# Creating a National Nonmotorized Traffic Count Archive

## Process and Progress

Krista Nordback, Kristin A. Tufte, Morgan Harvey, Nathan McNeil, Elizabeth Stolz, and Jolene Liu

**Robust bicycle and pedestrian data on a national scale would help promote effective planning and engineering of walking and bicycling facilities, build the evidence-based case for funding such projects, and dispel notions that walking and cycling are not occurring. To organize and promote the collection of nonmotorized traffic data, a team of transportation professionals and computer scientists is creating a national bicycle and pedestrian count archive. This archive will enable data sharing by centralizing continuous and short-duration traffic counts in a publicly available online archive. Although other archives exist, this will be the first archive that will be national in scope and enable data to be uploaded directly to the site. This archive will include online input, data quality evaluation, data visualization functions, and the ability to download user-specified data and exchange the data with other archives and applications. This paper details the first steps in creating the archive: (*a*) review count types, standard formats, and existing online archives; (*b*) list primary functional requirements; (*c*) design archive architecture; and (*d*) develop archive data structure. The archive's versatile data structure allows for both mobile counters and validation counts of the same traffic flow, an innovation in design that greatly expands the usefulness of the archive.**

Robust bicycle and pedestrian data on a national scale would serve numerous purposes. Access to a centralized nonmotorized traffic count archive will open the door for innovation through research, design, and planning; provide safety researchers with a measure of exposure; provide fundamental performance metrics for planning and funding decisions; and allow policy makers and transportation professionals to better support the public's desire for livable communities. Numerous jurisdictions have initiated nonmotorized traffic count programs. However, many agencies and policy makers, who need data to support investment decisions, are in locations without a centralized count program. This lack of access to count data may lead some decision makers, planners, and engineers to assume that cycling and walking levels are close enough to zero to be ignored. Providing reliable numbers may reveal that a surprising amount of walking and bicycling is taking place.

The lack of a centralized data archive and common data formatting inhibits data sharing and access, thereby greatly reducing the utility of this growing, but dispersed, data set. To remedy this problem, a team of researchers at Portland State University in Portland, Oregon, is creating a national online nonmotorized traffic count archive (*1*). For the purposes of this paper, the term "archive" is used to indicate a data storage and access system. While other online bicycle and pedestrian traffic count archives exist, this one is the first that is both national in scope and enables data to be directly uploaded to the site (*2*). This archive will include online input, data quality evaluation, data visualization functions, and the ability to download user-specified data and exchange the data with other archives and applications. This archive is unique and the first of its kind because of both front-end and back-end software functionality that wraps around the database. This archive addresses the need for a national nonmotorized traffic data one-stop location that provides multiagency data access, distribution, and archiving.

The need for a centralized archive is demonstrated by the current state of bicycle and pedestrian counts in which many counts never leave the hard drives and servers of the agencies collecting them and many such data are lost. For example, when compiling continuous count data for the state of Colorado, the team identified data from at least six local agencies in addition to that collected by the state department of transportation (*3*). None of these data had been included in a centralized archive and were thus difficult to share with other agencies that might be searching for them. While the National Bicycle and Pedestrian Documentation Project (NBPDP) does provide a standard count data collection format and invites participants to submit their counts, these data are not available to the public and are not compiled into a database (*4*). How would a nonmotorized traffic count archive be created and what are the basic elements that should be included? This paper provides some answers.

This paper details the first steps in creating such an archive and associated functionality: (*a*) review of count data types, standard data formats, and existing online archives; (*b*) list of primary functional requirements of the archive; (*c*) definition of the basic architecture for the archive; and (*d*) details of the archive data structure. The paper concludes with next steps for the effort. The goal of this paper is to inform others about the nonmotorized traffic volume data archive efforts and to illustrate a framework for building an open-source, practice-ready application for all public and private entities to use.

K. Nordback and J. Liu, Oregon Transportation Research and Education Center; K. A. Tufte and M. Harvey, Maseeh College of Engineering and Computer Science; and N. McNeil, Center for Transportation Studies, Portland State University, P.O. Box 751, Portland, OR 97207. E. Stolz, Sprinkle Consulting, 1624 Market Street, Suite 202, Denver, CO 80202. Current affiliation for E. Stolz: Ready4Wellness, LLC, P.O. Box 22990, Denver, CO 80222. Corresponding author: K. Nordback, nordback@pdx.edu.

*Transportation Research Record: Journal of the Transportation Research Board, No. 2527,* Transportation Research Board, Washington, D.C., 2015, pp. 90–98.
DOI: 10.3141/2527-10

## BACKGROUND

This background section is divided into three topics: a description of some of the types of nonmotorized count data available, a summary of some of the main standard data formats, and a review of existing online nonmotorized traffic count archives. Data formats are the requirements for how the count data and associated metadata are to be organized. Because the majority of nonmotorized count data consists of pedestrian and bicyclist counts, this paper often refers to these modes in lieu of the term "nonmotorized."

