Roadway Data Improvement Program: Supplemental Information Resource

RDIP

Roadway Data Improvement Program

June 13, 2012

U.S. Department of Transportation

Federal Highway Administration

Notice

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The U.S. Government assumes no liability for the use of the information contained in this document.

The U.S. Government does not endorse products or manufacturers. Trademarks or manufacturers' names appear in this report only because they are considered essential to the objective of the document.

Quality Assurance Statement

The Federal Highway Administration (FHWA) provides high-quality information to serve Government, industry, and the public in a manner that promotes public understanding. Standards and policies are used to ensure and maximize the quality, objectivity, utility, and integrity of its information. FHWA periodically reviews quality issues and adjusts its programs and processes to ensure continuous quality improvement.

ACKNOWLEDGEMENTS

The authors of this document were Ms. Barbara Hilger DeLucia and Dr. Robert Scopatz of Data Nexus Inc (DNI), supported by Ms. Nancy Lefler of Vanasse Hangen Brustlin, Inc (VHB) under contract DTFH61-10-D-00022. Mr. Robert Pollack was the Federal Highway Administration (FHWA) Contract Manager for this project.

The project team would like to thank the following individuals for their participation in the Roadway Data Improvement Program (RDIP) Noteworthy Practices Working Group Meeting:

- Peggi Knight, Iowa Department of Transportation
- Brian Mayhew, North Carolina Department of Transportation
- Brian Murphy, North Carolina Department of Transportation
- John Arnold, North Carolina Department of Transportation
- Jon Farley, North Carolina Department of Transportation
- Chris Harris, Tennessee Department of Transportation
- In-Kyu Lim, Virginia Department of Transportation
- Ralph Jones, Virginia Department of Transportation
- James Hall, University of Illinois-Springfield, retired Illinois Department of Transportation
- Dale Lighthizer, Consultant, retired Michigan Department of Transportation
- David Harkey, University of North Carolina Highway Safety Research Center

The project team would also like to thank the FHWA RDIP Advisory Group:

- Joseph Glinski, Resource Center
- Stephan Gaj, Office of Asset Management
- Joseph Hausman, Office of Highway Policy Information
- Rick Drumm, Indiana Division Office
- Lincoln Cobb, Office of Safety Research and Development
- Brian Gardner, Office of Planning
- Heather Rothenberg, Office of Safety

Their contributions helped shape the guidance that States will be receiving to improve their roadway data to help support data driven safety decisions.

TECHNICAL REPORT DOCUMENTATION FORM

1. Report No.	**2. Government Accession No.**	**3. Recipient's Catalog No.**	
4. Title and Subtitle Roadway Data Improvement Program: Informational Resource		**5. Report Date** June 2012	
		6. Performing Organization Code	
7. Author(s) Barbara Hilger DeLucia, Robert A. Scopatz, Nancy Lefler		**8. Performing Organization Report No.**	
9. Performing Organization Name and Address Vanasse Hangen Brustlin, Inc. 8300 Boone Boulevard Suite 700 Vienna, VA 22182		**10. Work Unit No.**	
		11. Contract or Grant No. DTFH61-10-D-00022	
12. Sponsoring Agency Name and Address Federal Highway Administration Office of Safety 1200 New Jersey Ave., SE Washington, DC 20590		**13. Type of Report and Period** Final Report, June 2011 - August 2012	
		14. Sponsoring Agency Code FHWA	
15. Supplementary Notes The contract manager for this report was Robert Pollack.			
16. Abstract			
17. Key Words: benefit cost analysis, safety data, roadway inventory data, traffic data, data collection, data quality		**18. Distribution Statement** No restrictions.	
19. Security Classif. (of this report) Unclassified	**20. Security Classif. (of this page)** Unclassified	**21. No. of Pages:** 106	**22. Price** N/A

Form DOT F 1700.7 (8-72) Reproduction of completed pages authorized

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
		NOTE: volumes greater than 1000 L shall be shown in m^3		
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
°F	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	°C
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact degrees)		
°C	Celsius	1.8C+32	Fahrenheit	°F
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380. (Revised March 2003)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONYMS

4 Es	Engineering, Education, Enforcement, Emergency Medical Services
AADT	Annual Average Daily Traffic
AASHTO	American Association of State Highway and Transportation Officials
ADT	Average Daily Traffic
AH	Milepost Ahead (denotes the original mileposts on a realigned roadway)
ALDOT	Alabama Department of Transportation
ANSI	American National Standards Institute
BIKESAFE	Bicycle Countermeasure Selection System
BK	Milepost Back (refers to new mileposts on a realigned roadway)
C&P	Condition and Performance
Caltrans	California Department of Transportation
CDIP	Crash Data Improvement Program
CFR	Code of Federal Regulations
CHP	California Highway Patrol
CMF	Crash Modification Factor
CODES	Crash Outcome Data Evaluation System
CRF	Critical Rate Factor
CSD	Context Sensitive Design
CSS	Context Sensitive Solution
DFD	Data Flow Diagram
DGPS	Differential Global Positioning System
DHMS	Digital Highway Measurement System
DOT	Department of Transportation
EMS	Emergency Medical Services
FARS	Fatality Analysis Reporting System
FDE	Fundamental Roadway and Traffic Data Elements
FHWA	Federal Highway Administration
FRA	Federal Railroad Administration
GIS	Geographic Information System
GPS	Global Positioning System
GUI	Graphical User Interface
HAL	High Accident Location

HD	High Definition
HOV	High Occupancy Vehicle
HPMS	Highway Performance Monitoring System
HRGX	Highway-Railway Grade Crossing
HRRR	High Risk Rural Roads
HSIP	Highway Safety Improvement Program
HSIS	Highway Safety Information System
HSM	Highway Safety Manual
IHSDM	Interactive Highway Safety Design Model
ISAT	Interchange Safety Analysis Tool
IT	Information Technology
LiDAR	Light Detection And Ranging
LRM	Linear Referencing Method
LRS	Linear Referencing System
MAP-21	Moving Ahead for Progress in the 21st Century
MDEV	Million Daily Entering Vehicles
MIDRIS	Model Impaired Driving Records Information System
MIRE	Model Inventory of Roadway Elements
MIS	Management Information System
MIRE	Model Inventory of Roadway Elements
MMUCC	Model Minimum Uniform Crash Criteria
MOU	Memoranda of Understanding
MPO	Metropolitan Planning Organization
MVM	Million Vehicle Miles Traveled
NBI	National Bridge Inventory
NCHRP	National Cooperative Highway Research Program
NEMSIS	National Emergency Medical Services Information System
NHPN	National Highway Planning Network
NHS	National Highway System
NHTSA	National Highway Traffic Safety Administration
NS	Non-regulatory Supplement to the U.S. Code
PBCAT	Pedestrian and Bicycle Crash Analysis Tool
PDA	Personal Digital Assistant
PEDSAFE	Pedestrian Safety Guide and Countermeasure Selection System

PMS	Pavement Management Systems
RDIP	Roadway Data Improvement Program
RI	Roadway Inventory
RPC	Regional Planning Commission
RSDP	Roadway Safety Data Program
SAFETEA-LU	Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users
SHSP	Strategic Highway Safety Plan
SPCS	State Plane Coordinate System
SSAM	Surrogate Safety Assessment Model
STIP	State Transportation Improvement Program
TASAS	Traffic Accident Surveillance and Analysis System
TAT	Technical Assistance Team
TFHRC	Turner-Fairbank Highway Research Center
TRCC	Traffic Records Coordinating Committee
TRS	Traffic Records Systems
U.S. DOT	United States Department of Transportation
USC	United States Code (as in Title 23 USC 109)
VHS	Video Home System (Analog Recording Videotape-Based Cassette)
VIN	Vehicle Identification Number
VMT	Vehicle Miles Traveled

CHAPTER 1—INTRODUCTION TO THE ROADWAY DATA IMPROVEMENT PROGRAM (RDIP)

PURPOSE OF THE RDIP

The purpose of the Roadway Data Improvement Program (RDIP) is to help transportation agencies improve the quality of their roadway data to support their safety initiatives. These improvements may be in terms of the data elements collected, data collection practices, geo-spatial data referencing, data storage, data maintenance, and linkage of roadway-related data with other safety data.

RDIP provides roadway database managers and other traffic safety professionals a tool to assist them in identifying, defining, measuring, and ultimately improving, the quality of the data within their roadway databases. The quality of the data can be characterized by the timeliness, accuracy, completeness, consistency, integration, and accessibility of the roadway data.

Purpose of the RDIP Supplemental Information Resource

The intent of the RDIP Supplemental Information Resource is to provide an overview of data collection practices, transportation planning and coordination, data management, and using roadway data for safety. The information provided in this Supplemental Information Resource is a result of a thorough review of current and emerging roadway data practices, and incorporates information provided by a Technical Working Group made up of representatives from States on the forefront of roadway data collection, analysis, management, and integration. These States included Illinois, Iowa, Michigan, North Carolina, Tennessee, and Virginia.

Intended Audience

The intended audience of this Supplemental Information Resource is any State or local department of transportation (DOT) professional interested in improving the roadway and traffic data in their agency. This includes data collectors, managers, and users from safety, planning, roadway inventory, traffic, asset management, operations, maintenance, pavement, information technology (IT), Traffic Records Coordinating Committees (TRCC) members, and others. While the primary focus is on improving roadway data for use in safety, improving the quality of data can benefit all involved in the collection, management, and use of roadway-related data.

BACKGROUND

Highway safety analysis is evolving, and the importance of quality data has never been more apparent. Quality safety data are the foundation for highway safety decisions. Much of the effort in the safety community in previous years has concentrated on crash data; however, crash data are only part of the picture. Roadway and traffic data are also essential. By incorporating roadway and traffic data into their network screening, prioritization, and countermeasure selection analysis, agencies can better identify safety problems and prescribe solutions to improve safety and make more efficient and effective use of their safety funds.

Crash data alone are useful, but leave safety practitioners with purely reactive approaches—identifying the locations where crashes have already happened. With the addition of traffic volume data it is possible to develop estimates of the expected crash frequency and compare crash rates for roadways with vastly different levels of service. As detailed roadway inventory information is added to the mix, safety practitioners can now develop a more in-depth understanding of the roadway attributes that contribute to crash risk thus allowing them to adopt a proactive approach seeking out those factors associated with a high risk of crashes and addressing sites that share those features in common. This process is known as the "systemic approach" to safety.

In addition to numerous safety analysis tools and methods developed by the States, a new generation of tools is being developed to help States to identify safety issues and provide recommendations for improvements. These tools include the 2010 Highway Safety Manual (HSM), the Interactive Highway Safety Design Model (IHSDM), *Safety Analyst*, as well as American Association of State Highway and Transportation Officials' (AASHTO) National Cooperative Highway Research Program (NCHRP) Series 500 Data and Analysis Guide, which all require crash, roadway, and traffic data to achieve the most accurate results.

The Federal Highway Administration (FHWA) has been on the forefront of the effort to improve the quality of safety data, including roadway and crash data. In 2007, FHWA released the initial Model Minimum Inventory of Roadway Elements (MMIRE), a recommended listing of roadway and traffic data elements crucial for States and local agencies to collect and incorporate into their safety programs. A recently released version, Model Inventory of Roadway Elements - MIRE Version 1.0, shows FHWA's continued commitment to the evolution of quality safety data, see http://safety.fhwa.dot.gov/rsdp/.

As part of this evolution, the FHWA Office of Safety developed the Roadway Safety Data Program (RSDP) to advance State and local safety data systems and expand safety data analysis and evaluation capabilities. The RSDP includes data guidance, resource development, technical assistance and training all aimed at improving the collection, analysis, management, and expansion of roadway data for use in safety programs and decision-making. Technical assistance programs that fall under the RSDP umbrella, in addition to

RDIP, include the Roadway Data Extraction Technical Assistance Program (RDETAP), and customized Technical Assistance. More information about these programs may be found at http://safety.fhwa.dot.gov/rsdp/.

FHWA patterned the RDIP after the Crash Data Improvement Program (CDIP), which it initiated to help States improve the quality of their crash data. The CDIP teaches a State about methods to improve their crash data. The CDIP is now operated by the National Highway Traffic Safety Administration (NHTSA). Similar to CDIP, the RDIP concentrates on providing recommendations for improving the quality of their roadway-related data. The RDIP also provides examples of State and Federal methods and tools for safety analyses to illustrate to practitioners the necessity of quality roadway and crash data.

In 2012, FHWA completed the Roadway Safety Data Capabilities Assessment, which documented the status of State roadway data. The assessment allowed FHWA to identify national gaps and focus resources accordingly, see *http://safety.fhwa.dot.gov/rsdp/*. The RDIP builds upon the findings from the Capabilities Assessment. It conducts more in-depth examinations of state practices and procedures surrounding the main Capabilities Assessment topics.

These recent efforts at FHWA build on a strong history at the State and federal levels of developing the data necessary for data-driven decision making for safety. The following sections provide further background and historical context for the present efforts.

Traffic Safety Databases

At the time of passage of the Highway Safety Act of 1966, State centralized Traffic Records Systems (TRS) generally contained basic files on crashes, drivers, vehicles, and roadways. Some States added data on traffic safety-related education, either as a separate file or as a subset of the driver file. As traffic safety programs matured, many States incorporated emergency medical services (EMS) and citation/conviction files for use in safety program decision making. Additionally, some States and localities maintain a safety management file that consists of summary data from the central files that can be used for problem identification and safety planning.

As the capabilities of computer hardware and software systems increased and the availability of powerful systems expanded to the local level, many States adopted a more distributed model of data processing. For this reason, the model of a TRS needs to incorporate a view of information and information flow, as opposed to focusing only on the files in which that information resides.

Under this more distributed model, it does not matter whether data for a given system component are housed in a single database on a single computer, or spread throughout the State on multiple local systems. What matters is whether the information is available to users, in a form they can use, and that these data are of sufficient quality to support its intended uses. Figure 1 shows the major data components of a TRS that include:

- Crash Information.
- Roadway Information.
- Exposure (traffic volumes).
- Driver Information.
- Vehicle Information.
- Citation/Adjudication Information.
- Injury Surveillance Information.

Figure 1. TRS Components.

Together, these components provide information about places, property, and people involved in crashes and about the factors that may have contributed to the crash or traffic stop. The TRS also contains information to help judge the relative magnitude of problems identified through analyses designed to account for differences in exposure (normalization), cost effectiveness, and performance level data to support countermeasure management.

Figure 2 is a data flow diagram (DFD) that provides an overview of the potential sources of these data and data flows. The DFD provides an example of the possible interactions among all of the TRS data sources and serves as a first step in designing a system to integrate the various data sources. See Appendix A for a key to interpreting the DFD.

