

ITS Archived Data User Service in Portland, Oregon: Now and Into the Future

Robert L. Bertini, Spicer Matthews, Steve Hansen, Andrew Delcambre, Andy Rodriguez

Abstract—The implementation and status of the recently developed archived data user service for the Portland, Oregon metropolitan area is described. This includes a discussion of the hardware and software implementation used in building this system. The basic functionality of the data archive is presented along with an update of future design plans for the system. Along with the implementation details, a short tour of the user interface is included. The description of several of the data archive's features are complemented with comparisons of other systems that are being implemented elsewhere.

I. INTRODUCTION

INTELLIGENT Transportation Systems (ITS) are known for producing large amounts of data. Data can come from many different sources, anything ranging from loop detectors installed in the roadways, Global Positioning Systems installed in motor vehicles, to human data collection. Often the data are used in only a real-time fashion. The data collected are used at that instant and then discarded. One data source that falls into this category is global positioning system data used by the Oregon Department of Transportation (ODOT) as part of their incident response program. The data collected by ODOT are used to manage the system in real time but are not exploited in systematic way after the fact for operational analysis.

Partially in response to this opportunity, the Portland State University (PSU) ITS laboratory is archiving ODOT freeway inductive loop detector data on a systematic basis. The data are being streamed to servers located at PSU and are then archived in a relational database. This system is known as the Portland Transportation Archive Listing (PORTAL). Once the data are retained, individuals and agencies such as transportation planners, administration research by key stakeholders, including metropolitan planning organizations (MPOs), state transportation planners, traffic management operators, transit operators, and transportation researchers will benefit from the archived data. The PORTAL system is being developed in an educational environment and therefore it will also contribute

toward improved engineering educational outcomes.

This archived data user service (ADUS) project has been derived from the U.S. Department of Transportation (DOT) National ITS Architecture. The idea behind ADUS was “the unambiguous interchange and reuse of data and information throughout all functional areas.” ADUS requires that data from ITS systems be collected and archived for historical, secondary, and non real-time uses, and that these data be made readily available to users. Interest has been sparked throughout the U.S. to start developing archiving systems. The ITS community in the Portland region has designated PSU as the official ITS data archiving entity, which partially drives the development of PORTAL.

The Federal Highway Administration (FHWA) maintains a website with links to other online ADUS projects (see <http://www.fhwa.dot.gov/policy/ohpi/travel/adus.htm>), including implementations in Arizona, California, Maryland, Minnesota, Texas, Virginia and Washington. Further, the ITS Plan for Canada includes an ADUS component. In addition to past research projects related to archived ITS data, the European Commission is sponsoring the development of the European Transport policy Information System (ETIS), which is envisioned as a “flexible, scalable and open” transportation data archive (www.etis-eu.org).

II. DATA SOURCES AND ARCHIVE STRUCTURE

A. National ITS Architecture

As defined in the National ITS Architecture [2], ADUS provides the historical data archive for all relevant ITS data and incorporates planning, safety, operations, and research communities. ADUS provides the data collection, manipulation, and dissemination functions of these groups, as they relate to all ITS data. The larger ADUS umbrella includes the Historical Data Archive (HDA) repositories, and five major functions: Operational Data Control (ODC), Data Import and Verification (DIV), Automatic Data Historical Archive (ADHA), Data Warehouse Distribution (DWD) and the ITS Community Interface (ICI).

The PORTAL system currently focuses only on freeway data and is passive in the sense that the archive receives the data from ODOT as an external user. Thus it is not possible to claim that the entire ADUS Architecture umbrella is fully implemented. However, as PORTAL has been developed, a design goal has been to adhere to the National ITS Architecture, including these five functions:

1. Store data in format received from ITS subsystems.

Manuscript revised Jun. 21, 2005. The National Science Foundation supported this project under Grant No. 0236567.

R.L. Bertini is with the Department of Civil & Environmental Engineering, Portland State University, Portland, OR 97207-0751 USA (phone: 503-725-4249; fax: 503-725-5950; e-mail: bertini@pdx.edu).