### Duration and Types of Count Data

With a data archiving system, it is critical to understand how data that will populate the archiving system are gathered. This understanding is important because an archiving system's back-end functionality is dependent on how data suppliers will provide information. In an effort to understand the back-end data suppliers of nonmotorized traffic volume data, a review of the nonmotorized count duration and data types is offered below.

Traffic count data, in general, are diverse, and nonmotorized traffic count data are especially so because nonmotorized travel is more complex (e.g., pedestrians may cross a road or intersection in a variety of ways) and because these data have lacked an effective central organizing mandate. However, there are some important function-based distinctions that inform data classifications. Bicycle and pedestrian count data can be classified by method of collection into several categories.

#### *Classification by Duration and Data Collection Method*

First, manual counts are counts conducted by individuals. Generally they consist of an individual, often an agency staff member, volunteer, or intern, manually counting people walking or bicycling. These counts may be taken on location in the field or while reviewing video recordings or feeds on a computer. They may be taken by hand by jotting lines on a paper schematic of the count location, with a handheld electronic counting board, by using a smartphone, or by various other ways. Manual counts are often conducted over a short duration (e.g., 2 h) at infrequent intervals (e.g., once annually). Because they are short in duration, these counts are also biased by weather conditions, events, and weekly and seasonal variation.

Next, short-duration automated counts are distinct from manual counts in that they are collected by machine and typically have a longer duration than manual counts (24 h to multiple weeks). They are collected by mobile automated counters. Because they are short in duration, they are also subject to biases, such as seasonality and weather.

Finally, permanent automated counts may provide the best data about a location. Permanent automated counts are collected continuously, 24 h per day, 7 days per week, 365 days a year, by counters that are permanently installed at a location. Permanent automated counts do not need to be adjusted for seasonality and, in contrast with a 48-h count that could be affected by an unusual weather pattern, the long duration of permanent counts limits the impact of special events and unusual weather. However, they are still susceptible to under- or overcounting because of occlusion, improper setup, or other technical issues related to the specific technology, site, and traffic flow.

Each type of count may contain bias. Manual counts are often conducted by volunteers or otherwise non-full-time counters and may be subject to counter bias. Automated counters require appropriate installation and maintenance and need validation. While such counts are not subject to human bias, they may have substantial biases because of occlusion, improper installation or setup, and other systematic errors, depending on the specific technology, site, and traffic flow (*5*). However, they also allow for much longer-duration counts than are possible with manual counting and thus provide a valuable source of count data. With such count records, one can study hourly, daily, and monthly traffic patterns, which would be impossible to study with only manual counts.

#### *Intersection Versus Segment Counts*

Another important way to classify count data is by whether the count is collected at an intersection or on a road or path segment. Intersection counts are sometimes broken out into turning movements. Segment counts are also known as screen line counts because they count every bicyclist or pedestrian who crosses an imaginary line drawn perpendicular to the facility. Many manual counts are collected at intersections, and most automated counts are collected on segments. Recent developments in video image processing are making intersection counts feasible for longer durations.

Intersection counts can be much more complex than segment counts because of traffic turns at intersections. The paths of nonmotorized traffic flow at intersections is not as channelized as it is for motor vehicles. For example, bicyclists may either act as on-street vehicles or use crosswalks like pedestrians. This situation leads to a greater degree of complexity in nonmotorized intersection counts than is required for motor vehicle counts.

### Description of Standard Data Formats

This section provides an overview of current standard data formats. The project team reviewed these formats, in addition to the formats of data provided by partner agencies, to inform the database structure design and where possible to be compatible with other count archives. The goal is to create a robust data structure that can handle data from a variety of input sources, including manual and short-duration and permanent automated counts. For this reason, an understanding of existing standard data formats is important.

Nonmotorized count data formats vary by jurisdiction in the case of manual counts and by counter manufacturer in the case of automated counts. To standardize these data, there are at least three main ongoing efforts in the United States. The oldest is that offered by the NBPDP (*4*). The other two have only been established in the past 2 years: the Los Angeles, California, Bike Count Data Clearinghouse (*2*) and FHWA's *Traffic Monitoring Guide* (TMG) data format (*6*). The NBPDP and Los Angeles formats are designed for manual counts. The TMG format includes the ability to adapt to both manual and automated counts but is best adapted to automated count data.

This section begins with a discussion of the variety of raw count data and then details each of the three main efforts to standardize these data.

### Raw Data Formats

Despite efforts by the NBPDP and others to standardize manual counts, they are collected in many formats. The forms provided by NBPDP are often modified to suit the needs of local jurisdictions, which can lead to lack of compatibility both in the raw data themselves and in the formats in which the data are stored. For example, the NBPDP intersection form is specific for bicycle counts only, but some jurisdictions have modified it to collect pedestrian counts as well. Similarly, the NBPDP intersection form does not record gender or helmet use but has been modified by some jurisdictions to include these fields, adding to the complexity of the data produced.