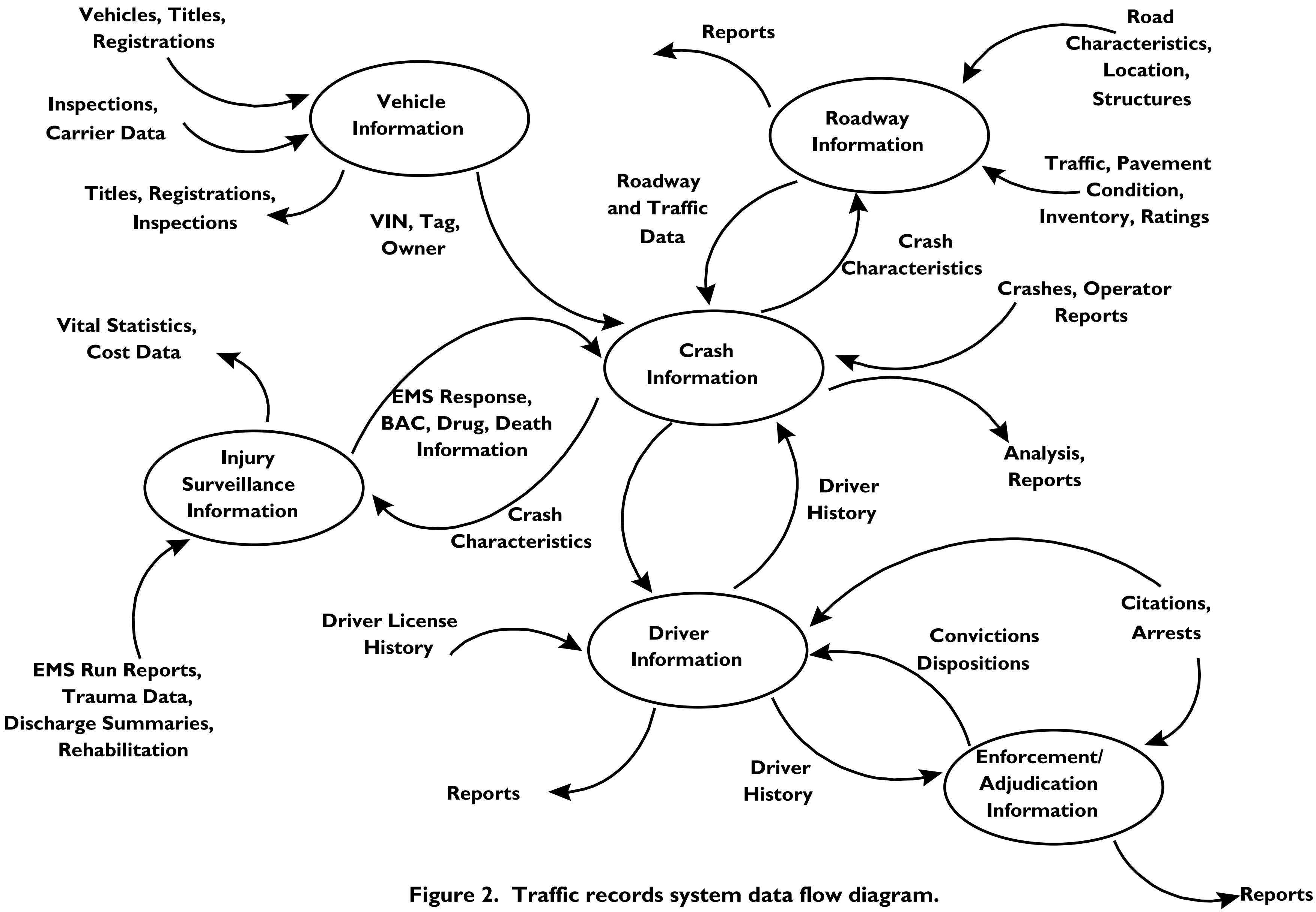


Figure 2. Traffic records system data flow diagram.

In each of the major components of the TRS, there is an obvious need for aggregate data to "roll up" from local, to State, to national levels. Agencies meet these needs at the State level by imposing standards on data collection. The Federal level was slower to evolve such standardization, but there are now recommended guidelines for crash data (the Model Minimum Uniform Crash Criteria—MMUCC), impaired driver tracking system (the Model Impaired Driving Records Information System—MIDRIS), and roadway data (the MIRE guideline already referenced). Further, there are national standards for fatal crashes (the Fatality Analysis Reporting System—FARS), commercial motor vehicle crashes (the SAFETYNET system), and EMS run reports (the National Emergency Medical Services Information System—NEMSIS).

There have also been advances in how we approach safety analysis and approach the problem of reducing the frequency and severity of crashes. The Haddon Matrix, named after its developer William Haddon, the first Administrator of the National Highway Traffic Safety Administration (NHTSA), provides an overview of the contents of a TRS. It provides a valuable framework for viewing the primary effects of human, vehicle, and environmental factors and their influence before, during, and after a crash event. Table 1 shows an expanded Haddon Matrix that was developed by R. Quinn Brackett, Ph.D. and staff from Data Nexus, Inc., for the original Traffic Records Assessment reports starting in the early 1990s. This matrix has been incorporated in numerous safety-related publications since that time.

The Haddon Matrix provides a meaningful way to examine primary effects of contributing factors on crash frequency and severity. It helps decision makers consider countermeasures designed to address specific contributing factors. In recent years, with availability of more detailed data analyses, awareness has grown about the interactions among contributing factors. For example, the factors that exist prior to a crash may result in a crash while post-crash measures may affect time required to transport injured parties to a medical facility for treatment.

Table 1. Expanded Haddon Matrix with example highway safety categories.

	Human	Vehicle	Environment
Pre-Crash	· Age · Gender · Experience · Alcohol/Drugs · Physiological Condition · Psychological Condition · Familiarity with Road & Vehicle · Distraction · Conviction & Crash History · License Status · Speed	· Crash Avoidance · Vehicle Type · Size & Weight · Safety Condition, Defects · Brakes · Tires · Vehicle Age · Safety Features Installed · Registration	· Visibility · Weather/Season · Lighting · Divided Highways · Signalization · Geographic Location · Roadway Class, Surface, Cross-Section, Alignment, etc. · Structures · Traffic Control Devices, Signs, Delineations, and Markings · Roadside Appurtenances, Buildups, Driveways, etc. · Volume of Traffic · Work Zone · Animal Range Land & Seasonal Movements
Crash	· Belt Use · Human Tolerance · Size · Seating Position · Helmet Use	· Crash-Worthiness · Passenger Restraints · Airbags and Airbag Shutoff	· Guardrails · Median Barriers · Breakaway Posts · Rumble Strips and Other Safety Devices · Maintenance Status of Roadway and Devices
Post-Crash	· Age · Physical Condition · Insurance Status · Access to Health Care · Driver Control Actions · Court Actions · Probation	· Post-Crash Fires · Fuel Leakage · Power Cell Securement · Hazardous Materials · Title	· Traffic Management · Bystander Care · EMS System · First Responders · Hospital Treatment · Long-Term Rehabilitation

The Haddon Matrix remains the pre-eminent summary of the factors that contribute to the frequency and severity of crashes. It also must be recognized that the factors interact. Modern safety analysis looks at all the factors, and their interactions, seeking to find not *the single* cause of any individual crash, but the multiple factors that combine and interact to produce the pattern of crashes and crash outcomes that we see.

The tools and strategies for safety analysis have been growing both in number and sophistication since Haddon's pioneering days. The HSM, and the associated tools and resources including *Safety Analyst*, the IHSDM, the Crash Modification Factors (CMF) Clearinghouse, summarize these advances. These advanced methods and tools represent the current state of the practice arrived at through decades of research and implementation by States and the federal government. They are covered in greater detail later in this document.

CHAPTER 2—THE ROADWAY DATA COMPONENT

The roadway data component of a TRS refers to the collection, storage, and use of various databases that contain information about how an agency constructed and maintained the roadway throughout its lifecycle. In the more specific area of traffic records and safety analyses, it is crucial to know the location where a crash occurs and any roadway and traffic characteristics that may have contributed to the crash. As part of a traffic records system, roadway data helps to identify the environment where motor vehicle crashes occur, any characteristics of the location that may have contributed to the occurrence or severity of the crashes, and suggest where engineering and/or enforcement may be changed to improve the safety of that location. In advanced safety analysis, the knowledge of where crashes have *not* occurred is also crucial to our ability to identify those roadway characteristics that are associated with increased crash risk. There must be a method of linking roadway data with crash data to support analyses, research, and public policy decisions.

> ***The roadway data component is a means by which States can describe the location and physical characteristics of roadways where crashes occur.***

Roadway data is also essential for agencies to be able to use the systemic approach to safety. The systemic approach involves implementing improvements based on the identification of high-risk roadway features correlated with specific crash types. This approach helps agencies broaden their safety efforts and consider crash risk as well as crash history when making determinations for safety improvements.

Figure 3 illustrates the sources of roadway data for the TRS and how these sources could potentially be linked to form a comprehensive roadway data management system. The roadway DFD is an expansion of the roadway process shown as one of the integrated processes in the overview DFD of Figure 2. This roadway DFD provides greater detail of the sources of roadway-related information and the interactions among them.

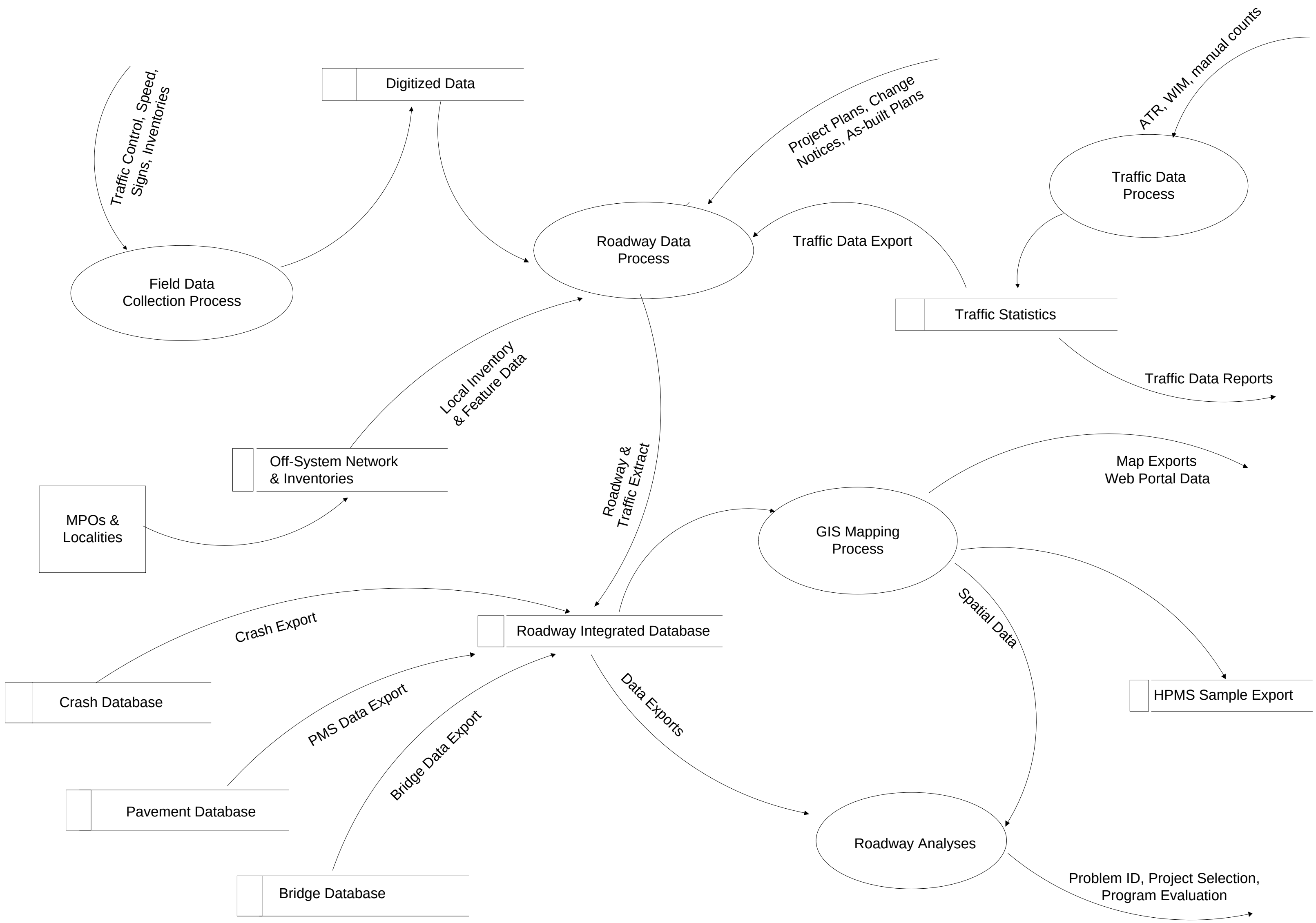


Figure 3. Sources of roadway data.

The following provide examples of various roadway databases and what elements may be included in each of them.

Roadway Physical Characteristics

Data about a roadway's physical characteristics include elements such as:

- Lanes (e.g., total number of lanes, number by direction, average width).
- Shoulders (e.g., shoulder type and average width).
- Medians (e.g., median type and width).
- Speed limit (e.g., design speed, posted speed).
- Design features (e.g., curbs, sidewalks, striping, bikeways, structures).
- Pavement (e.g., age and type, serviceability index, roughness index).

Roadway Access Control

Data about a roadway's access control may include elements such as:

- Access control (e.g., publicly, or privately controlled).
- Toll type and charge (e.g., toll road charge or not).
- High occupancy vehicle (HOV) operations (e.g., identifies segments with HOV operations).
- Land control (e.g., control of land through which a roadway segment passes).

Roadway Intersections

Data about a roadway's intersections may include elements such as:

- Intersection type (e.g., circle, roundabout, T-intersection, on or off ramp, etc.).
- Number of lanes (e.g., number of mainline through lanes or turn lanes).
- Traffic control (e.g., signal, sign).
- Annual average daily traffic (e.g., AADT of intersection and legs, year of count).

These various roadway databases may be located at a city or county public works department, a regional planning agency, or a State department of transportation. These inventories generally will exist in the form of a relational database used with a geographic information system (GIS), an asset management system, or individual legacy systems that focus on one type of roadway component (e.g., bridges, rail grade crossings, pavement management). In some instances, GIS shapefiles are transported among jurisdictions such as, from local systems to state systems. The following are examples of sources of roadway-related data. Not all of these roadway data sources are available in every jurisdiction or for every section of roadway.