S. Matthews, A. Delcambre and A. Rodriguez are with the Department of Computer Science, Portland State University, Portland, OR 97207-0751, (email: spicer@pdx.edu, adelcambre@gmail.com, rodrigur@pdx.edu).

S. Hansen is with the School of Urban Studies & Planning, Portland State University, Portland, OR 97207-0751 (email: stevenh@pdx.edu).

2. Accommodate aggregation and reduction of the data flows, depending on the type of data represented.
3. 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.
4. Apply quality control procedures to the data, including the flagging of suspect data and the editing of data.
5. Distinguish between the following data types: unprocessed (raw), edited, aggregated, and transformed.

As PORTAL develops, the design continues to adhere to these functional goals. With regard to meta-data, PORTAL currently does satisfy the ADUS Architecture requirements by including a meta-data repository and maintaining the meta-data schema for all data entering the system, including information generated in the field at the controller level and in the traffic management center.

B. Case Study

The California Department of Transportation (Caltrans) and the University of California at Berkeley have partnered to create the PeMS system, (performance measurement system, see <http://pems.eecs.berkeley.edu/Public/>). PeMS is used for monitoring freeway speeds as well as estimating travel times and volumes. PeMS is a statewide system and includes archived data from most Caltrans districts, plus incident data and other reporting capabilities. PeMS is stored on-line with an Oracle database containing over two terabytes of data. The database grows at two gigabytes per day.

The PeMS data comes from California's 23,138 freeway loop detectors measuring 30-sec vehicle count and occupancy [4]. Detector data are transmitted via phone lines and other communications media to a central computer and then sent to PeMS via the Caltrans Wide Area Network (WAN). These data are aggregated into 5-minute intervals and loaded into the PeMS data warehouse.

The development of the PORTAL project has been inspired by the California's PeMS project, though PORTAL currently operates in a research lab/educational environment. PORTAL is currently archiving data from inductive loop detectors in the Portland metropolitan area, including Vancouver, Washington. There are future plans for incorporating data from other parts of Oregon and Washington.

C. Portland's ADUS Design

In July 2004, PORTAL "went live," streaming data from ODOT Traffic Monitoring Operations Center (TMOC) to PSU. Currently, the data come from 485 inductive loop detectors installed in the freeway mainline and on-ramps in the Portland region. These detectors are in place as part of ODOT's advanced traffic management system that includes ramp metering, incident management, traveler information and a central traffic management operations center. As the system grows, more detectors and ramp meters are being added. PORTAL is designed to handle changes in the loop

detector configuration as the system grows. As shown in Fig. 1, these data are transferred over a secure direct fiber optic connection between ODOT and PSU. Every 20 seconds XML data are pulled down from the ODOT servers to the PSU archiving servers. Once retained at PSU, the 20 second data are inserted into a PostgreSQL relational database management system (RDBMS) running on a Fedora Core Linux system. The raw 20 second data are not altered. Every morning at 3 a.m., the 20 second data are aggregated to 5 minute data in order to allow for faster processing via the web interface described below. These data are appended to a table for the relevant month. The data are also aggregated to 1-hour bins and appended to a table that contains data for an entire year. The one hour and 5 minute data allows for initial queries to be faster, allowing the user to fine tune their queries and work their way down to the 20 second data.