For automated counts, the data formats used by different equipment manufacturers can vary and are often not interoperable. These raw data usually consist of a date and time followed by a count, or counts for directional counts or equipment that counts multiple locations or modes. However, even minor differences between data files require modifications to the loading script for each raw file type to be added to an archive. Once such a script is created, uploading future data can be automatic, unless the manufacturer changes the data output file format even slightly. This situation becomes more problematic with the more raw data formats that an archive supports and the more often manufacturers make format changes.

For this reason, standardized formats are desirable to archive data. If all data were always produced in one standard format, adding them to an archive would be significantly easier.

### National Bicycle and Pedestrian Documentation Project

The first effort on the national level to create a standard format for bicycle and pedestrian counts was the NBPDP. The NBPDP was initiated by a joint effort between ITE and Alta Planning and Design in 2004 as a response to the lack of available bicycle and pedestrian data for use in analysis, estimation, and other purposes (*7*). The NBPDP website provides standard forms, instructions, and other information for agencies interested in counting nonmotorized traffic. The NBPDP has helped and encouraged many jurisdictions around the nation to start bicycle and pedestrian counting programs.

The NBPDP accepts and stores data files submitted by e-mail to the project's administrators. It encourages submitters to use its standard data format for such submission but does not require it. The format includes contact information for the person responsible for data as well as data fields summarized in Table 1. The format asks for general information on the area in which the count is collected and count location–specific information as well as count data. While these fields would provide helpful metadata for those studying the area, they can sometimes be found in other databases and may not be readily accessible to the count data provider. This problem can lead to few data providers submitting data such as population density or number of visitors to an area.

While the NBPDP data collection methodology is meant to provide guidance on data collection methods, it does not address the need for electronically managing the data in an organized, standardized, easily accessible database and associated archiving system. Access to data collected using the NBPDP method is found by request only, can be paper format only, and does not yield an electronically efficient way to access data in a practice-ready format.

### Los Angeles County Bike Count Data Clearinghouse

Another effort to standardize and collect bicycle count data comes from the Los Angeles area. The University of California, Los Angeles, Luskin School of Public Affairs' Bike Count Data Clearinghouse project began in 2012 with the goal of housing bike volume data from the Los Angeles County region (*8*). The project is cosponsored by the Southern California Association of Governments and the Los Angeles County Metropolitan Transportation Authority. This data archive offers a user-friendly interface featuring a web-based geographic information system tool to make housed data accessible for use. Data are standardized for municipalities in Los Angeles County. To the authors' knowledge, this archive is the only publicly available, online bicycle count archive that also enables no-cost online data uploads from agencies within a region.

While the Los Angeles Data Clearinghouse provides access to data electronically, the project database structure is focused on handling primarily 2-h count data. With a lack of continuous count volume data or a minimum of 24 h of consecutively collected hourly traffic count data, conclusions about time of day, day of week, and travel volume trend patterns cannot be reached. Data handling and uploading of data are restricted, and data suppliers must first obtain approval to upload data to the system.

### Traffic Monitoring Guide

FHWA's TMG is "intended to provide the most up to date guidance to State highway agencies in the policies, standards, procedures, and equipment typically used in a traffic monitoring program" (*6*). Chapter 7 of the recently updated TMG gives instructions for coding and entering collected nonmotorized traffic count data in the TMG format. The TMG's main goal is to help states manage and improve their traffic monitoring programs, including all related business processes, technology, and equipment. Unlike the previously discussed data formats, the TMG format has precise requirements for the number and type of characters in each field in a data file.

The TMG format includes two types of data files: nonmotorized station description records and nonmotorized count records. Data fields in the station description include state and county codes, station identification code, functional classification of road (including two new categorize for trails and general area counts), and other specifics as listed in Table 1. The count record includes 24 h of data per record and optional weather information and repeats some of the same fields also included in the station description.

## Existing Online Nonmotorized Traffic Count Archives

There are a multitude of existing online archives of bicycle and pedestrian count data in the United States and abroad. They fall into four categories as described in the matrix in Figure 1 depending on if the source code is publicly or privately owned and if the data are available to the public or if access is restricted.

This review focuses on those archives that are both publicly available and public-agency owned with open-source code because they are most similar to the scope of this project. Examples are listed in Table 2. These archives are usually managed by local or regional