ROAD INVENTORY MAPS

Road inventory (RI) maps, or straight-line charts, are the linear representation of a roadway segment drawn from as-built plans or the most current construction or design plans[1]. Agencies often identify locations in terms of control/section and job number. Whether drawn manually or with specialized software, these diagrams and the as-built plans are valuable sources of roadway characteristics because of their portability and accessibility by field personnel. Figure 4 is an example of a manual RI map, see http://onlinemanuals.txdot.gov/txdotmanuals/szn/developing_strip_maps.htm. Figure 5 is an example of a digitized road inventory map from Michigan's RoadSoft software, see http://www.roadsoft.org/asset-management/road-inventory.

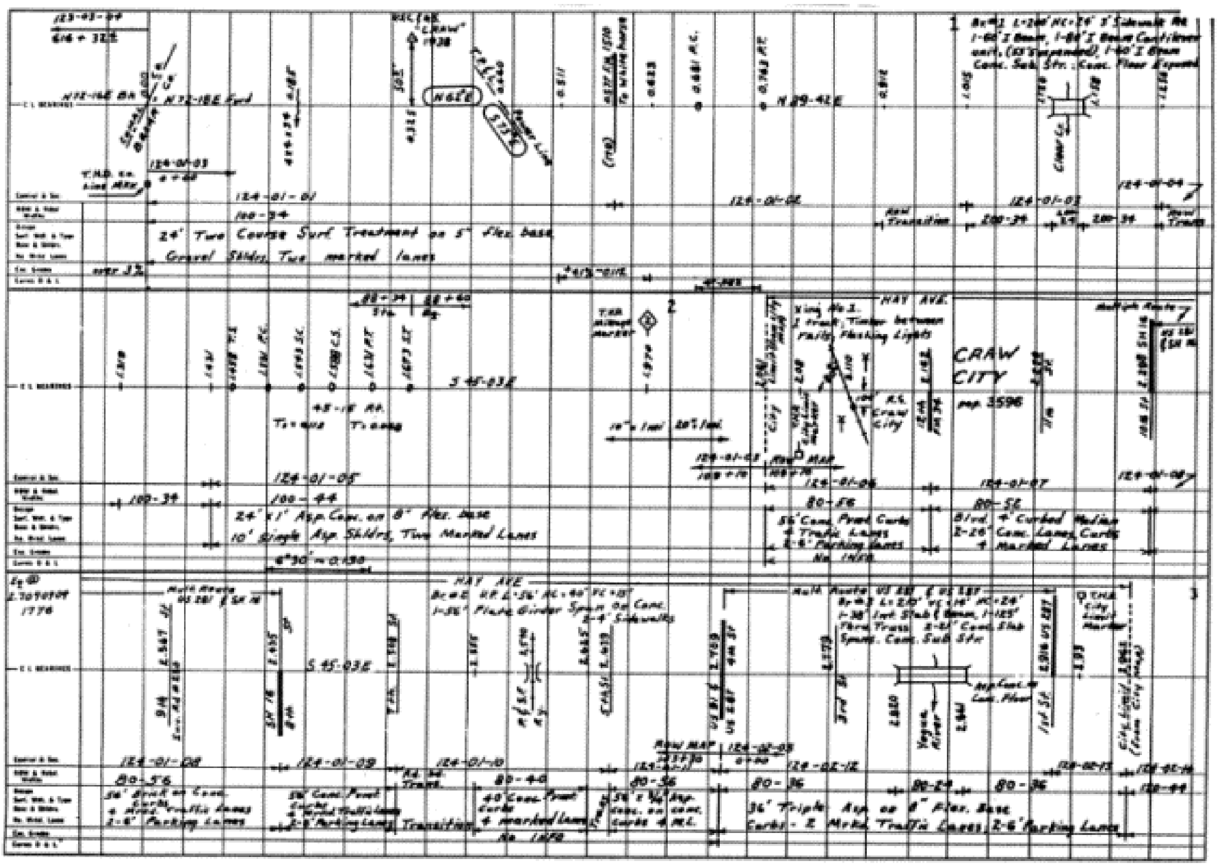


Figure 4. Road inventory straight-line diagram.

[1] For the purposes of this Guide, the term as-built plans refers to the as-built plans or the most current construction or design plans available.

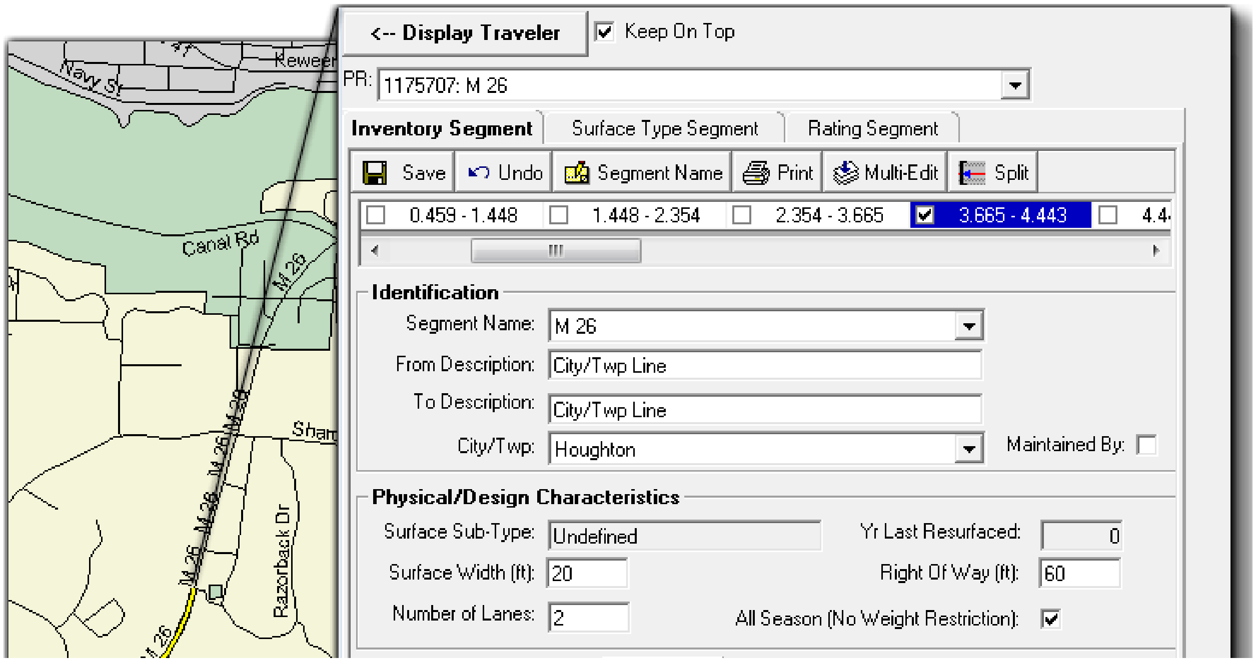


Figure 5. Automated road inventory map.

Roadway Inventory Database

An example of one way an agency could automate its roadway inventory records is by using the characteristics derived from the roadway inventory maps. Each record describes characteristics of a homogeneous roadway segment. A new section number indicates a change in the physical characteristics or geometry of the roadway; for example, transition to wider lanes, more lanes, a bridge feature, different pavement material, etc. If future construction or maintenance changes only a portion of the roadway segment, agencies break the roadway inventory record into two records representing homogeneous segments. Traditionally, some agencies may have coded roadway inventory records with the control/section numbers matching the inventory maps, and the beginning and ending milepoint or reference marker for the segment. Agencies add the new location coding to reflect the new segments and allow the changes to be reflected on any GIS mapping. Aspects of a roadway inventory database may include:

- Milepoints are represented to thousandths of a mile from the beginning of a semi-fixed reference point (e.g. a state or county boundary line). As a document-based calculation of distance, measures are useful in determining distances from automated database records. Note: While .001 miles is a common point of accuracy, it is not always used. States generally measure their roads in a cardinal direction, starting at a jurisdictional point that is stable, like a State or County line.
- Reference markers (also called mileposts) are physical markers or tags that can be found along the roadway at pre-defined intervals; e.g., on interstate highways these are located

consistently at some interval from one-tenth mile to one-mile. On other types of roadways, the markers may be further apart and may not be as visible to the traveling public (e.g., a metal tag attached to the back of other signage. Using reference markers, a location is defined in terms of some fixed distance from the marker. Reference markers are useful to law enforcement officers or highway workers who must identify a location while in the field. However, reference markers are not always geospatially accurate.

- Coordinates (latitude/longitude), obtained global positioning systems (GPS) or automated map locators, are becoming more commonly used to identify roadway locations. This may be for initial creation of a roadway network for use with a GIS or may be used to identify the location of individual incidents. Even when coordinates are used to identify a location, other location coding methods may be needed to link to older legacy databases. In fact, most highway departments maintain extensive cross-reference tables to link data stored with different methods of location coding.

Traffic Characteristics Database

Traffic data reflects the volume and characteristics of vehicles that travel along the roadway. Agencies often store traffic counts, such as AADT, in a manner that allows analysts to integrate traffic data with the roadway inventory database. The FHWA maintains an extensive library of research and operational studies about traffic characteristics see http://www.fhwa.dot.gov/policyinformation/hpms.cfm.

Data about a roadway's traffic may include elements such as:

- AADT (e.g., whether actual/estimated and year recorded).
- Counts (e.g., average daily traffic [ADT] hourly/weekly).
- Adjustments (e.g., compensation, seasonal and directional factors, etc.).
- Site (e.g., location of counts, project, interstate corridor, etc.).
- Vehicle classification (e.g., number of buses, trucks, axles, trailers, etc.).

Structures Database

The FHWA National Bridge Inventory (NBI) is a database of approximately 600,000 of the Nation's bridges located on public roads, as well as publicly-accessible bridges on Federal lands. States provide data to FHWA for the NBI. There is a unique identification code (e.g. structure file number) for each bridge that allows it to be located spatially for mapping. The database includes:

- Material of the bridge components, deck, and deck surface.
- Operational conditions such as structure age, construction year, rehabilitation year.

- Type of services and traffic carried over and/or under the structure.
- Number of the lanes over and/or under the bridges.
- ADT.
- Average daily truck traffic.
- Information relating to bypass or detours.

Furthermore, the bridge inventory contains information about inspection data, ratings assigned by inspectors and appraisal results. For additional information about bridges and other structures, see *http://www.fhwa.dot.gov/bridge*.

Highway-Railroad Grade Crossing Database

FHWA established a formalized Highway Safety Improvement Program (HSIP) database to ensure that safety programs are carried out in an organized, systematic manner where the greatest benefits can be achieved. The Highway-Rail Grade Crossings Program is one of the infrastructure-related HSIP programs.

Data about a highway's railroad grade crossings may include elements such as:

- Crossing number (i.e., from U.S. DOT / American Association of Railroads).
- Crossing signal type (e.g., flashing lights, bells, signs, etc.).
- Number of tracks (e.g., mainline and other).
- Train speeds (e.g., maximum speed or lower speed for trains at crossing).
- Materials (e.g., road surface before or between tracks or rails).

All crossings (i.e., public, private and pedestrian, both at-grade and grade separated underpasses and overpasses) have a U.S. DOT Crossing Inventory Number assigned and posted at the crossing. In addition to State-maintained databases, the Federal Railway Administration (FRA) maintains railroad safety data including accidents and incidents, inspections, and highway-rail crossing data. Agencies include coordinates (longitude and latitude) in the record to identify the location along with the crossing inventory number. For more detailed information about these data, see *http://safety.fhwa.dot.gov/xings*.

Pavement Management Systems (PMS)

GIS can be used as a platform for a pavement management system (PMS) to visually display the pavement attributes for a section of roadway and its location. Reference markers may be included because of the need to monitor or test pavement sections in the field. A PMS may include video logs, as well as data about pavement type, pavement condition, friction, failures, and serviceability. For more detailed information, see FHWA's Pavement Publications at *http://www.fhwa.dot.gov/pavement/pub_listing.cfm*.

A PMS is one element of roadway maintenance. Pavements represent the largest capital investment in any modern roadway system. Maintaining and operating pavements on any roadway system typically involves complex decisions about how and when to resurface or apply other treatments to keep the roadway performance and operating costs at reasonable levels. A PMS usually consists of three major components:

1. A system to collect roadway condition data on a regular basis.
2. A computer database to sort and store the collected data.
3. An analysis program to evaluate repair or preservation strategies and to suggest cost effective projects to maintain roadway conditions.

Most jurisdictions will combine the PMS analysis results with planning needs and political considerations to develop annual roadway repair/preservation programs. Data collection focuses primarily on pavement condition and ranges from simple "windshield surveys" (often with subjective results) to the use of elaborate testing vehicles that measure smoothness, skid resistance, faulting, and cracking in the road surface. Agencies may link to other data in a PMS, such as traffic volumes and physical descriptions of the roadway (width, sub-surface, shoulder/edge, etc.).

The analysis component of a PMS attempts to predict how long a pavement segment will last with a certain kind of repair under the given traffic loads, climate, and other factors. This analysis is based primarily on the collective experience of roadway experts and on the historical costs incurred for repairs or reconstruction. More sophisticated analysis packages also predict annual repair costs, overall system performance, and expected pavement conditions on related routes within planning corridors. The intent of the analysis is to identify the most cost-effective ways to maintain a roadway system in satisfactory condition. Temporality is important in a PMS since it's critical in analyzing pavement data similar to its importance in analyzing safety data.

Roadway Assets

Data about a roadway's characteristics or attribute may include elements such as:

- Hardware Type (e.g., signal, guardrail, sign, crash barrier, light pole, etc.).
- Maintenance (e.g., State-maintained or local-maintained).
- Condition (e.g., appraisal of condition, conformity, distance from right-of-way).
- Electronic (e.g., cameras, message signs, meters, advisory radio, monitor).

Not all of these data are available in every jurisdiction or for every section of roadway. However, a national guideline, MIRE, to provide a model of what roadway and traffic data elements a jurisdiction should collect to support data-driven decisions for safety. In many States asset management systems are being used to coordinate the management of various inventory databases.

Various specialized inventories may be available, particularly for a local or regional transportation agency. These inventories define numerous features and activities, for example, billboards, adopt a highway, permits, work zones, roadway type, speed limits, safety devices (guard cable), signals, lighting, flashers, construction projects, and other types of roadway-related data.

The Tennessee DOT has been able to successfully create asset inventories. The DOT extracts the data collected from both Light Detection and Ranging (LiDAR) and Photolog. Figure 6 is an example using LiDAR and related software.