The tables in the RDBMS are extracted nightly to text files in which the data are compressed and shipped to a backup server that is located several blocks from the main archiving server. It was felt that having the backup server off-site from the main archiving server was best to ensure fail safe storage and backup of the archived data. In addition to the backup server, a yearly backup is made to a DVD-ROM. The two server rooms are climate-controlled, secure and managed as

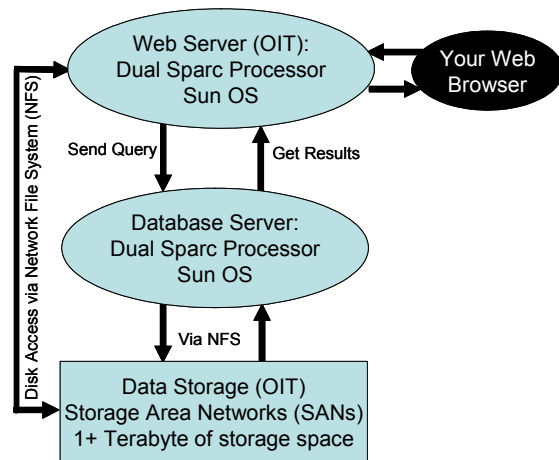


Fig. 1. Network layout of PORTAL

part of the PSU office of information technologies system.

D. The PORTAL Network

The raw inductive loop detector data is streamed via different network technologies from each particular loop detector station to the ODOT traffic management operations center. In most cases the data are streamed over the extensive regional fiber optic network or telephone lines. In some cases outside of the Portland region wireless networks are being tested for data transmittal to ODOT.

As part of the regional ITS backbone, the transportation agencies in the Portland region are finalizing plans to install a private ITS fiber network linking the key regional ITS organizations. Some of the major nodes on the network are

ODOT, the City of Portland, the Port of Portland, TriMet and PSU. This network is designed to be a private and secure way for ITS organizations to share data.

E. PORTAL's User Interface

As suggested in [3], user access to the data archive is through a web-based interface providing point-and-click access to both raw data sets and a wide range of common summary data and performance measures. Users have the option of graphically viewing data or outputting the data in tabular form. Additionally, PORTAL has the feature of downloading all queried data as a comma separated text (CSV) file, since users may want to perform individualized manipulations off-line.

Fig. 2 shows the PORTAL home page. The PORTAL web interface (<http://portal.its.pdx.edu>) Persons interested in using PORTAL can register for an account from the home page. An administrator of the PORTAL team can approve the account or request more information to confirm that the new user in question truly has a need to access the data. As of this writing there are 90 registered PORTAL users. Because of the research lab/educational setting, we have many students using the data to supplement their textbook and classroom presented data. Because of the interest in transportation performance in the Portland region, several elected transportation decision-makers are also PORTAL users.

Users login into PORTAL for many reasons. Some are looking for raw data that they can import and process with personalized algorithms. Others use PORTAL to organize data in a fashion that satisfies their interests. From the web interface, different specialized queries of the RDBMS can be

graphs are produced using the open source tool Gnuplot. For example, on the data fidelity page a user can see a line graph showing the percentage of the data that is determined to be valid versus the percentage of error-prone data for a particular location and time period.

As with any widespread data collection system, data fidelity problems have many different causes. PORTAL includes a data fidelity page that allows users to understand the quality of the data, both at the detector level and system-wide. This is useful to an organization such as ODOT. An employee of ODOT can log in and see that a loop detector has malfunctioned and dispatch maintenance personnel out in the field to examine the problem. In a real-time system in which data is discarded, it is not as easy to spot broken loop detectors.

F. Performance of PORTAL

The RDBMS stores data physically on a redundant array of independent disks (RAID) providing both high-speed access and increased reliability through redundancy in the event of hardware failure. Currently PORTAL has a capacity of 1.2 TB, with approximately 200 MB added each day. The technology chosen to implement this system is completely open source and has been proved to be the top quality technology in industry currently. PORTAL uses an Apache web server with processing powered by the PHP hypertext preprocessor module.

III. A TOUR OF PORTAL

A. The PORTAL Breakdown

As shown in Fig. 2, when a user logs onto PORTAL, there are several menu options, including Timeseries, Grouped Data, Data Fidelity, Raw Data and Weather. Each section contains allows users to fine tune a particular query. New functionality is being developed, details are provided below.