**TABLE 1 Summary of Data Fields Included in Standard Data Formats**

| Information Type | NBPDP | Los Angeles Bike Count Data | TMG |
|---|---|---|---|
| Identification | Location Description | Location ID<br>Data set name | Station ID |
| Locational | Land uses (1 to 2 mi)<br>Jurisdiction<br>Population density, bike–ped mode share, median age and income, number of visitors to area, type of setting, scenic quality, visitor destinations (1 mi) | Land use | <br>State, county |
| Route | Motor traffic volumes<br>Posted speed limit<br>Intersecting traffic volume<br>Crossing protection<br>Topography | Road class<br>Speed limit | Functional class, national highway, direction of route<br>Posted speed limit, route signing, route number<br>Intersection<br>Crosswalk |
| Facility | Facility type<br>Length of facility | Bikeway type<br>Type of other users | Exclusive facility<br>Sidewalk |
| Network | Connecting facility quality<br>Quality of network | None | None |
| Counter | None | None | Year established, year discontinued<br>Latitude and longitude<br>Type of sensor<br>LRS ID, LRS location point station location<br>Location relative to road |
| Count description | None | Count method<br>Direction | Count type (walk or cycle)<br>Direction of travel<br>Method of counting<br>Factor groups<br>Count purpose, notes |
| Temporal | Date<br>Time | Date<br>Day<br>Period<br>Interval begin | Year, month, and day<br>Count start time<br>Count interval (min) |
| Weather | Weather | Raining (yes or no) | Precipitation (yes or no)<br>High and low temperature |
| Count | Count of bicyclists<br>Count of pedestrians<br>Count of other nonmotorized traffic | Count of bicyclists<br>Female bicyclists<br>Sidewalk bicyclists<br>Wrong-way bicyclists<br>Count of other | Count per interval |

NOTE: ped = pedestrian; LRS = linear referencing system.

agencies who desire multiagency data sharing. With the exception of the Los Angeles Bike Count Data Clearinghouse, these sites do not allow users to upload data to the system but only allow users to view or download data (*2*). The project team found no states with an easily accessible public online bicycle and pedestrian count database. This finding illustrates, once again, the need for a nationally accessible system that can provide publicly available online nonmotorized traffic volume data.

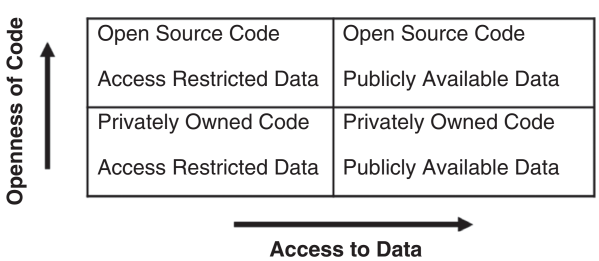


**FIGURE 1 Matrix of count archive types.**

The Travel Monitoring Analysis System operated by FHWA is an archive of motor vehicle traffic data, the new version of which is being designed to include nonmotorized traffic counts (*6*). While this database is operated by a public agency, it currently does not include nonmotorized count data and is not available to the public. For these reasons, it is not included in Table 2.

The NBPDP provides a standard data format and encourages participants to send in data files. However, it does not archive these data into one database and does not make these data available publicly. For this reason, it is not included in Table 2.

In addition to the publicly available data archives listed in Table 2, there are also many privately available data archives and online tools. Most of these data management products cater primarily to motor vehicle traffic, but some include bicycle and pedestrian traffic. Traffic count database products are available through detector manufacturers and traffic data software providers. These proprietary software

TABLE 2 Examples of Publicly Available Online Nonmotorized Traffic Count Archives

| | Data Types | | | | | |
|---|---|---|---|---|---|---|
| Agency | Duration | Automated? | Map? | Graph? | Allows Data Download | Other |
| Delaware Valley Regional Planning Commission (*9*) | 1 week | ✓ | ✓ | ✓ | ✓ | Includes weather |
| Arlington, Virginia (*10*) | Permanent | ✓ | ✓ | ✓ | ✓ | Includes weather |
| Portal (*11*) | Permanent | ✓ | ✓ | ✓ | | |
| Central Lane Metropolitan Planning Organization (*12*) | >24 h | ✓ | ✓ | | ✓ | Weather in output file, includes photos |
| Los Angeles Bike Count Data Clearinghouse (*8*) | 2 h | | ✓ | | ✓ | Allows data input |
| Seattle, Washington (*13*) | Permanent | ✓ | ✓ | ✓ | | Only two sites, but includes weather |

NOTE: ✓ = category included in agency archive; blank space = category not included in agency archive.

products provide data analysis tools and often produce reports, and some are able to export the data for use in other software.

## FUNCTIONAL REQUIREMENTS

One of the first steps in building any national archiving system or data warehouse is to develop functional requirements. Building this national archiving system has included obtaining funding for the basic creation of a bicycle and pedestrian count data clearinghouse. With this initial project, a basic nonmotorized traffic count clearinghouse will be established as part of Portal, an existing transportation data archive operated at Portland State University (*14*). The new archive is referred to as the "Bike–Ped Portal" in this paper. The basic functions that the Bike–Ped Portal is being designed to include are listed in Table 3 in the first phase of the work.

Managing the development of a data archiving system requires careful thought and prioritization to produce a working product within a given budget and time frame. While the vision for the project is to include all types of count data, to produce a working data archive within the budget and time frame of the project, which is to be completed in November 2015, some data types are prioritized over others.

