, TTI, w-frawley@tamu.edu, (817) 462-0533
Darryl Puckett, TTI, d-puckett@tamu.edu, (713) 686-2971

TxDOT Project Director: Regina Flores, Project Director, TxDOT, rflores@dot.state.tx.us, (817) 370-6820

To obtain copies of the report, contact Dolores Hott, Texas Transportation Institute, Information & Technology Exchange Center, (979) 845-4853, or e-mail d-hott@tamu.edu. See our on-line catalog at <http://tti.tamu.edu>.

TxDOT Implementation Status January 2002

As a result of their participation in this research project, the North Central Texas Council of Governments (NCTCOG) has committed to developing a regional data archive in Dallas-Ft. Worth. As of November 2001 NCTCOG has allocated some of its resources and is preparing a budget and scope for this archive development.

For more information, please contact the Research and Technology Implementation office at (512) 465-7403 or e-mail RTImain@dot.state.tx.us.

Related Research...

The TransLink® Research Program (TxDOT Project 0-1752) at TTI developed a prototype Internet-based data archiving system for San Antonio (see Report 1752-2) several years ago. TransLink® researchers are currently developing a prototype data archiving system that works with TxDOT's ATMS software.

TTI researchers have documented data archiving issues in a study conducted for the Federal Highway Administration in a report entitled "ITS Data Archiving: Case Study Analyses of San Antonio TransGuide Data" (Report No. FHWA-PL-99-024). This FHWA study and the TransLink® program jointly produced a CD entitled "Resources for Implementing the Archived Data User Service (ADUS)."

In an interagency contract with TxDOT's Austin District, TTI developed a CD-based data archive using Austin's year 2000 ITS freeway detector data. This simple, CD-based archive could be considered a prototype similar to that suggested in the recommendations.

The University of Texas and Texas Southern University are currently investigating the use of archived ITS data in transportation planning (TxDOT Project 0-4054). This project will conclude in August 2002.

FHWA has an ongoing data archiving research and implementation program. Ralph Gillmann and James Pol at FHWA Headquarters are the data archiving contacts.

YOUR INVOLVEMENT IS WELCOME!

Disclaimer

This research was performed in cooperation with the Texas Department of Transportation (TxDOT) and the U.S. Department of Transportation, Federal Highway Administration (FHWA). The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Texas Department of Transportation (TxDOT) or the Federal Highway Administration (FHWA). This report does not constitute a standard, specification, or regulation. Trade names are used solely for information and not for product endorsement. The engineer in charge of the project was Shawn Turner, P.E. #82781.