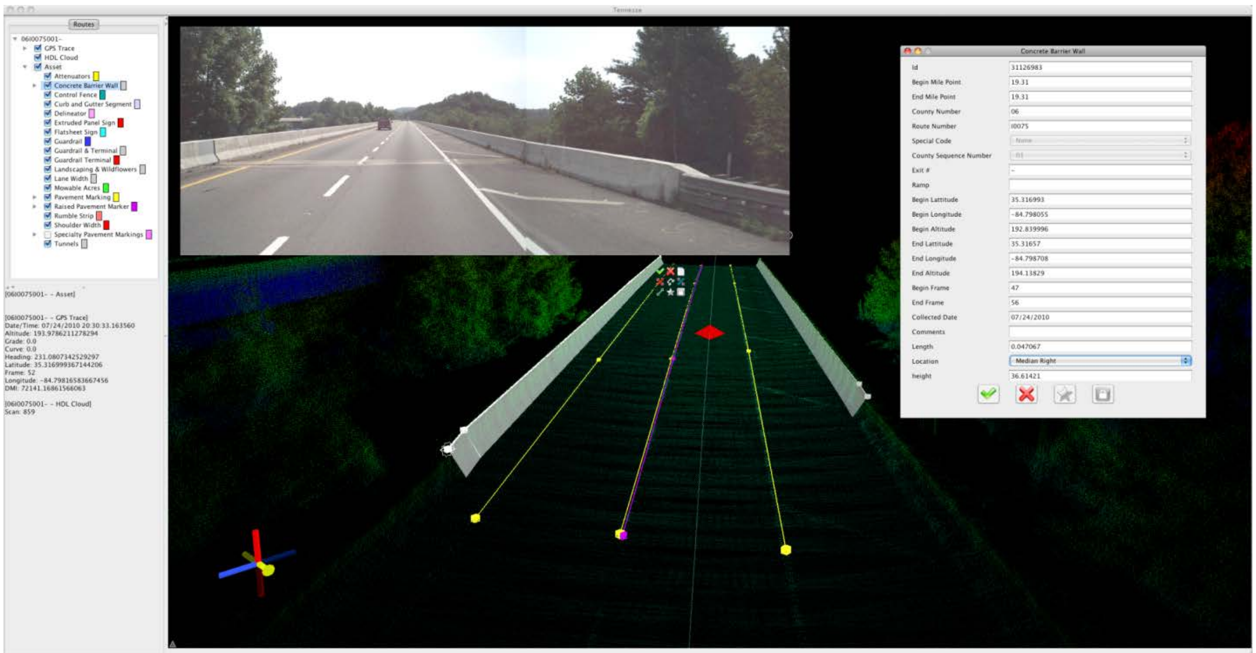


Figure 6. Roadway view using LiDAR by Tennessee DOT

Figure 7 shows an example of how assets can be mapped to illustrate their location on the roadway. For a more detailed look at asset management and inventories, see http://www.fhwa.dot.gov/asset/.

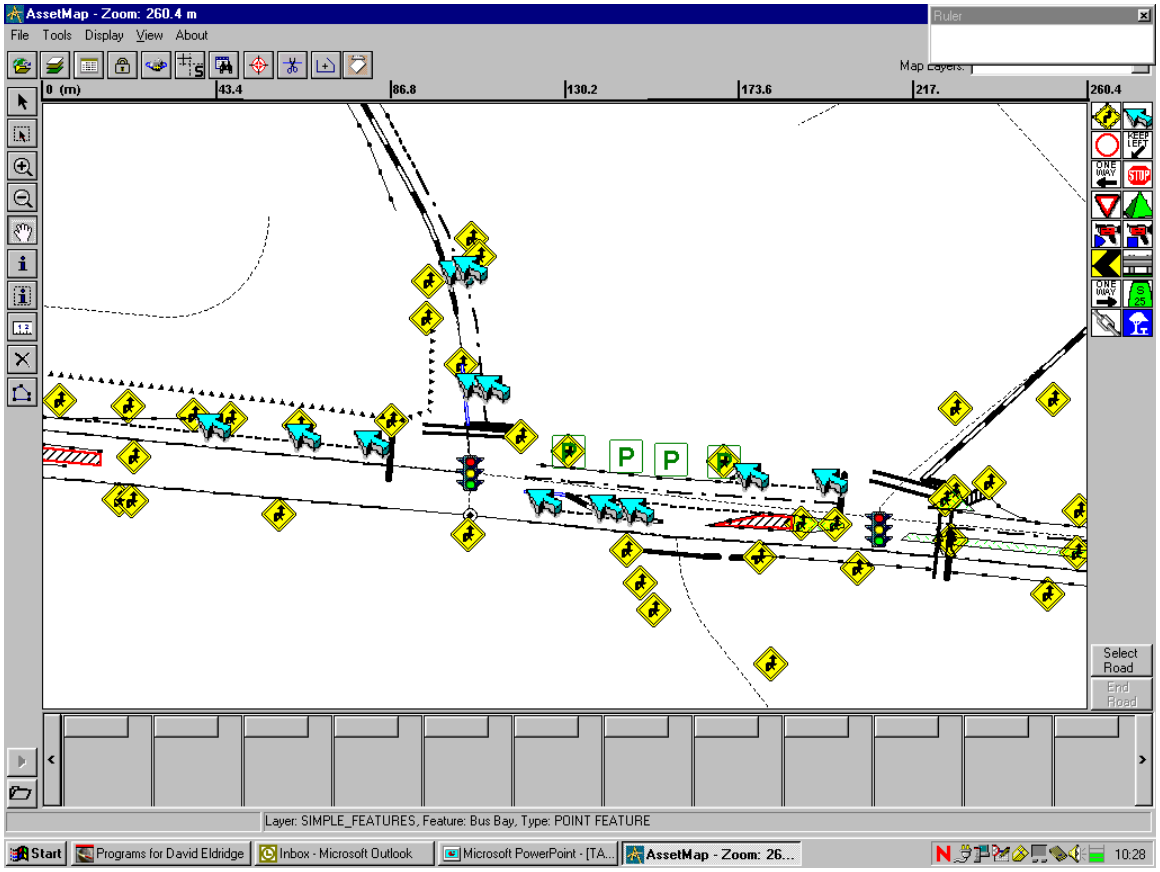


Figure 7. Example of asset management reporting.

Highway Performance Monitoring System (HPMS)

The Highway Performance Monitoring System (HPMS) is a national level highway information system program that provides data on the extent, condition, performance, use, and operating characteristics of U.S. highways. The FHWA Office of Highway Policy Information manages the HPMS. Through the HPMS, FHWA requires States collect roadway characteristic data and report those data to FHWA. Currently, the HPMS database contains over 123,000 highway segments.

The number of data elements and extent of roadway that they have to be collected on vary based on the type of roadway. The data required for the annual submittal of HPMS includes: (1) limited data on all public roads (Full Extent), (2) more detailed data for designated sections of the arterial and collector functional systems (Sample Panel), and (3) area-wide summary information for urbanized, small urban and rural areas (Summary). In addition, with 2010+ Reassessment, States are also required to submit a linear referencing system for the Full Extent and Sample Panel data on selected highway functional systems, see *http://www.fhwa.dot.gov/policyinformation/hpms/fieldmanual/*.

The major purpose of the HPMS is to support a data driven decision process at the national, State, and local level for meeting the nation's transportation needs. FHWA uses the HPMS data for highway system performance assessment, condition and performance (C&P) reporting, apportionment of Federal-aid highway funds, reporting of highway statistics, and other transportation related analysis.

HPMS information is also available for highway and transportation planning and other purposes through the report of annual Highway Statistics or through other media, see *http://www.fhwa.dot.gov/policyinformation/hpms.cfm*.

Highway Safety Information System (HSIS)

The Highway Safety Information System (HSIS) is a roadway-based system operated and maintained by FHWA that provides quality data on a large number of crash, roadway, and traffic variables. This multi-State system contains information already being collected by agencies for managing the highway system and studying highway safety. In some cases, special inventories such as intersections, interchanges, and roadside hardware are available. The data are acquired annually from a select group of States and one municipality, processed into a common computer format, documented, and prepared for analysis. Data users (practitioners, researchers, etc.) can request specific data from FHWA for approved transportation research. See *http://www.hsisinfo.org/* for more information.

Data-Driven Decisions

Many other systems could be included in the discussion in this Chapter. The important point is not just that the data exist, but that they exist to be used. Data-driven decision-making is an approach that has permeated all levels of government. Practitioners recognize the fact that we make better quality decisions when we base those decisions on valid facts describing the real world. It is an approach that reduces the reliance on individual intuition and focuses more attention on quantifiably proven strategies—the things we know work.

This is no easy endeavor. In fact, decision-making is still as much an art as it is a science. There are often not complete and accurate data; many of the data sources are out-of-date; and it is often difficult to convince decision-makers that the investment in data and advanced data analysis is worthwhile. It also is true that having data does not necessarily simplify the decision-making process. In highway traffic safety, in fact, the more we develop our quantitative understanding of the situation, the more factors we realize should be incorporated into our analyses. The Haddon Matrix is not a description of discrete influences sorted neatly into rows and columns, but rather a framework reminding us that the crashes and crash outcomes we see are the result of multiple interacting influences. The growth of multi-disciplinary approaches to traffic safety is the natural consequence of our increased understanding of the situation. The state-of-the-practice is to work jointly with engineering, enforcement, emergency care providers, educators, and others to solve the problems. Each specialty has a contribution to make to solutions that are most effective when coordinated.

CHAPTER 3—TRANSPORTATION PLANNING AND COORDINATION

States conduct continuing, comprehensive, and collaborative intermodal statewide transportation planning that facilitates the efficient and economic movement of people and goods in all areas of the State, including metropolitan areas. Planning a transportation system is a continuous process influenced by changes in traffic patterns, congestion, air quality, safety, technology, politics, events, culture, legislation, and rule making. Transportation planning occurs at the State, region, and local levels of government. The FHWA has broadly defined the planning process as a series of steps:

- Monitoring existing conditions.
- Forecasting future population and employment growth.
- Assessing projected land uses in the region/identifying major growth corridors.
- Identifying problems and needs, and analyzing through detailed planning studies, various transportation improvements.
- Developing alternative capital and operating strategies for the movement of people and goods.
- Estimating the impact of the transportation system on air quality within the region.
- Developing a financial plan that covers operating costs, maintenance of the system, system preservation costs, and new capital investments.

Transportation agencies need to make decisions in an environmentally sensitive way, using a comprehensive planning process that includes the public and considers land use, development, safety, and security. The Volpe National Transportation Systems Center, *Analysis of State Long Range Transportation Plans*, identified the following options to take into account when making decisions:

- Factors required by legislation.
- Type of plan (such as needs based, vision-based, policy, project, corridor, or fiscally realistic).
- Multi-modal planning (includes water, aviation, transit, rail, bike, & pedestrian).

See *http://www.fhwa.dot.gov/planning/processes/statewide/practices/anaswplans.cfm* for more information.

The current transportation legislation, MAP-21, identifies the following performance factors that need to be considered in the statewide transportation planning process:

- Safety
- Infrastructure conditions

- Congestion reduction
- System reliability
- Freight movement and economic vitality
- Environmental sustainability
- Reduced project delivery delays

As is evident from this list, data play a central role in the decision making in all aspects of a project plan, and the information needs include both roadway and non-roadway data. Since planning will vary by State and local needs, the data used vary by jurisdiction. Current roadway data collected for the planning process may include ADT, crash history and severity, and specific roadway characteristics such as number of lanes, pavement type, intersection type, and rail crossing signal type. The information available for planning purposes should also include historical data to assist planners in monitoring trends and forecasting future mobility needs. Examples of non-roadway data necessary for the planning process include air quality, demographics, household income, environmental detail, land use, population growth, and political policy.

Context Sensitive Solutions (CSS)

FHWA defines "context sensitive solutions" (CSS) as "a collaborative, interdisciplinary approach that involves all stakeholders in developing a transportation facility that complements its physical setting and preserves scenic, aesthetic, historic, and environmental resources while maintaining safety and mobility. When transportation planning reflects community input and takes into consideration the impacts on both natural and human environments, it also promotes partnerships that lead to 'balanced' decision making."

Note that, in a typical road project (see Figure 8), the planning phase is followed by the "project development" phase, which includes these basic steps:

- Refinement of purpose and need.
- Development of a range of alternatives (including the option not to build).
- Evaluation of alternatives and their impact on the natural and built environments.
- Development of appropriate mitigation.

The project development phase includes a context sensitive design (CSD) component, which is related closely to the earlier CSS activity. CSD reflects an emerging national trend that strives to preserve and enhance a sense of place when building or expanding road, transit, bicycle, and pedestrian projects. The intent is to include broad community participation and comment in a project's selection and design. Figure 8 shows an example of the context-sensitive process. Additional information and the partnerships involved, along with detailed technical information about CSS is available at http://contextsensitivesolutions.org/.

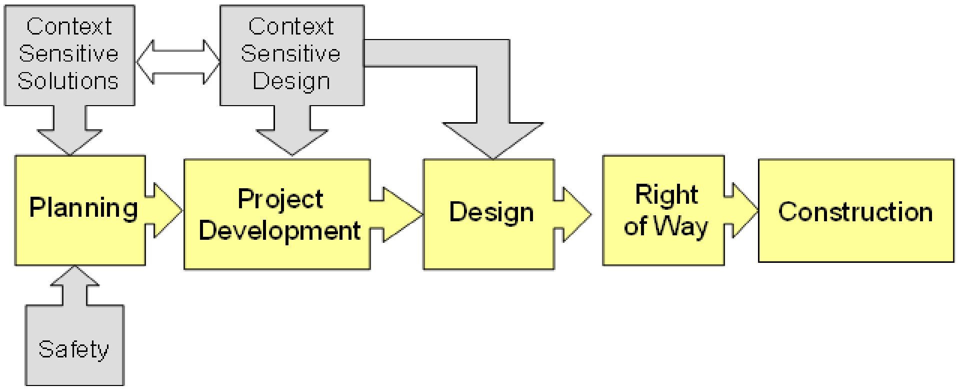


Figure 8. An example of a context-sensitive process.

The planning process in most jurisdictions also determines the functional classification of the roadway. The functional classification is used in transportation planning, in roadway design and for allocation of federal roadway improvement funds. The functional class hierarchy is:

- Principal Arterial - Interstate.
- Principal arterial other freeways and expressways.
- Principal arterial other.
- Minor arterial.
- Major collector.
- Minor collector.
- Local.

The concepts and criteria for classification were formulated based on the Federal-aid Highway Act of 1973. "Functional classification is the process by which streets and highways are grouped into classes, or systems, according to the character of service they are intended to provide. Most travel involves movement through a network of roads. It becomes necessary then to determine how this travel can be channelized within the network in a logical and efficient manner."

Roadway Design

Beyond the planning and the project development phases is design. At this stage, agencies select a preferred design alternative, determine the functional class, and evaluate other design considerations for inclusion into the final design.

Federal law governs roadway design; Title 23 United States Code (USC) 109 requires that the Secretary of Transportation, in cooperation with the State highway departments, approve design standards for projects on the National Highway System (NHS) (*http://www.fhwa.dot.gov/legsregs/directives/cfr23toc.htm*). On the federal level, this authority has been delegated to the Federal Highway Administrator.