TABLE 3 Basic Functional Requirements for Phase 1

| Requirement | Phase 1 Priority |
|---|---|
| Input tool | Provide user accounts so they can upload data.<br>Create translation tools and ability to upload continuous count files online.<br>User input tool for counts on road or path segments. |
| Quality assurance/quality control (QA/QC) tool | Automated check for completeness.<br>User validation based on review of plots and automated flags.<br>User fixes errors and resubmits. |
| Archive documentation and metadata | Flexible architecture to expand to future uses.<br>A set of required and optional fields for locations, detectors, and data records. |
| Output tool | Interactive map of count locations.<br>Data visualization with two basic graphs for a user-chosen time period.<br>Statistics: compute annual daily traffic if a full day of data are available and annual average daily traffic if a full year of data are available.<br>Export: allow data to be exported in table format or in TMG format.<br>Basic application programming interface. |
| Data to include | Prioritize segment count sites with at least 24 h of data. |

The priorities, listed below, were made on the basis of the recommendations of the project's technical advisory committee and their desire to focus on the most complete, quality-driven, and manageable data first:

- Data sets with 24 h of consecutively collected hourly counts or greater per location are a first priority. Shorter-duration counts are a second priority.
- Data collected on segments are a first priority. Data collected at intersections are a second priority.

Counts longer than 24 h are prioritized because they provide views of travel patterns over the course of the day that cannot be known with shorter-duration counts. Traffic statistics can be calculated with 24-h counts, and continuous count volume data can provide conclusions about time of day, day of week, month of year, and year-to-year travel volume trend patterns.

Counts at intersections are inherently more complex and are usually associated with counts collected for less than 24 h. For the first phase of work, if automated count data are collected at intersections, each approach to the intersection will be treated as a separate road or path segment. The ability to archive individual turning movements is a task left for future phases of work.

The archive will be publicly available through the Portal website and any interested agency will be encouraged to upload data. The first phase of work will support data input in a few specific data formats, the specifics of which are yet to be determined.

## ARCHIVE ARCHITECTURE

The diagram shown in Figure 2 illustrates the system architecture for the Bike–Ped Portal. The Bike–Ped Portal has two primary components: the Bike–Ped Portal web interface and the Bike–Ped Portal archive.

When building any archiving system, developing a strategic database architecture that is solid, sustainable, and maintainable is critical. For this project, a system architecture that provides front-end,

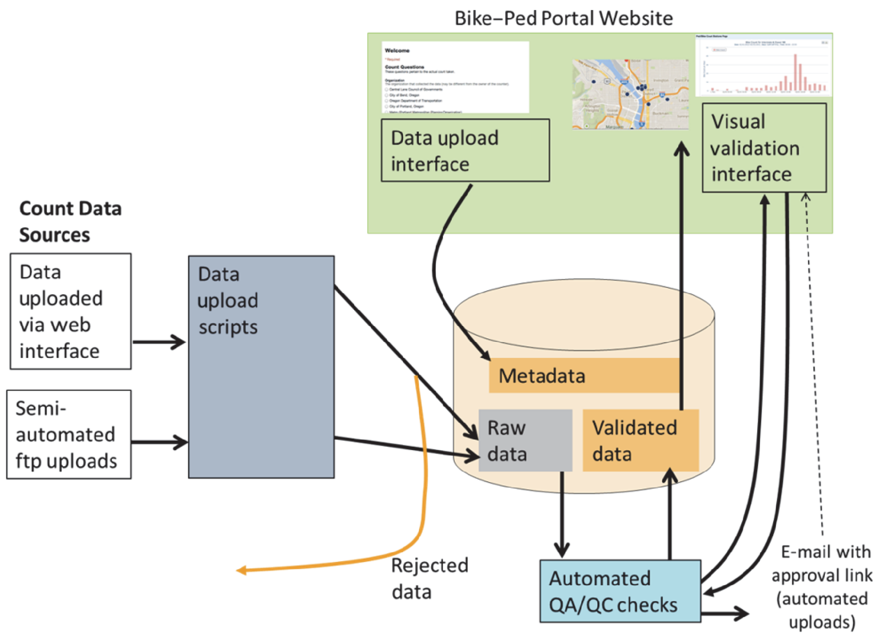


FIGURE 2 Archive architecture (ftp = file transfer protocol).

back-end, and middle-ware database functionality is required. The following are system architecture components with descriptions of the anticipated functionality.

### Data Sources and Upload

Count data will arrive into the archive from two primary sources: files uploaded by agencies through the web interface and files automatically uploaded to a file transfer protocol (ftp) site. The count data will be accompanied by metadata that describes the count segments and detectors. All metadata will be uploaded through the web interface. As part of the upload process, the data will be checked to ensure that it meets basic formatting criteria and "sanity checks," such as verifying that the start date of the file precedes the end date and that required data fields are not null. If the data pass these initial checks, the data are loaded into the "raw data" portion of the Bike–Ped Portal archive. If the data do not pass the check and the data are rejected, the user will be notified either that the data have passed the initial checks and the quality assurance/quality control (QA/QC) process has begun or that the data have been rejected.

### Bike–Ped Archive

The Bike–Ped Portal uses a PostgreSQL database that is shared with the existing Portal traffic data archive. The archive stores both the raw uploaded data and the validated data.