B. Time Series

As shown in Fig. 3, the time series section allows the user to perform the queries on the 5 minute aggregated data to extract volume, speed, occupancy, vehicle miles traveled, vehicle hours traveled, travel time, and delay. The figure shows an example of one day's volume measured at one location (the y-axis is presented in veh/hr) The user can specify a specific loop detector station as well as a full freeway segment. Additionally, the user can choose the time period and specific travel lanes. The user has the option of plotting the data as shown in the figure, viewing the data in a table, or exporting that data as a comma separated text file.

C. Grouped Data

The Grouped Data section of the PORTAL user interface allows the user to display summaries of data from multiple days. The data can be grouped by hour of day, day of week, or week of year. The user can also determine the particular statistics that are of most interest or relevance, including the mean, minimum value, maximum value, and standard

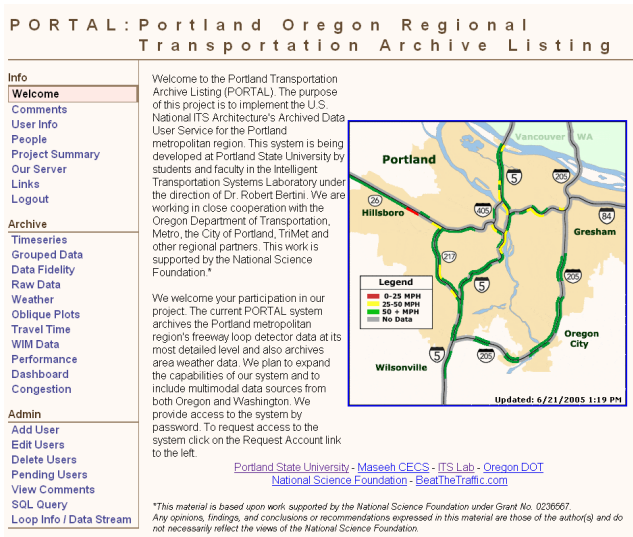


Fig. 2. PORTAL homepage

performed with just a few clicks of the mouse. For example, if a user only wanted to see the loop data from a certain loop detector on a certain day, this is possible with little effort.

Many users are not interested in seeing large amounts of data but do want to see graphs of the collected data. The

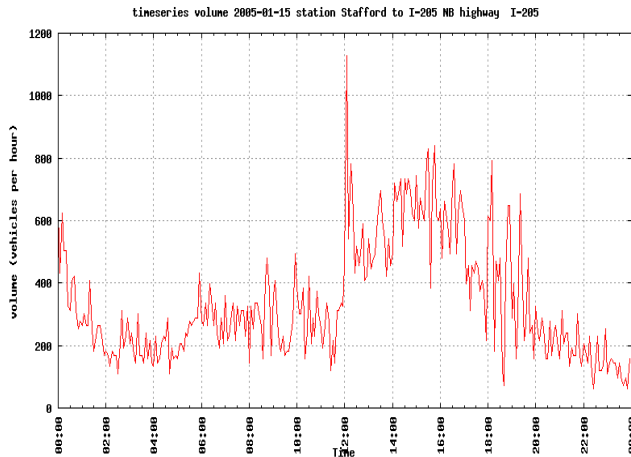


Fig. 3. Sample time series plot generated from the PORTAL system.

deviation. As in the time series section, the user has the choice of plotting the data, displaying a table of the data, or downloading a comma separated text file. Fig. 4 shows an

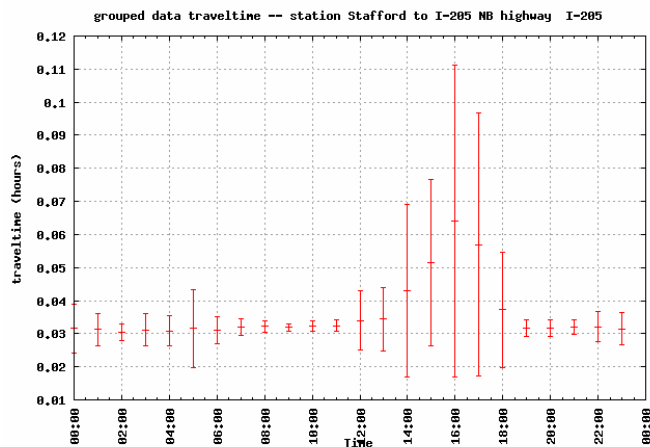


Fig. 4. Sample Plot of grouped data, from PORTAL system

example of grouped data for travel time across one freeway segment over a six month period. The plot shows the mean and plus/minus one standard deviation.