State highway departments, working through AASHTO, develop design standards through a series of committees and task forces. FHWA contributes to the development of the design standards through membership on these working units, sponsoring and participating in research efforts, and many other initiatives. Following development of the design standards, FHWA uses a formal rulemaking process to adopt those it considers suitable for application on the NHS.

The design standards currently adopted by the FHWA can be found at 23 CFR 625 (Code of Federal Regulations under Title 23 USC). Additional design guides and references can be found in the Federal-aid Policy Guide, Non-regulatory Supplement (NS) 23 CFR 625, paragraph 16. The CFR and Non-regulatory Supplement point to a list of publications that contain the standards. Probably the most frequently used document in 23 CFR 625 is A Policy on Geometric Design of Highways and Streets (often called the AASHTO Green Book).

Title 23 USC 109 also requires that projects (other than highway projects on the NHS) shall be designed, constructed, operated, and maintained in accordance with State laws, regulations, directives, safety standards, design standards, and construction standards. Design standards for each State are available at *http://www.fhwa.dot.gov/programadmin/statemanuals.cfm*. At this point in the project with the final design complete, land acquisition is fulfilled, and the actual construction phase is bid and awarded. This ultimately leads to physical construction and maintenance of the roadway.

State Transportation Improvement Program (STIP)

As defined by FHWA, the State Transportation Improvement Program (STIP) is a "short-term transportation planning document covering at least a three-year period and updated at least every two years. The STIP includes a priority list of projects to be carried out in each of the three years. Projects included in the STIP must be consistent with a long-term transportation plan, must conform to regional air quality implementation plans, and must be financially constrained (achievable within existing or reasonably anticipated funding sources)."

The STIP itself is not a plan, per se, but a budget document that is used to schedule and fund projects. Projects listed in the STIP typically come from local or State approved plans. All projects that will receive federal funding must be included in the STIP and may include programs such as:

- Pavement management.
- Bridge management.
- Safety.
- Transit.
- Congestion management, etc.

Many States and local jurisdictions publish their STIP to the internet and solicit public comment on the proposed next version of the STIP. The CSS concept usually has already involved all stakeholders in the planning process before the project is eligible for inclusion in the STIP. The STIP process is determined by a jurisdiction, using guidelines set by federal law, and includes input from all the relevant stakeholders, including the public.

Highway Safety Improvement Program (HSIP)

The HSIP emphasizes a data-driven, strategic approach to improving highway safety that focuses on results. Through HSIP, FHWA provides federal funding to States for the purpose of reducing traffic fatalities and serious injuries on all public roads through the implementation of infrastructure-related highway safety improvements. Since roadway data are crucial in evaluating infrastructure programs, these funds can be used for data collection. MAP-21 Guidance developed by the FHWA Office of Safety identifies eligible uses of for HSIP funds to support data collection, analysis and interoperability. More information on funding eligibility can be found at http://www.fhwa.dot.gov/map21/guidance/guidesafetydata.cfm. Each State should contact their FHWA Division Office for additional information on available funding sources. Contact information for the FHWA Divisions Offices can be found at *http://www.fhwa.dot.gov/about/field.cfm*.

Additional information on the HSIP is available on the FHWA Office of Safety website at *http://safety.fhwa.dot.gov/hsip/*.

Strategic Highway Safety Plan (SHSP)

The purpose of an SHSP is to identify the State's key safety needs and guide investment decisions to achieve significant reductions in highway fatalities and serious injuries on all public roads. The SHSP allows all highway safety programs in the State to work together in an effort to align and leverage its resources and positions the State and its safety partners to address the State's safety challenges on all public roads.

An SHSP is a statewide-coordinated safety plan that provides a comprehensive framework, and specific goals and objectives, for reducing highway fatalities and serious injuries on all public roads. This statewide document is developed by the State DOT. The SHSP is a data-driven, four to five year comprehensive plan that integrates the "4 Es" (i.e., engineering, education, enforcement, and emergency medical services). The SHSP establishes statewide goals, objectives, and key emphasis areas and is developed in consultation with Federal, State, local, and private sector safety stakeholders" ((*http://safety.fhwa.dot.gov/hsip/shsp/*).

NHTSA Traffic Records Planning

The purpose of a *Traffic Records Strategic Plan* is to guide the State's TRCC and its member organizations in fulfilling a shared vision of what the State's traffic records system should be: its organization, management, contents, and functions. The major components of a comprehensive traffic records system include:

- Crash information.
- Roadway information.
- Exposure (traffic volumes).
- Driver information.
- Vehicle information.
- Citation/adjudication information.
- Injury surveillance information.

Traffic Records Strategic Plan

MAP-21 identifies complimentary data programs between FHWA, NHTSA and FMCSA as a requirement. One means to help achieve complementary data programs at the state level is the development of the state Traffic Records Strategic Plan. This is a document that provides a roadmap from the current traffic records system to a desired future system that best meets the identified needs of all stakeholders in the State. Strategic planning must translate the desired system into reality, including an organizational structure, operating budget, and resource decisions.

NHTSA's requirements for a Strategic Plan relate to eligibility for funding under current traffic records grant programs (Section 405c), but may be interpreted more broadly as a requirement to link the State's safety plans to the traffic records strategic plan. The support for making this linkage is that all of these planning efforts require data and require some statement about the quality and reliability of the data used in the plan. Where deficiencies are noted, a State is expected to address those in the plan with a series of activities that will result in improvement. More information is available on NHTSA's website at http://www.nhtsa.gov/Data/Traffic+Records.

NHTSA Requirements for a Traffic Records Strategic Plan

- *Be comprehensive and multi-year.*
- *Have the State's TRCC approval.*
- *Be reviewed annually to ensure that project goals are defined and are being met.*
- *Address existing deficiencies, how they were identified, and priorities for corrective action.*
- *Identify performance-based measures and matrices for measuring progress, including benchmarks.*
- *Indicate what funds will be used and how they will be used to address the goals of the plan.*
- *Establish timelines and accountability for components of the plan.*
- *Integrate the State data needs and goals with the State's SHSP.*

The plan is not the only place that some of this information should be found. Benchmarking and performance measures describing the current status of the traffic records data (its timeliness, accuracy, completeness, etc.) are not items created solely for insertion into the Strategic Plan, but are intended to be part of an ongoing quality control measurement effort – the data are to be referenced in the plan, not created solely for the plan.

Traffic Records Coordinating Committee (TRCC)

Under the NHTSA guidelines, the State TRCC is responsible for reviewing, updating, and approving the Traffic Records Strategic Plan. Updates are called for annually. As part of this role, the TRCC is expected to have the authority to review agencies' plans for changes to any component of the traffic records system. That review authority is expected to be recognized through formal Memoranda of Understanding (MOU).

The responsibility for developing and maintaining traffic records information that supports a State's traffic safety mission is shared among many organizations and literally hundreds of individuals. Without a coordinated process of strategic planning, these organizations and individuals may take actions that may inhibit the State's ability to manage its traffic records systems in an efficient and effective manner. Simple decisions to add or delete specific information items require careful consideration and planning if the traffic records system is to remain viable. Strategic planning offers a means by which these decisions may be taken in a coordinated way.

Coordinating the management and improvement of a traffic records system is the role of the TRCC. This takes on many different forms among the States, but one of the most common assignments for the TRCC is to oversee the development and maintenance of the *Traffic Records Strategic Plan*. To accomplish this, the TRCC will have to take time to define and communicate its vision, relate its activities to this vision, measure its progress in achieving this vision, and identify changes in its behavior that must be accomplished to maintain this vision. This activity will prepare the organization for the future and permit it to act in advance of events that may affect the system rather than be driven by those events in a direction that the TRCC and its members do not wish to proceed.

By virtue of this coordination, the strategic planning effort focuses the TRCC on promoting a system-wide approach to traffic records and its role in helping its member agencies move forward. The TRCC cannot and should not act alone. To be truly effective, the TRCC must work with its member agencies to achieve mutually agreed upon goals. Defining the active and facilitative roles of the TRCC is one of the primary uses of a good Strategic Plan.

Traffic Records Assessments

Traffic records may be viewed as a product that is supplied by the agencies. It is important to look at the quality of that product, list the customers, their needs, and their systems' strengths and weaknesses. It is also important to be aware of opportunities for improvement and threats to the product's quality or the system's ability to supply that product.

A traffic records assessment takes stock of the status of all of the related traffic records system components, the environment in which the system operates, and considers the current service provided. An agency must conduct an assessment to identify system strengths, weakness, and areas of potential improvements. Once an assessment is completed, the detailed review provides awareness to all of the stakeholders that have a vested interest in the system and helps bring a practical focus on identifying the various goals to be pursued by projects identified in the Strategic Plan.

A vision has evolved from these and other efforts of a TRS that is not just a hypothetical construct—easily described but never actually existing—to become a reality of integrated data, easily accessible, and capable of supporting analyses that foster the coordinated approaches to traffic safety we know work best. Realization of this vision takes place in the context of multiple custodians over multiple systems, most of which have a primary purpose other than traffic safety. Bringing together data from systems designed to produce revenue, track assets and expenses, monitor insurers and hospitals or care utilization, and a host of other systems and purposes requires both the will to share data and the technology to make data sharing affordable. There are State and federal regulations securing data privacy that must be addressed. Moreover, it is common to find systems of vastly different vintages and capabilities that must be modified or worked around in order to develop the desired data access and sharing processes. This vision of shared data relies on the data management practices that are part of each TRS component system, and imposes yet another layer of data management practices required for the integrated datasets that are created because of the data sharing process itself.

CHAPTER 4—DATA MANAGEMENT

Data managers and users must make the most of the resources available to them, and where data needs are identified, the first task should be to look for already existing electronic sources of the necessary information so that it may be obtained as economically and efficiently as possible.

DATA INTEGRATION

The needs of users and the constraints of tight budgets are best satisfied by integrating data from available sources and making the resulting combined data available to as many users as possible. As noted in the white paper developed from the Traffic Safety Information Systems International Scan, see http://www.tfhrc.gov/safety/pubs/06099/06099.pdf:

> "Excellent crash, roadway inventory, and traffic data are critical to making decisions concerning roadway planning, roadway design and improvement, vehicle design, and driver programs—all of which affect the safety of the driving public. Safety data will become even more critical … to more fact-based safety decisions and to performance-based programs."

The Data Integration Primer from Federal Highway Administration's Office of Asset Management is an excellent introduction from a roadway data management perspective, see http://isddc.dot.gov/OLPFiles/FHWA/010393.pdf. The benefits of integrating data from disparate agency systems may not seem obvious. Some familiar examples of combined datasets used in roadway analyses may help make the point that these efforts are worthwhile. Figure 9 shows an example of the data integration process.

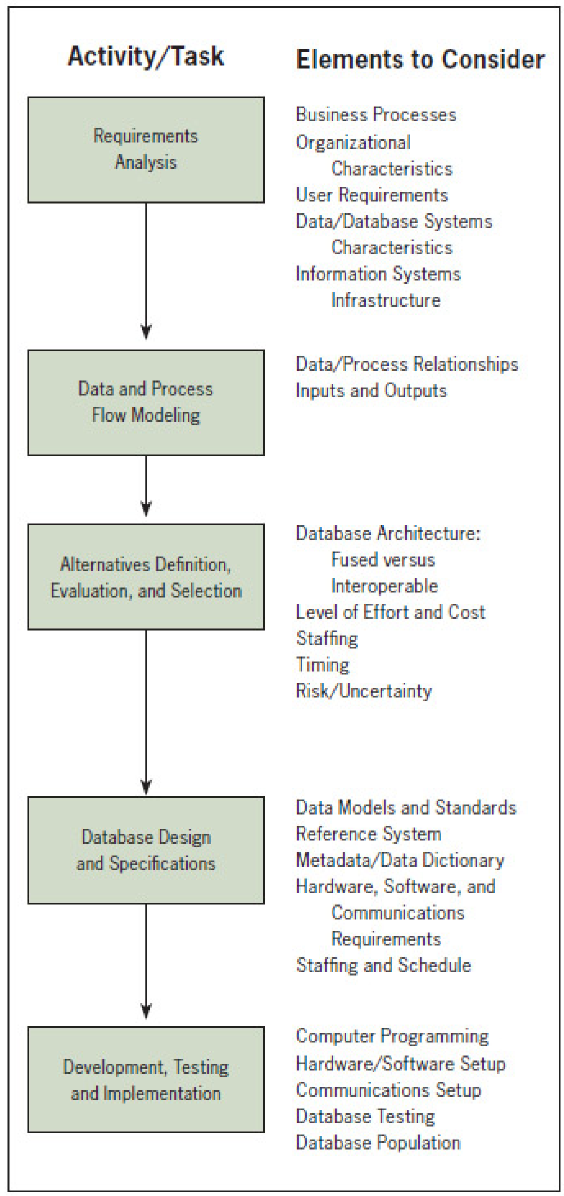


Figure 9. Data integration process.

Results of analysis from integrated systems can provide answers to questions such as *what is the impact on the crash rate by upgrading the intersection design to a roundabout? Which rumble strip* *design is the most beneficial when compared to similar roadways with equal amounts of average daily traffic?* Questions such as these require data from multiple shared data sources. In these examples are the familiar combination of roadway inventory, crash, and traffic volume databases used to support analyses that could not be possible without integrated data.

DISTRIBUTED DATA SYSTEMS

Traditionally, States still often concentrate roadway data collection and maintenance efforts based on whether roadways are state-maintained or non-state-maintained "local" roads, often giving the highest priority to State-maintained roads. The caveat being for HPMS data, which creates a three-tiered system of roadways within a State on which data are collected: Federal-aid roads, non-Federal Aid State-maintained roads, and local roads. However, some States often maintain primary and roadway traffic data inventories independent of their HPMS data submittal. Managing these various datasets can present a challenge.

A potential form of distributed data management of roadway information is a single statewide system covering only those roadways that are managed by the State DOT and a separate system, unconnected to the first, designed to store data on local roadways. In many States, this latter (local) system has been developed in an ad hoc fashion and it is built from multiple local roadway data sources. The so-called "local file" is usually a much less complete dataset incorporating a smaller number of data elements, and often missing data even in the data elements that are included in the system. However, this is not to say that such a two-tiered approach cannot work. It is certainly possible and perhaps even desirable from a cost effectiveness standpoint to develop a distributed system that shares the data collection and management burden among multiple agencies; State, local, Metropolitan Planning Organization (MPO)/Regional Planning Commission (RPC), etc.

From an IT viewpoint, a modern distributed model would consist of source data that are collected and managed locally by those who have the greatest stake in its overall quality. These varying sources of data are linked to a master statewide system that can poll all of the local systems and obtain updated information on an as-needed/as-available basis. For such a distributed system to work well, it is imperative that all participating agencies adhere to a set of agreed-upon standards for data quality, especially consistency across jurisdictions.