### Data Quality Process

The data uploaded to the Bike–Ped Portal will be checked through a QA/QC process. The diagram in Figure 2 shows some important features of the QA/QC process. The QA/QC process will identify suspect data on the basis of simple flags, such as counts that are unusually high for a given hour or day and unusually high numbers of consecutive identical counts. Data will arrive to the archive either through the web interface or through automated upload to an ftp site. In either case, the user will be notified of the suspect data.

The notification will be either directly through the web-upload process for data being uploaded through the website or through an e-mail for data that are uploaded to the ftp site. The process for validation will be the same for data sets uploaded through the website or through the semiautomatic ftp upload. In either case, the user will be asked to investigate and validate suspicious data. The user will be provided with information about the data quality tests that fail and will be able to view simple plots of the data.

The user will be given the ability to add notes to the data. Each note will be associated with a count detector and a range of time. The notes will give the user a chance to record information about events or other things that affect the counts; particularly of interest are observations that may not be available from other sources. After the user has reviewed the data and QA/QC information, the count data will be accepted into the validated data in the archive or may be included but marked as invalid data, which will be hidden from public use.

### Bike–Ped Portal Web Interface

The Bike–Ped Portal will have a web interface that supports data upload and the QA/QC process and also provides plots and maps of the data. The plots and maps in the web interface query the validated data in the archive.

## ARCHIVE DATA STRUCTURE

As with all data archiving systems, the archive data structure describes how the metadata and validated data are stored within the archive. The data are stored with an eye toward minimizing redundancy while preserving as much detail as available in the raw data. The data structure was designed to allow easy access to the data for querying and easy and efficient future data processing.

As shown in Figure 3, the data structure consists of the following basic elements: segment areas, detectors, facilities, flows, count descriptors, and the data records themselves. While the specific fields included in these basic elements may be adjusted, the elements themselves and their relationship to one another have been finalized and represent a novel and versatile approach to archiving bicycle and pedestrian count data. Each element is described below.

The segment area represents the section of roadway or path on which the count is collected, including all associated transportation facilities. In the upper portion of Figure 3, the largest rounded rectangle represents the segment area. For example, if the count was conducted on a bridge, the segment area would include the entire bridge, roadway, sidewalks, and paths. This area will be input by the user and used for future efforts to combine count data with various line-based data sets. The segment area has high-level attributes, such as name, state, county, and observed land use. In addition, TMG attributes were included in the data structure for the purpose of exporting to TMG format. Finally, for the geographic attribute, segment area is spatially represented as a polygon in the archive.

In this data structure, the detector element represents the device that is used to collect counts. A detector may be a pneumatic tube for bicycle counts, an inductive loop bicycle detector, an infrared device, or a person. The detectors are shown as large dots in Figure 3. Multiple detectors may to be associated with one segment area. The attributes of the detector device include a description and information about the device, such as if it is automated or permanent and its make, model, and serial number. As with the segment area, TMG fields have been included to support output to TMG format.