D. Data Fidelity

For any data archive, understanding, documenting and communicating the data quality characteristics is critical. The National ITS Architecture emphasizes the importance of meta-data that clearly describes any known problems with the data that is distributed. This section of the PORTAL site allows the user to determine whether particular loop detectors reported valid data according to the data fidelity algorithm. ODOT's system wide adaptive ramp metering system (SWARM) reports several data quality measures and PORTAL has also implemented several other quality checks. The user can either acquire information on the health of a specific detector, or can obtain the percentage of healthy loop readings on a specified highway section. Fig. 5

is an example of the detector health information available. This has been useful for assisting ODOT with the upgrade of some of its traffic management system software and in updating some of its ramp meter controller hardware.

E. Daily and Monthly Reports

This section of PORTAL was requested by the various agencies around Portland, so that they could assimilate a gauge of the total system health for a specific day or month. This includes the average congested travel time for a road segment, the percent of time that a segment was congested during the peak period, and the buffer index (the 95th percentile travel time for the month for a road segment). These performance measures can provide traffic managers and transportation planners a good idea of the system's health at any time. Fig 6 shows a sample travel time reliability plot for northbound I-5 for May 2005. The mean and 95th percentile travel times are shown along with the percentage of time that each 5-min time slice was congested for that month. Fig. 7 shows a sample monthly report developed by the FHWA to be a national report card for ten key cities. Currently, the FHWA must extract large raw data files and perform their own calculations. If FHWA desired Portland data, they could display the monthly report shown in figure 7 with one click of the mouse.

F. Weather

In an effort to provide additional information about traffic and freeway conditions, weather data are also archived on PORTAL Archived data includes rainfall and temperature collected by the National Oceanic and Atmospheric Administration, (NOAA). These weather stations are located at the Portland International Airport, as well as at the airports in suburban Hillsboro and Troutdale, on the west and east sides of the region respectively. The data are collected with the Python based module pyMETAR. The pyMETAR module provides easy access to the NOAA weather reports [5]. These data are presented much like our other data in a graphical format, data table or raw comma separated format. Fig. 8 shows an example of weather data from one day. The line graph shows the hourly temperature variation over the day (left hand axis) while the bar chart (right hand axis) shows the hourly measured precipitation. A weather "pop up" button appears on each PORTAL report screen so the user can rapidly view whether the day in question was typical or not. Weather data is and will continue to be used in monthly and daily reports. The idea is when a section of road seems to be congested a researcher can look at the weather information and determine if weather was a potential cause. As PORTAL grows to include data outside of the Portland metropolitan area, our weather archiving will expand as well, using the road weather information system (RWIS) sites in Oregon.

IV. LOOKING TO THE FUTURE

Future PORTAL activities will include development of meta-data schema, consistent with the national ITS

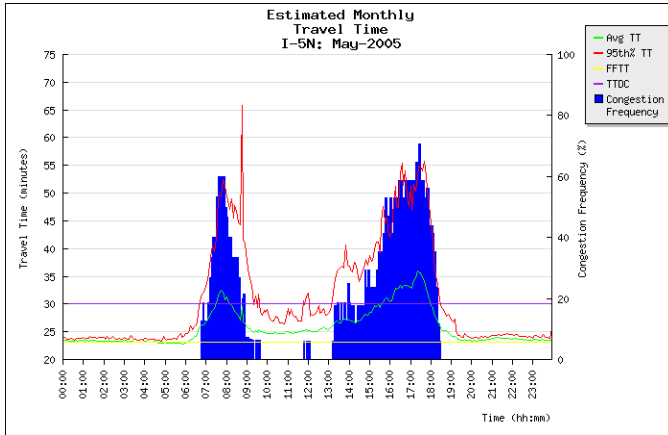


Fig. 6. Monthly Reliability Report for Northbound I-5

Architecture, particularly with regard to the aggregation methods used. Meta-data are critical components of any data archive and as PORTAL from research to deployment, this will be the primary means to assure future scalability.