CENTRALIZED SYSTEMS

In an attempt to create a complete database of all public roadway locations, many States have opted to develop centralized systems that store records on State and local roads in a single resource. These systems have the obvious advantage over distributed systems of having data quality management reside in a single agency. The realization of such centralized systems has, in many States, fallen short of the ideal because, ultimately, data quality relies heavily on a variety of data providers who (just as in the distributed data systems) may not all adhere to established data quality standards. Only when a State's DOT controls all the data collection and data management can the processes for roadway data collection and management be said to be truly centralized. Such systems do exist, but their costs are born entirely by the DOT and still there are often important differences between the quality of information obtained for state-maintained and locally-maintained roads.

The FHWA Office of Highway Policy Information and Office of Planning, Environment, and Realty issued the Memorandum on Geospatial Network for All Public Roads on August 7, 2012. This Memorandum identified a Highway Performance Monitoring System (HPMS) requirement for States to update their Linear Referencing System to include all public roadways within the State by June 15, 2014, in accordance with the HPMS information collection approval from the Office of Management and Budget (2125-0028). This Linear Referencing System will provide a means to geolocate all safety data on a common highway basemap that includes all public roads.

Enterprise-Wide Data Systems

Some State DOTs have developed a single, centralized data application that incorporates all (or almost all) of the traditional data sources found in a transportation agency. Agencies design such "enterprise" databases to provide a common platform, structure, and tools for all of the department's data resources. This type of system can reduce costs and certainly improves the integration of data among the various sources that make up the components of the enterprise-wide system. Such systems are expensive, but they may represent a cheaper solution than one in which the same department must separately maintain applications software and databases for each of the functions separately. They also help to reduce duplication in data storage and often in data collection. Agencies realize the most benefits when they design a system to meet the needs of multiple users and build the system using strict data governance processes.

Figure 10 shows the information flow for an enterprise-wide roadway data system in a State DOT. This example illustrates an enterprise view of the Illinois DOT based on their roadway data systems.

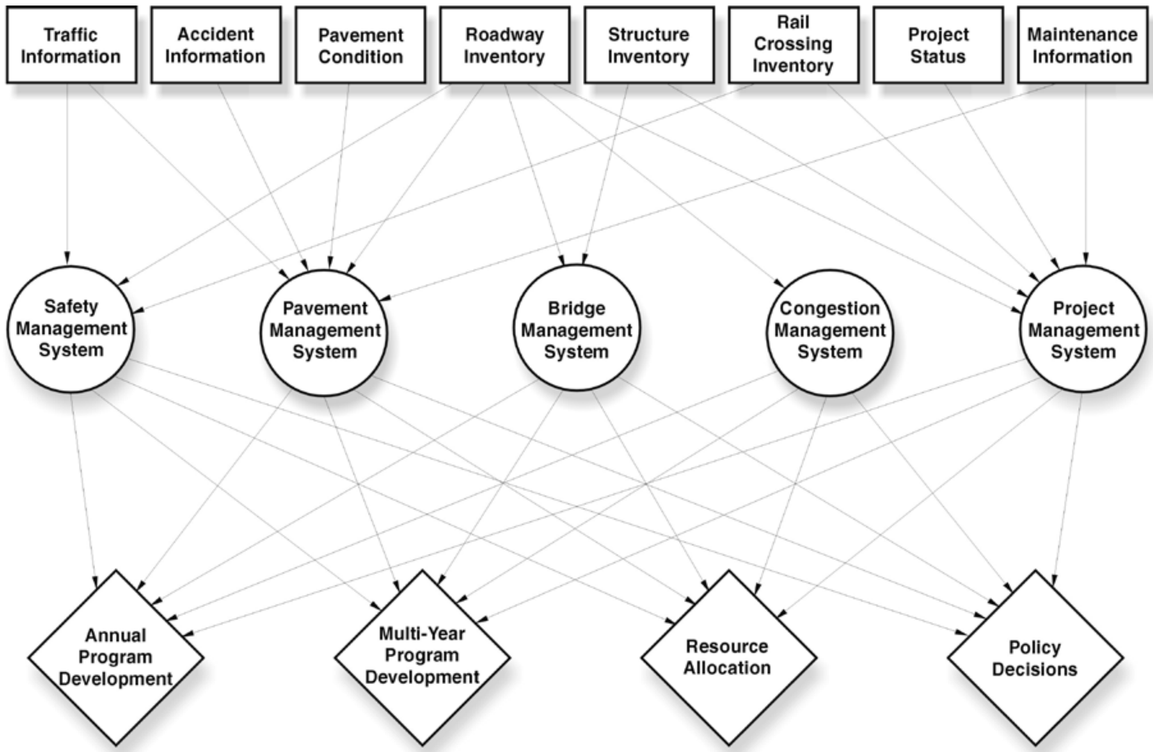


Figure 10. Illinois DOT information flow for asset management.

Even with an enterprise-wide roadway system, many states have created the opportunity for data owners to enter and maintain their own data. The challenge is to keep all of the data in sync over time so temporal analyses can occur and the associated location remains anchored on the correct highway and reference point. For example, in the North Carolina DOT, the Planning Division enters basic inventory data from the original plans, plan changes, and as-built plans and the maintenance operations division updates records as any data is added or updated. In other cases, a DOT's Structures Division may provide all of the data entered into the system about bridges and culverts.

Data Governance

Data governance is a term used with increasing frequency in the government IT area, and especially within State DOTs. The term can mean multiple things; however, the common theme of data governance efforts is to impose common standards for all data elements and data collection efforts throughout a broad community (such as an entire DOT or the DOT plus all local partners). The goals of data governance efforts are to improve coordination among multiple systems by ensuring that data definitions and data collection methodologies are standardized, and that multiple systems using the same (or similar) data are sharing the data to the greatest extent possible. Data governance efforts seek to reduce redundant data collection and simplify the processes of data integration while, at the same time, helping to ensure accuracy and completeness.

A formal data governance process typically includes a set of data governance policies and a data governance "board" made up of executive-level and practitioner-level staff from all stakeholder offices and agencies. An example for a DOT would be a data governance board including the Director of IT, the stewards/custodians of each of the agency's major systems, and the IT managers or staff assigned to support those systems. At a statewide level, data governance boards might include the heads of each State agency, the State Chief Information Officer, and a set of representative data stewards and agency IT staff responsible for the major systems. The functions of the board include setting standards, reviewing system designs, assessing data definitions and requested changes to those definitions, and determining the sequencing and resource allocations for data-related projects.

Because of the historic methods used by most State agencies (DOTs included) for developing data resources, data governance practices and policies often are seen as a barrier to system improvement rather than the aid that they should be. This is unfortunate since a well-run data governance process can save the agencies and the State a great deal of time and money and make much better use of scarce IT resources. Submitting agency plans to a data governance board; however, may involve losing a measure of internal control over project scheduling, scope, and funding.

The decision of whether or not to establish a data governance process/policy and whether or not to submit all of an agency's planned system development efforts to the data governance process is one that is made at the State and individual agency level. System designs and changes that are developed and implemented in a coordinated manner have a better chance of meeting the widest variety of users' needs at the least possible cost. It is possible for States to develop this kind of coordination in a less formalized manner, perhaps through the work of the TRCC and/or through the establishment and maintenance of various strategic plans (e.g., the *Strategic Plan for Traffic Records Improvement;* the *Strategic Highway Safety Plan* data quality emphasis area; etc.). The formal data governance practices discussed here have advantages over the other methods mentioned in that they are specifically designed to include the IT perspective and they are capable of going beyond the traditional set of systems included in the definition of a traffic records system.

The remainder of this RDIP document will focus more exclusively on the Roadway Data Component of a TRS. For the purposes of this report, the term "roadway data" incorporates a broad array of data sources and information systems. As with the overall TRS, the roadway data component systems each exist to fulfill a specific purpose. Often that primary purpose has only a tangential relationship to traffic safety. Fortunately, as with the TRS overall, advances in database architecture and analytic software are making the job of data extraction and integration much easier. Modern databases are more flexible than the legacy systems they are rapidly replacing. Throughout this discussion, it is important to remember that most of the analytic needs for roadway data require spatial information. It is not enough to know the spatial coordinates of a crash or of a particular set of roadway features; the information must also be aggregated into relationships that define particular roadways, routes, and networks of roadways.

CHAPTER 5—ESTABLISHING THE ROADWAY NETWORK

At the most basic level, knowing where the roads are in given a jurisdiction is fundamental to the daily operations of many DOT functions. Location reference methods provide "a way to identify a specific location with respect to a known point" and includes three elements:

1. Identification of a known point.
2. Measurement from the known point.
3. Direction of measurement.

The two basic location reference methods described in those studies are still in use today:

- Sign-oriented methods (milepost, reference post).
- Document-oriented methods (calculated milepoint, route log, and straight-line diagrams).

Linear Referencing System

The location referencing methods (LRM) described above are used by State and local municipalities to develop a Linear Referencing System (LRS) that consists of the LRMs and the procedures to store, manage, and retrieve information about the roadway location data. Historically, State and local linear reference methods were designed and maintained to meet specific, individual business needs. In many cases, several LRMs are used in a single agency, and variations of a LRS may exist within the same agency. For effective safety analyses on all public roads, roadway data inventories must be able to able to be logically linked with other traffic records systems across all State and local agencies; with this linkage most likely occurring based on location.

Resolution, scale, and precision used in various LRMs may introduce accuracy problems; for example, roadway location data in one LRM may use milepoint data in tenths, while another uses milepoint data in hundredths, and a third uses State plane coordinates, and a fourth uses latitude and longitude data from GPS. Assuming that the data in each of the LRMs are correct to the desired degree, transforming each of these methods into a single LRS, or multi-level LRS introduces a wide variance in the resulting accuracy of the location information. If errors exist in the individual LRMs, these are compounded in the results. However, with advancements in technology, this is less of an issue. Techniques have been developed to process and minimize the differences in the data, but confidence with the resulting data and visual displays still vary from LRS to LRS.

When agencies integrate the LRS with the GIS, additional accuracy variance is introduced, depending on the source and precision of the GIS base mapping. Agencies may digitize GIS mapping from various sources or collected using commercially available maps. Ideally, agencies construct precision GIS base mapping using enhanced GPS coordinate data and roadway characteristic data that are collected using GPS tools in the field with a comparable degree of accuracy.

These basic methods have several variations as noted below:

- Route-Milepost.
- Route-Reference.
- Link-Node.
- Route-Street Reference.

These four methods are described in more detail in the flowing sections. This list is not all-inclusive, as variations of these methods are used at times in each agency and jurisdiction.

Route-Milepost

The r*oute-milepost* is the most commonly used method. A measurement generally begins at the beginning of a route or a specific known point such as a jurisdictional boundary (e.g., county border) and continues to the location of interest (e.g., end of guardrail, crash site, sign location, etc.). This is a location method used in the field with mileposts or other physical reference markers positioned accordingly. On major routes, the reference markers or posts are placed where they can be seen for use by maintenance crews, law enforcement, etc. On other routes, the reference marker may be placed in less conspicuous locations, such as on the backs of traffic signs.

These reference markers may get out of sync because the road is changed and the marker placement is not corrected. If they are correctly placed, jurisdictions should be sure to correct their inventories regarding the repositioned markers. This practice may result in degradation of accuracy when encoding or using data that describes the location. When relying on mileposts or other reference markers along the roadway, measurements are generally collected anywhere from tenth of a mile to thousandths of a mile increments, depending on the type of data that are being collected. Figure 11 illustrates a guardrail measurement taken from the beginning of a route. The guardrail begins at 4/10th of a mile past milepost one 1 and terminates at 3/10th of a mile past milepost 3.

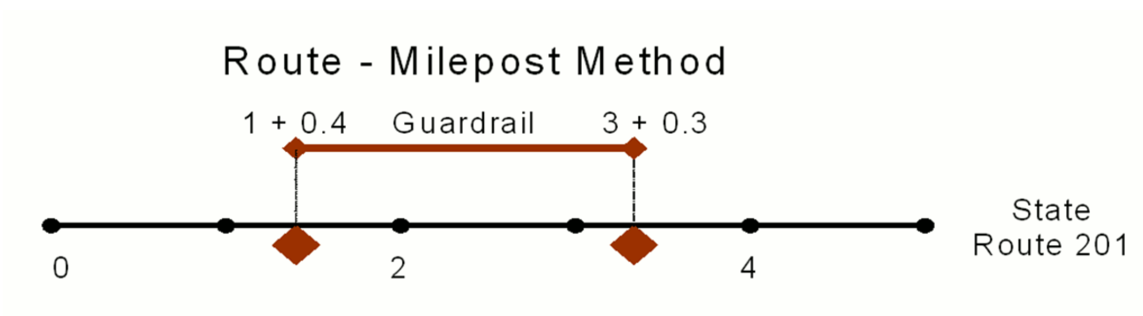


Figure 11. Example of a roadway feature location measured by route and milepost.

ROUTE-REFERENCE

The *route-reference or control/section-milepoint* method is a document-oriented location method with calculated distances from a known location. It relies on known physical landmarks (e.g., markers, bridge, an intersection, water tower, fire hydrant, or a marker such as a numbered pole embedded in the ground). Initial measurements of roadway features (i.e., reference or milepoint) come from the as-built or the more recent construction of design plans available and generally are expressed in hundredths of a mile or thousandths of a mile. Events such as crashes are located using a distance from these markers that is converted to the documented calculations of locations to determine the degree of accuracy. Figure 12 shows a guardrail that begins at 400 thousandths (400/1000) of a mile from the documented marker and terminates at 2.320 thousandths (2 and 320/1000) of a mile from that known point.

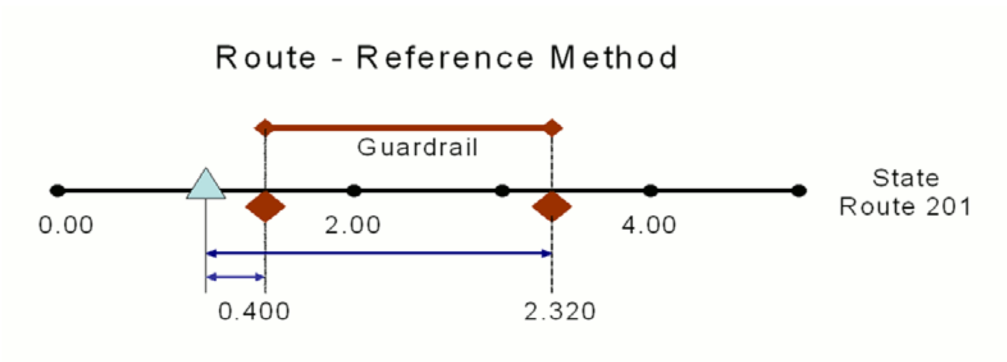


Figure 12. Example of a roadway feature location measured by route-reference.