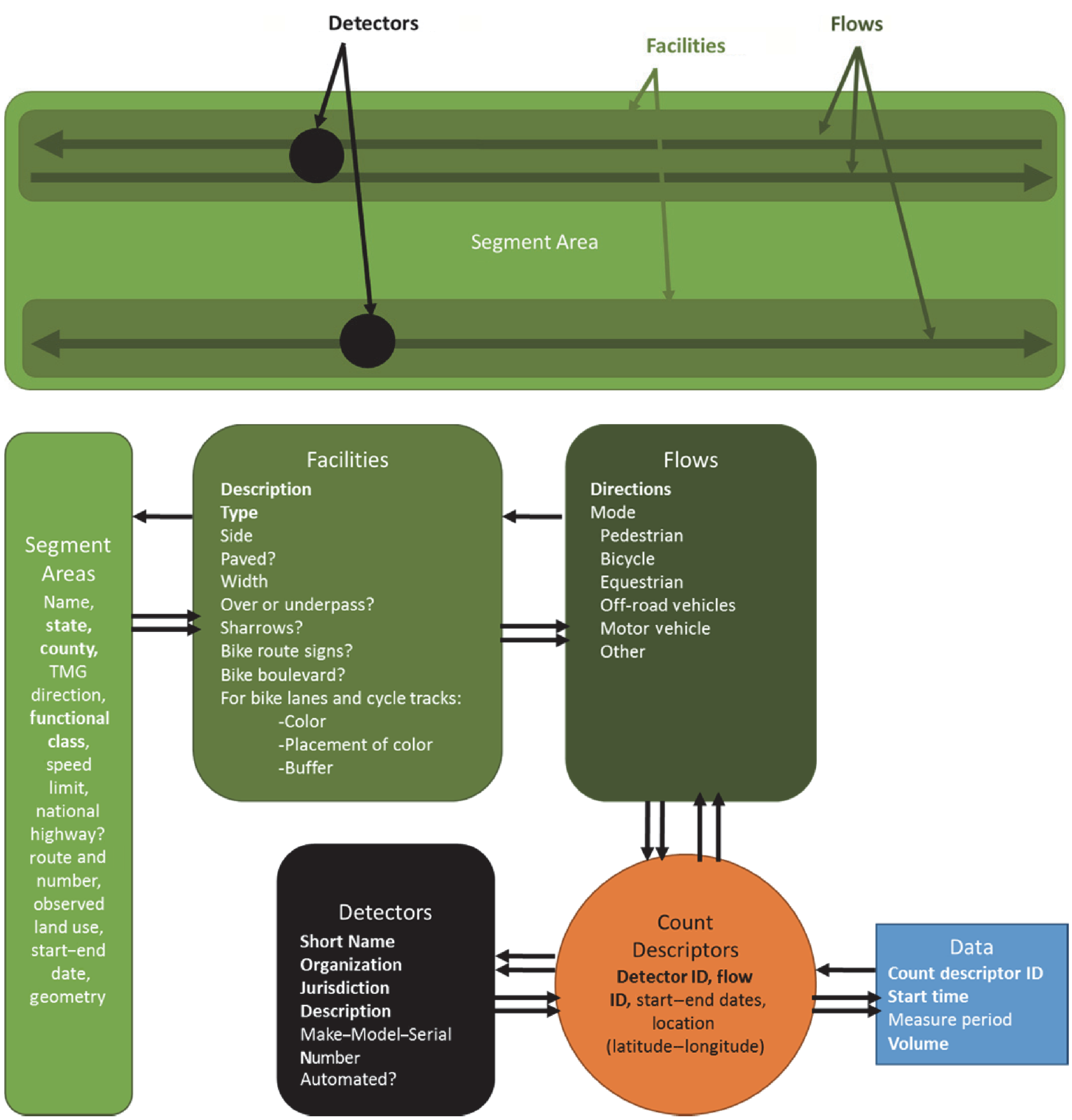


FIGURE 3 Archive data structure.

The facility represents the facility on which traffic is being counted. For example, a facility might represent the north sidewalk of a roadway or a bicycle lane on the roadway. Facilities are shown as two boxes, one for each sidewalk on either side of the roadway represented by the segment area in Figure 3. Facility data include information describing the type of facility, its width, and pavement type.

The flow represents the traffic flow that is being counted. A single detector may count multiple flows. For example, a single detector may count both bicycles and pedestrians, or a single detector may count both northbound and southbound traffic. Thus there are typically multiple flows associated with each detector. There can also be multiple detectors associated with a given flow. The flow includes the travel direction and mode of travel.

To accommodate multiple detectors for a given flow and multiple flows for a given detector, a table was created to link a given detector to a given flow, called the count descriptor. The count descriptor table also includes the location of the detector, which is represented as a point.

The final item in the data structure is the data records themselves. These records contain simply the time interval of the count and the counts recorded in that time period. A file to be uploaded to the site is likely to contain many data records.

The data structure is currently focused on counts collected on road or path segments. However, in the future phases of work, the team plans to expand the basic data structure to include intersection-specific data. This plan can be accomplished by making modifications to the segment area to define an intersection area instead of a segment, minor modifications to detector, and expansion of flows and facilities to include specific movements and crosswalk counts. These modifications illustrate the inherent flexibility in the archive data structure.

## CONCLUSION

The project team has taken on the development of a national nonmotorized traffic volume data archiving system. Many agencies have developed archiving systems, but none have attempted to provide public access to a national standardized nonmotorized data set. With this being the first attempt, much will be learned during implementation that offers agencies access to a practice-ready nonmotorized data set.

This paper provides a survey of the existing nonmotorized count data landscape and proposes a design to create a national online nonmotorized traffic count archive. The review of existing count types, formats, and publicly available archives reveals that, though there are many archives, none are both national in scope and enable data to be directly uploaded to the site. The archive outlined in this paper will fill that gap and provide a needed forum for sharing nonmotorized count data beyond jurisdictional boundaries.

This paper details how such an archive would be created, including the basic architecture, features, and database structure. This archive will include online input, data quality evaluation, data visualization functions, and the ability to download user-specified data and exchange the data with other archives and applications. The archive's versatile data structure allows for both mobile counters and validation counts of the same traffic flow, an innovation in design that greatly expands the usefulness of the archive while minimizing data storage requirements.

The archive will enable data sharing such that the data collected by one agency can be shared with many others and not be lost. This archive would provide safety researchers with a measure of exposure, allow local and regional agencies to share count data and estimate daily nonmotorized traffic, and provide policy makers and transportation professionals with basic information on cycling and walking to inform decisions and plans around the country.

## NEXT STEPS

The next steps for the project may include improving the user interface, more refined data quality checking processes, and expanding tools to compute nonmotorized annual average daily traffic (AADT). The existing scope only allows nonmotorized AADT values to be computed if sufficient data are available to use the AASHTO method (*15*). However, if seasonal adjustment factors were computed using the permanent count data available in the archive, they could be applied to short-duration counts to estimate AADT values at additional locations. The project team hopes to include development of such functionality in future phases of this work. This development might include further investigation of appropriate techniques for automated grouping of count stations as well as resolving issues of computing seasonal factors specific to the data included in the archive.