A. Adding Data Sources

Some possible added data sources are: TriMet bus data from the on-board GPS data loggers, automated traffic recorders along different sections of the highways in Oregon, and data from weigh in motion weight enforcement stations around Oregon that record information about trucks as they are weighed at 12 locations in Oregon.

Each TriMet vehicle includes an on-board memory card that is pre-programmed with that bus' schedule information. Therefore, the bus "knows" where it is located at all times, and links that with the schedule, enabling a schedule status (late/early) report to the operator. Further, as part of the archived data component of the bus dispatch system (BDS), a data record is written to the memory card at each bus stop. These records contain the route number, direction, trip number, date, vehicle number, operator ID, bus stop ID, stop arrival time, stop departure time, boardings, alightings, and passenger load. Also recorded is whether the doors of the bus were opened, whether the lift was used, the dwell time, maximum speed between the previous and current stop, and

GPS coordinates. The memory card data are downloaded at the end of each day and compiled into a database. This information is used by TriMet to measure system performance on an ongoing basis and to aid in the bus route scheduling effort that occurs each quarter.

One of the major benefits of TriMet's real time and archived BDS data is that it can be used to extract performance characteristics from the region's arterial system. The PORTAL team plans to begin testing the use of both data sources as part of the Portland archive in the near future for this purpose. As the PORTAL project grows and seeks funding for more infrastructure, the team continues to maintain a scalable system that is dynamic enough to add archiving of different datasets with ease.

B. GIS Integration

PORTAL is implementing some geographic information systems (GIS) features into the system. PORTAL's current graphing capability includes color contour plots in the Timeseries and Grouped Data pages. Applying these data onto a map will be a powerful visual tool. One application of this technology might be to show average volume by highway segment displayed as varying line thickness on a map of the system. Other ideas include a click-able map that would allow the user to choose a particular station for analysis. The user also might be able to click on the map at two locations and learn the travel time between the two points.

C. Vehicle Length Implementation

The current loop controller firmware does not support vehicle length calculation. There is a firmware upgrade in progress to implement this feature. PORTAL will be expanded to allow queries involving minimum or maximum vehicle length. This would enable user to limit calculations only to long vehicles such as tractor-trailers.

D. Travel Time Algorithm

The PORTAL travel time algorithm right now is simply the average speed divided by the segment length. Although this does give an approximate travel time, it does not accommodate for changes in the travel time during the time period (five minutes). There are other algorithms that more accurately approximate the average travel time. These algorithms take into account elements such as the shock waves from downstream detectors when there is congestion

downstream. There are plans to implement alternative algorithms in PORTAL so that we may get a more accurate reading for travel time. There are also plans to deploy probe vehicles with GPS sensors to determine actual segment travel times, which can then be compared to the travel time from the system travel time algorithm(s).