LINK-NODE

Some State and local municipalities use the link-node method. This is the most generic document-oriented method of location coding and can be used to mimic other document-oriented methods rather easily. For example, using the link-node method, Figure 13 shows a crash occurring on link 102103, 104 feet from node 102. Known locations, typically an intersection, are designated as a unique node with an identifier, and the connector to the next node is a unique link. Links may be of unequal lengths. Roadway features are located by measuring the distance from the nearest node or nodes.

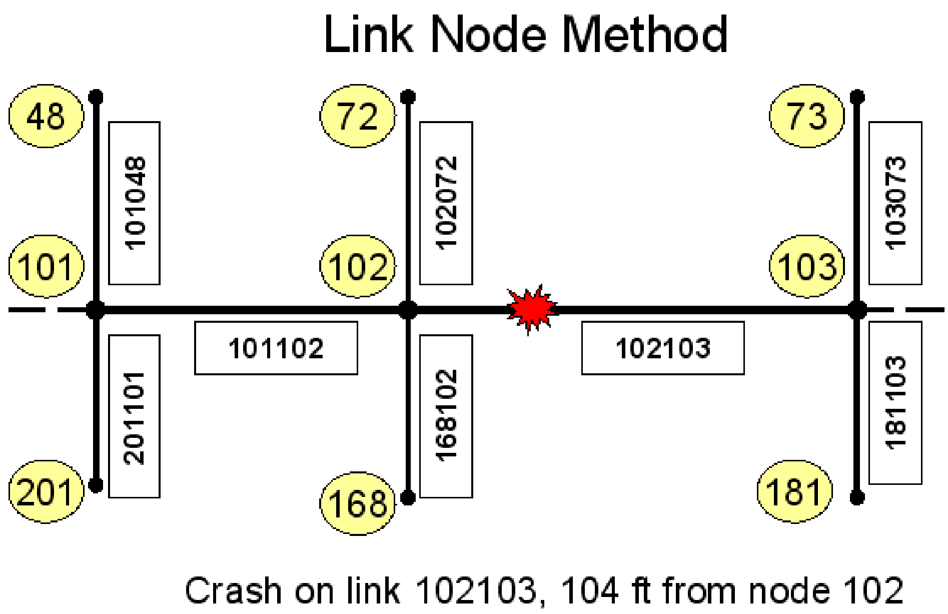


Figure 13. Example of link-node method of identifying locations.

The schema for designating node and link names can be complicated when changes are made to the existing system. This is particularly true for a legacy system that does not utilize a relational database. These legacy systems may not be flexible enough to add additional nodes and segment descriptions without losing historical data and essentially entering all of the data again that has accumulated over the years for that roadway segment.

Route-Street Reference

A similar method to link–node is route–street, typically used in urban settings. Rather than assigning node and link designations, the street names or unique street codes are used and distances then are measured from intersections or other fixed landmarks. Examples of complications when using this method include:

- Street names change over time or there are alias names for the street that are used interchangeably by data collectors.
- Streets intersect more than once with another street.
- Technicians or law enforcement officers fail to differentiate streets of similar names (e.g., Stanford Drive and Stanford Court, or West Prospect Road and Prospect Road).
- Block numbers can be used to segment a street and indicate where an event occurred, rather than measured distances from a known point.

GEOGRAPHIC COORDINATES

A different method used by both State and local municipalities is Cartesian coordinates along an x and y plane, often by means of the State plane coordinate system (SPCS, established in the 1930s for each State). Typically, a cartographic map is produced with the State plane coordinates fixed on the map, showing roads and other known fixed locations and geographic features. Locations of roadway characteristics are digitized manually to collect the coordinates of each feature.

Global Positioning System (GPS)

More recently, some agencies use GPS to automate the collection of road segment coordinates or individual point coordinates for roadway characteristics and events. These coordinates represent a 'point in space' or in this context, a spatial location relative to a specific point on the Earth's surface. A number of jurisdictions have noted that collection of LRM information in addition to the coordinates has helped to maintain the quality of location data.

Some agencies may have had difficulty integrating the legacy LRM data with new GPS coordinates due to variations in required scale for individual data systems and errors in original data. In some cases, though rarely documented, maintenance personnel have physically relocated mileposts without notifying other key stakeholders, resulting in a degradation of location data quality. Several States have developed complex algorithms or other data matching and integration tools (called *conflation* techniques) to merge the various LRMs while attempting to reduce the extent of error that may result in specific datasets.

Additional methods for denoting locations that usually are not used in a LRM include the following:

Roadway Stationing

Locations on a roadway may include the *station* information from the construction plans. For most State and local agencies, the station is a survey location denoting a distance and direction from a known point. Usually agencies convert this information to calculate milepost locations expressed in thousandths of a mile. Even with this degree of accuracy, location error can potentially be expected when comparing this information with GPS coordinate data for the same physical location.

Roadway Geo-coding / Address Interpolation

Agencies obtain geo-coding by capturing a street or road address from the field, entering that data into interpolation software or a web-based site that returns latitude and longitude coordinates. Accuracy may vary considerably using this method. However, Ohio was able to develop a highly accurate unified map data. Ohio's Location Based Response System (LBRS) is a model for other States considering a system. A White Paper on the LBRS is available at: http://gis3.oit.ohio.gov/%5Clbrs/_downloads/docs/White%20Paper-LBRS_2011.pdf.

Roadway Realignment

Location coding methods indicate how a spatial location is identified for purposes of associating attributes to a specific location. When a section of roadway is no longer located in the original space, adjustments must be made to the location scheme to indicate the new location of the roadway and, in some cases, to historical data that references that location.

Where agencies realign a roadway to shorten the road (e.g., when a curved section is straightened), the new roadway will have a gap in the mileposts (sometimes referred to as a gap equation). Figure 14 shows the resulting mileposts ahead (AH, which usually denotes the original mileposts) and back (BK, referring to new mileposts):

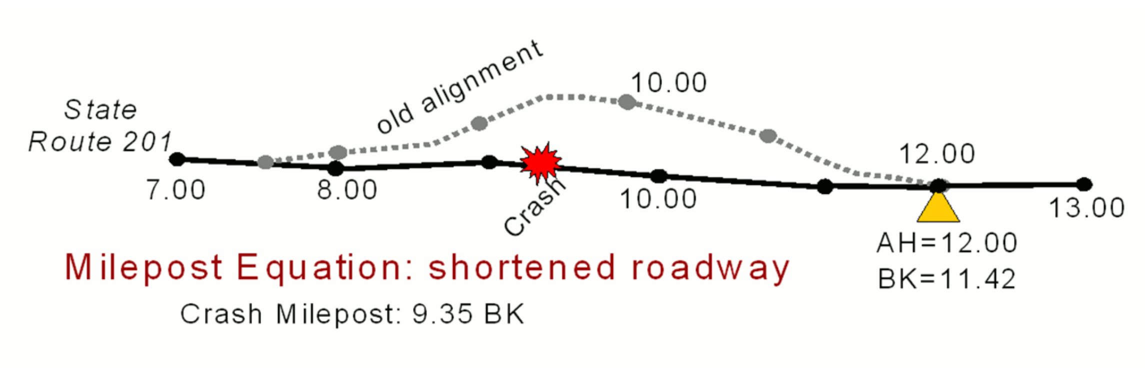


Figure 14. Example of a realignment that shortened a roadway.

With the new shortened alignment, the mileposts from 11.42 to 12.00 no longer exist. The crucial item is the *temporal* or time element that must be preserved to denote when the change to the roadway occurred and which segment is relevant to the features involved. This also applies to the scenario illustrated in Figure 15 that shows a circumstance where the road realignment is longer than the original, resulting in duplicate mileposts on the new roadway. The designation for ahead or back is critical, and must be added for the entire length where duplicate milepoints exist.

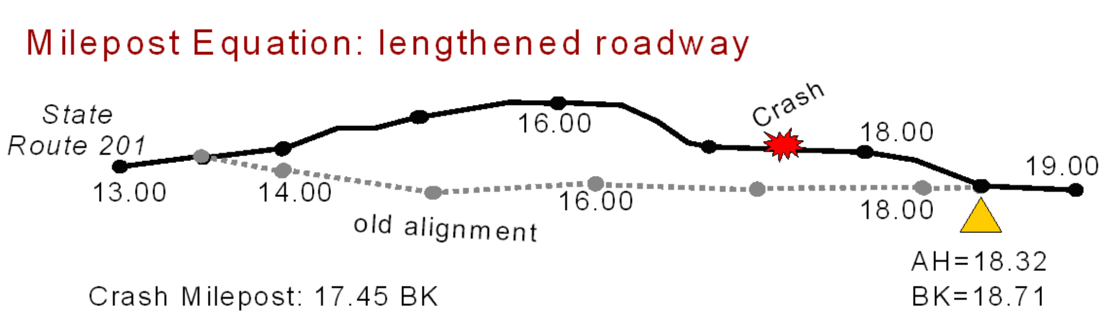


Figure 15. Example of a realignment that lengthened a roadway.

With the variety of LRMs, the next logical step is creating a linear referencing system that manages the various LRMs into a cohesive product that can be used in conjunction with GIS software.

Many transportation network specialists use a LRS approach where attributes are related to roadway segments in terms of linear measurements from defined locations (e.g., intersection, etc.). This technique reduces the need to re-segment the network geometry when new attributes are added or changes are made. Figure 16 graphically depicts a segmented road model. Nodes define the limits of each segment; i.e., where characteristics of the roadway change. Note in Figure 16 that the beginning and endpoints of the bridge are delineated with nodes.

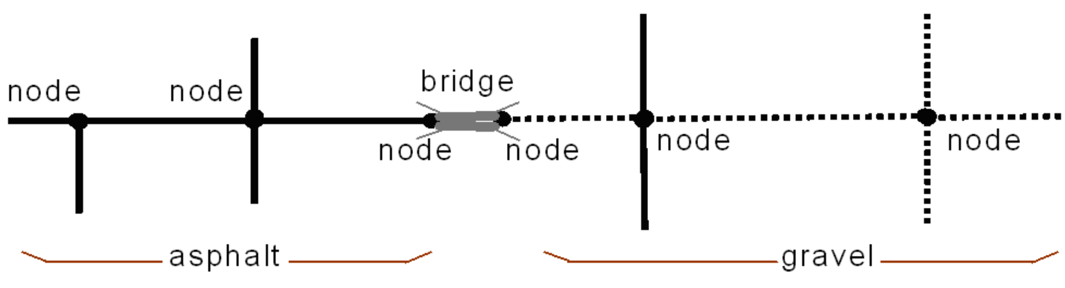


Figure 16. Example of a segmented road model.

Figure 17 depicts a roadway segment using LRS. In this example, the bridge is represented as a linear event that occurs along the roadway, rather than needing new nodes and a roadway segment to be defined. A more detailed explanation of this topic and graphical examples of road segmentation models are available in the American National Standards Institute (ANSI) Standard for IT. For information that is more detailed see *Geographical Information Framework Data Content Standard, Part 7: Transportation Base,* (http://www.fgdc.gov/standards/projects/incits-l1-standards-projects/framework/documents-1/Part_7%20-%20Transportation-Base%20-%2020060112-1846.pdf).

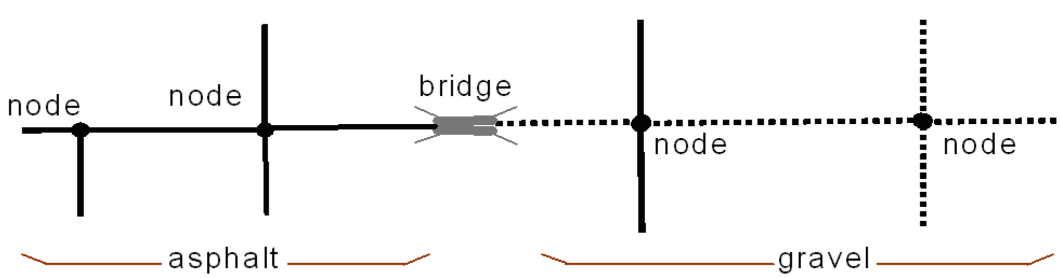


Figure 17. Example of a linear referencing system (LRS).

Geographic Information System (GIS)

The roadway data component of a traffic records system must provide a framework that defines the roadway network and provides a method for describing locations and characteristics of points and segments of the roadway system. With the widespread use of GIS software, the ability to define a roadway network and provide a means to link to locations on that network has become more comprehensive and accurate than ever before. Using a GIS graphical representation to cross-reference various location-coding methods and define roadway attributes makes even legacy data easier to report, analyze, and understand. Figure 18 shows an example of geographical analytical tools.

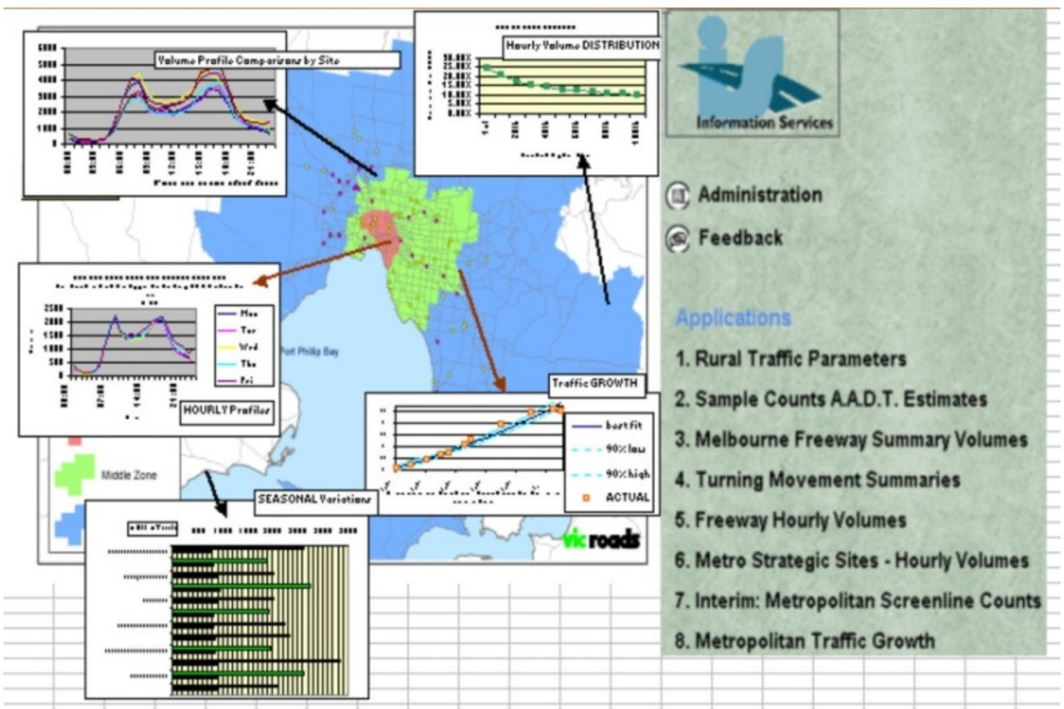


Figure 18. Example of geographical analytical tools.