Key to the success of the archive as a national resource will be encouraging its adoption by both data suppliers and data users. To this end, the project team will update TRB on its progress through the Bicycle and Pedestrian Data Subcommittee and will communicate with a set of national stakeholders. The success of this effort relies on the jurisdictions around the country finding it useful and continuing to fund it.

## ACKNOWLEDGMENTS

Many people have contributed to make this work possible. The authors are grateful for the work of Susan Payne of the Lane Council of Governments (LCOG), Eugene, Oregon; Jon Makler and Tara Weidner of the Oregon Department of Transportation (DOT); Marjorie Bradway of the City of Portland; and Jennifer Dill, Hau Hagedorn, and Susan Peithman of the National Institute for Transportation and Community (NITC), who got the project off the ground. The authors also acknowledge the support of the Modeling Program Coordination (MPC) subcommittee of the Oregon Modeling Steering Committee (OMSC), which included representatives from all metropolitan planning organizations (MPOs) within Oregon, including Rogue Valley Council of Governments (Medford, Oregon), Corvallis Area MPO (Corvallis, Oregon), and Mid-Willamette Valley Council of Governments (Salem, Oregon). The authors are especially grateful to the project partners: NITC; the Oregon DOT; LCOG (Eugene, Oregon); Metro (Portland, Oregon); Bend MPO (Bend, Oregon); the cities of Boulder, Colorado, Austin, Texas, and Bend, Oregon; the Oregon Community Foundation; and FHWA, Office of Planning. The authors are also very grateful for the outstanding technical advisory committee members: Susan Payne and Josh Roll of LCOG, Aleksiina Chapman of Austin, Don Crownover and Sheila Lyons of the Oregon DOT, Jeremy Raw of FHWA, Jovi Anderson of Bend MPO, Mike Gardner-Sweeney of Boulder, Lake McTighe and John Mermin of Metro, and Hau Hagedorn of NITC. The authors also thank IBM and the Saturday Academy for funding Jolene Liu's internship.

## REFERENCES


1. Nordback, K. *Online Non-Motorized Traffic Count Archive.* 2014. http://otrec.us/project/817. Accessed April 23, 2014.
2. Huff, H., and M. Brozen. Creating the Bicycle Count Data Clearinghouse for Los Angeles County, California. Presented at 93rd Annual Meeting of the Transportation Research Board, Washington, D.C., 2014.
3. Nordback, K., W. E. Marshall, and B. N. Janson. *Development of Estimation Methodology for Bicycle and Pedestrian Volumes Based on Existing Counts.* Colorado Department of Transportation, Denver, 2013.
4. National Bicycle and Pedestrian Documentation Project. 2009. bikepeddocumentation.org. Accessed Nov. 22, 2011.
5. Ryus, P., F. R. Proulx, R. J. Schneider, T. Hull, and L. Miranda-Moreno. *NCHRP Web-Only Document 205: Methods and Technologies for Pedestrian and Bicycle Volume Data Collection.* Transportation Research Board of the National Academies, Washington, D.C., 2015.
6. *Traffic Monitoring Guide.* FHWA, U.S. Department of Transportation, Washington, D.C., 2013.
7. National Bicycle and Pedestrian Documentation Project. *Fact Sheet and Status Report.* 2014. http://bikepeddocumentation.org/downloads/. Accessed July 22, 2014.
8. Lewis Center. *SCAG Bicycle Data Clearinghouse Project.* http://www.lewis.ucla.edu/2012/07/scag-bicycle-data-clearinghouse-project/. 2014. Accessed July 22, 2014.
9. Delaware Valley Regional Planning Commission. *Pedestrian and Bicycle Counts.* 2014. http://www.dvrpc.org/webmaps/pedbikecounts/. Accessed Aug. 1, 2014.
10. Arlington, B. *Bicycle and Pedestrian Counters.* 2014. http://www.bikearlington.com/pages/biking-in-arlington/counter-dashboard/. Accessed Aug. 1, 2014.
11. Ped/Bike Count Stations Page. *PORTAL.* 2014. http://demo.portal.its.pdx.edu/Portal/index.php/pedbike. Accessed Aug. 1, 2014.
12. Central Lane Metropolitan Planning Organization. *Regional Bicycle Count Program.* http://www.thempo.org/what_we_do/bicycle_counts_web_a.cfm. Accessed Aug. 1, 2014.
13. Seattle Department of Transportation. *Seattle's First Bicycle Counter.* 2014. http://www.seattle.gov/transportation/bikecounter_fremont.htm#detail. Accessed Aug. 1, 2014.
14. *What Is Portal?* 2013. http://portal.its.pdx.edu/newsandinfo/. Accessed May 15, 2013.
15. *Guidelines for Traffic Data Programs.* Joint Task Force on Traffic Monitoring Standards of the AASHTO Highway Subcommittee on Traffic Engineering, Washington, D.C., 1992.


*The Standing Committee on Highway Traffic Monitoring peer-reviewed this paper.*