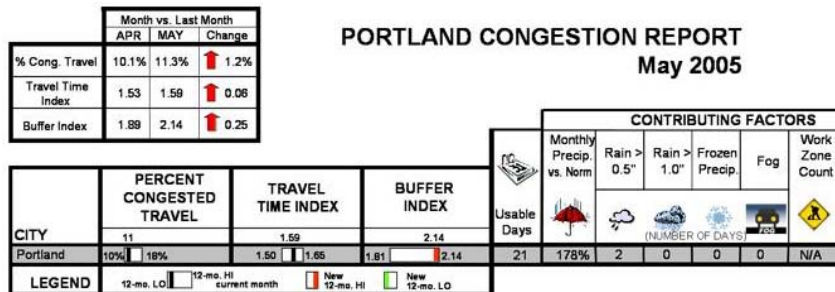


Fig. 7. Sample Urban Congestion Report. May 2005 Portland data

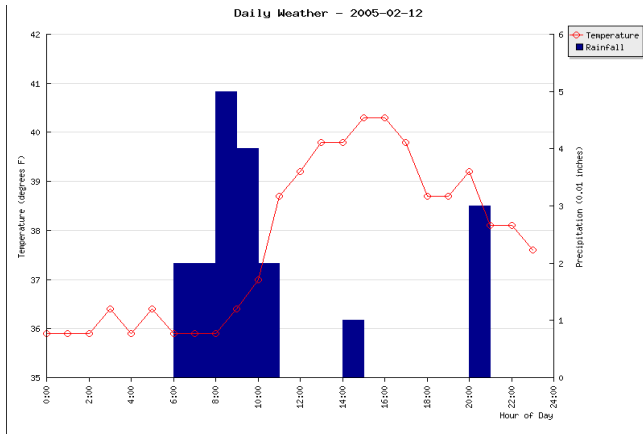


Fig. 8. Sample weather plot from the PORTAL system.

E. Computer Aided Dispatch Database

Portland's ATMS includes an incident management system [1], which in turn generates a large computer aided dispatch (CAD) database that contains information about all recorded incidents on Portland's freeway system. This information includes the type of incident, which lanes were blocked as a result of the incident, and the start and end time of the incident. The TMOC can record the X - Y coordinates of an incident, but it is only required if the incident were severe enough to cause significant delay. In 2001, coordinates were recorded for 10% of all incidents [1]. We plan to archive the CAD data to identify times and locations when incidents affected traffic flow.

F. User Benefits

Currently PORTAL has 90 registered users. One primary mission of the project will be to assess user reaction to the system and attempt to measure the user benefits of the data archive implementation. Currently, PORTAL is presented to prospective users in seminars, lab tours and one-on-one demonstrations, and reactions are documented for the PORTAL team on an internal website. Users can also readily submit comments and suggestions via the PORTAL website. We envision developing an interactive web-based user survey to periodically assess user reaction to the site, its usability, flexibility and performance.

V. CONCLUSION

PORTAL has brought students from different disciplines together to work on an important project. The PORTAL team utilizes students' skills from urban planning, civil engineering, and computer science. From an academic perspective, this NSF funded project has provided the PORTAL team members with real-world experience while providing civil engineering and urban planning students real traffic data to use in their studies as opposed to hypothetical text book data. At the industry level, the PORTAL project has provided professionals from all over the world with the wealth of archived data that can be used to better our roadways. This project has provided an opportunity to test

the National ITS Architecture in a real-world application, and development of this system will continue to meet the Architecture requirements. The PORTAL project is still in its infancy and its possibilities of growth are only beginning.

ACKNOWLEDGMENT

The authors gratefully acknowledge the National Science Foundation, ODOT, TriMet and the City of Portland as well as the TransPort organization for their ongoing support and guidance in the development of PORTAL.

REFERENCES

- [1] Bertini, R.L., Rose, M. and El-Geneidy, A. Using Archived Data to Measure Operational Benefits of ITS Investments: Region 1 Incident Response Program. Portland State University, Center for Transportation Studies, Research Report, June 2004.
- [2] National ITS ADUS Addendum
- [3] S. Turner. Guidelines for Developing ITS Data Archiving Systems. Report 2127-3. FHWA, U.S. Department of Transportation, Texas Department of Transportation and Texas Transportation Institute, 2001.
- [4] S. Turner. ITS Data Archiving: Case Study Analysis of San Antonio TransGuide Data. Texas Transportation Institute, August 1999.
- [5] <http://www.schwarzvogel.de/software-pymetar.shtml>, accessed July 30, 2000.