For over 30 years, State highway departments have worked with GIS software to define their roadway networks in sufficient detail to use for engineering activities. A GIS joins a computer software mapping application with a database of geographically enabled information that links spatial information (i.e., where things are) with demographic and environmental information (i.e., what things are like). Basic data elements and terminology for defining the roadway network include the following:

Shapefile

A shapefile is one format of a digital geo-referenced file that supports point, line, and area features to enable a "…user to apply various geographical analytical tools, such as interpolation, distance measurements, or data visualization." At a minimum, a shapefile will include a coordinate pair, an index number, and an attribute.

Coordinates

Coordinates are used to determine the relative position of points within a survey area or, in many cases, with respect to a much larger area. In order to make the coordinate system usable for engineering projects, the horizontal relationships should be defined as two dimensional on one (mapping) plane. The north-south axis is designated as the Y-axis and the east-west axis as the X-axis. The horizontal distance from the Y-axis is the "easting" coordinate and the vertical distance from the X-axis is the "northing" coordinate. A statement of the exact position of a point within the system can be expressed as X-Y, or easting-northing value of the point. Stated either way, the point location is being described by rectangular coordinates and the location is a grid location. For spatial analysis (using GIS tools), a Z-coordinate can be used to add a vertical third dimension to the analysis to express elevation.

Cartesian Coordinates System

Cartesian coordinates represent a two-dimensional planar coordinate system in which x measures horizontal distance and y measures vertical distance.

Coordinate Pair

A coordinate pair is a set of Cartesian coordinates (north-south and east-west distance, or latitude and longitude) that describe the two-dimensional location in terms of the earth's surface, from a known reference point.

Latitude

Latitude measures angles in a north-south direction and defines the y-coordinate of a Cartesian coordinate pair.

Longitude

Longitude measures angles in the east-west direction and defines the x-coordinate of a Cartesian coordinate pair.

Map Scale

A statement of a measure on a map and the equivalent measure on the earth's surface, often expressed as a representative fraction of distance; for example, 1:24,000 (one unit of distance on the map represents 24,000 of the same units of distance on the Earth). Map scale also can be expressed using different units: for example, 1 inch = 1 mile or 1 inch = 2,000 feet.

Attributes

Attributes are the characteristics of a geographic feature described by numbers, characters, images and drawings, typically stored in tabular format and linked to the feature by a user-assigned identifier or index.

There are numerous methods used to reference a roadway segment or section. An example provided by Illinois DOT includes:

- Annual program number.
- Multi-year program number.
- Design project number.
- Construction contract number.
- Federal job number.
- Local agency reference.
- Structure number.
- Rail crossing inventory number.

To integrate the various methods of locating data for safety program development, Illinois DOT uses the following methods for their enterprise spatial data integration:

- Crash data (safety marked route).
- High accident locations (paper maps).
- Program (program number).
- Ongoing construction (contract number).
- Roadway inventory (key route).
- Rehabilitation history (marked route).
- Effectiveness of safety countermeasures.
- Relationship to total program.

Locating data about a roadway segment may vary from state to state and among jurisdictions, but one or more method for identifying a roadway location is needed to develop safety programs and conduct safety analyses.

Figure 19 shows a GIS map example of Iowa's use of multiple referencing systems. Iowa DOT has the ability to cross-reference data easily to locate and analyze the effects of safety-related improvements. Iowa's mapping includes the capability to be used by cities for utilities and engineering, by counties for the tax assessor and engineering, by other state agencies such as Emergency Management, and by the Federal government for FHWA and non-transportation agencies such as the Census Bureau.

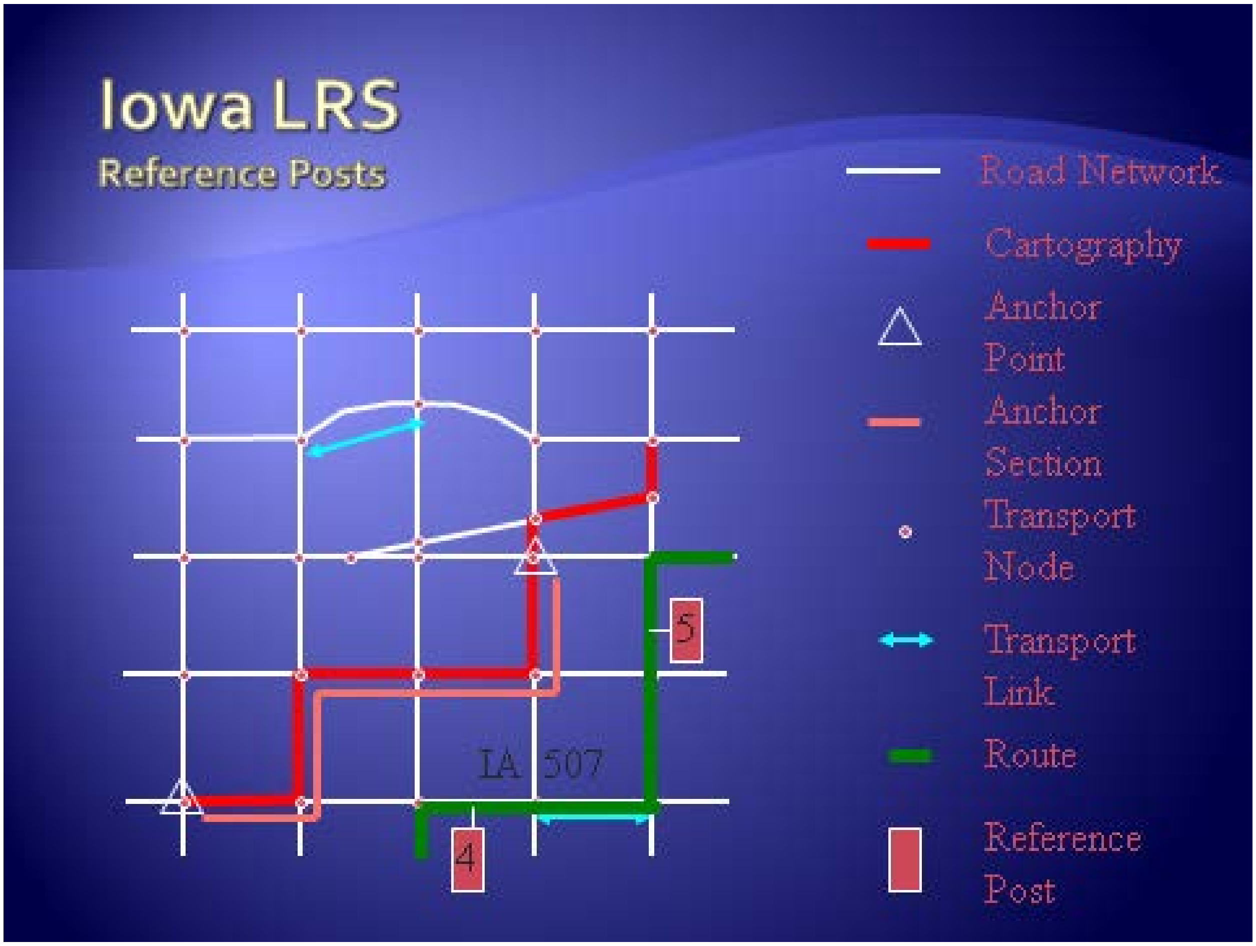


Figure 19. Iowa DOT example of using multiple location methods.

With the structure of modern databases designed to define roadway networks comes the ability to manage increasingly large amounts of data. Databases to support detailed analysis of roadway features and attributes are easily designed. Implementing and maintaining such systems is the more difficult task because it requires data. The next chapter presents common and emerging data collection tools and methods useful in populating the databases we use to construct the roadway network and support safety analysis.

Several NCHRP Synthesis reports on highway location reference methods have been published by AASHTO over the years. An example includes NCHRP Synthesis Report 460, *Guidelines for the Implementation of Multimodal Transportation Location Referencing Systems* (http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rpt_460.pdf). This report presents and describes the transportation multimodal, multidimensional location referencing system data model developed through NCHRP Project 20-27. A more recent report, that builds on this earlier work, is the NCHRP 20-07 Task 302 Report, *Multi-level Linear Referencing System (MLLRS) Cost/Benefit Value Analysis Study* (http://onlinepubs.trb.org/onlinepubs/nchrp/docs/NCHRP20-07(302)_FR.pdf).

CHAPTER 6—DATA COLLECTION TOOLS AND METHODS

Application of technology has been suggested as a way to overcome many of the limitations surrounding current safety data. As noted in the NCHRP 367 report *Technologies for Improving Safety Data* (http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_syn_367.pdf) incorporating technology will assist DOTs in meeting the current and future needs of managing safety on our road systems, with the following points identified:

- Technology alone cannot solve all of the problems associated with safety data systems, especially those related to inadequate institutional cooperation. Organizational issues should be addressed before considering technological advancements.
- Technology implementation and maintenance can be capital-intensive, requiring significant funding and programmatic support; therefore, the benefits and costs clearly must be evident across all affected agencies.
- Technologies are constantly evolving; therefore, agencies should seek to employ technologies that allow for flexibility. Additionally, a practical plan for maintaining and upgrading the technologies, as well as assessing their continued effectiveness, should be developed before initial investment and implementation of the technology.
- No single technology will allow for the collection and maintenance of *all* the varied data required for safety analysis — hence, an array of technologies should be considered.

This overview of technology is an introduction to the possibilities of collecting accurate, comprehensive, timely, and complete roadway inventory data that can be integrated into current GIS and is needed for the use of safety analysis tools such as IHSDM, *Safety Analyst*, and the HSM.

Roadway Photolog

The typical method of collecting roadway data until the 1970s was through 'windshield surveys' and manual data collection. These efforts are very time intensive and the resulting information is generally poorly distributed, incomplete, and lacking continuity. In the late 1970s, some DOTs adapted the concept of 'photo logging' that usually consisted of a 35 mm photo frame, taken every 52.8 feet (1/100th of a mile or 100 frames per mile), from a vehicle traveling on State roadways. As each photo was taken, usually a date stamp, highway identifier, and location by milepost were included in the frame. Photolog offers a photo inventory of the highways that can be referenced without an additional field trip to locate or verify an inventory item at a specific location.

Agencies used the resulting film images to determine general surface conditions and locations of signs, signals, driveways, and other roadside items or hazards. The storage and indexing of the film reels was cumbersome and time consuming, and usually limited to one agency office. The difficulties in storing and accessing 35 mm film images led to the use of video home system (VHS) technology and other analog storage such as laser optical platters providing for improved storage and reproduction. Now digital photography has replaced VHS, with the advantages of digital storage for faster and easier retrieval by users, typically over an agency's intranet. Several States make their imagery available to the public via the internet. States are now collecting digital images in high definition (HD) video, higher resolution imagery for improved clarity.

The digital photolog or Videolog systems can include additional cameras (as many as 18) and images that allow users a much wider view of the roadway and images of the road surface for pavement evaluation, etc. Agencies update the roadway photolog images typically every two to three years, with archival images retained for two to three cycles since storage is less expensive and easier to access in digital format. While some States obtain Videologs using internal staff, they also use commercial Videolog services.

Automating the Roadway Inventory

In the early 1980s, some of the early photolog systems included compass gyros, adapted from maritime and aviation use, along with other sensors to collect roadway data elements. As technology evolved, this equipment has become vastly more sophisticated and includes such devices as differential GPS, accelerometers, laser measurement, inertia detection, and other sensors to collect data such as:

- Sign type, location, and distance to the edge of the road.
- Sign nighttime reflectivity.
- Roadside barriers and other roadside hardware.
- Pavement roughness, rutting, texture, faulting or cracking, reflectivity.
- Road geometry (heading, grade, cross-slope).
- Lane width, shoulder width.
- Location via GPS coordinates.

The consistency and timeliness of information collected on roadway and pavement condition in conjunction with accurate GPS data are key to an effective pavement management system.

Agencies can also use data collected to determine horizontal (roadway curvature) and vertical alignment (hills and terrain) information to be used to generate roadway profiles for safety analysis. To date, automated collection of curve and elevation data has not achieved acceptable levels of accuracy or consistency.

Some DOTs purchase roadway Videolog vehicles with the roadway inventory equipment, while others contract for these services from commercial vendors. Although the fully equipped vehicle with a two-person crew is expensive to operate, the cost of obtaining and maintaining the resulting data is far less expensive than comparable manual pen and paper data collection.

Figure 20 illustrates a typical vehicle equipped with video logging and additional sensors for automated roadway inventory data collection. The van is fully equipped with cameras, computer equipment, and lasers to capture pavement conditions and road geometry.

Figure 20. A typical van used to collect the video images.
(http://www.fhwa.dot.gov/publications/research/safety/05073/chapt2.cfm)

LIGHT DETECTION AND RANGING (LIDAR)

LiDAR-equipped vans are capable of obtaining highly detailed digital images of roadway features and attributes. Agencies use the data for asset management, safety analysis and other purposes. Depending on how they are configured, such data collection systems can automatically collect x,y,z spatial coordinates along with images of the roadway surface, pavement width, shoulder width and type, median width and type, signs, pole locations and types, roadside hardware locations and types, and other information. The high-resolution images support automated and manual measurement of features. Surveys can be ground-based by adding LiDAR instrumentation to vans as that shown in Figure 20, as in Figure 21 at low altitude. Figure 21 shows downward-looking LiDAR lasers that have been mounted temporarily to the underside of a helicopter.

Figure 21. Aerial low-altitude LiDAR.
(*http://www.fhwa.dot.gov/publications/publicroads/01septoct/lidar.cfm*).

GLOBAL POSITIONING SYSTEM (GPS)

Example of a PDA/Phone with GPS.

A GPS is a satellite-based radio-beacon navigation system developed, owned, and operated by the U.S. Government. GPS uses a constellation of 24 satellites that transmit time signals continuously. Users equipped with the appropriate receivers can receive signals from the satellites to calculate their position (latitude, longitude, and elevation), time, and velocity. Originally, GPS accuracy for civilian use registered only within approximately 100 meters. This intentional degradation of the accuracy was a security measure instituted by the U.S. government for any non-military use of the system. This degradation was removed as of May 1, 2000, bringing accuracy of the basic GPS data to approximately 15 meters. Though GPS use is continuing to flourish, the accuracy of the basic civilian-accessible implementation of GPS often is insufficient for applications needed for transportation requirements.

To achieve sufficient accuracy, agencies can use the augmentation technique commonly known as the Differential Global Positioning System (DGPS). The DGPS technique is based on a highly accurate geodetically surveyed location of a GPS reference station. The reference station observes GPS signals in real time and compares their ranging information to the ranges expected to be observed at its fixed location. The differences between observed ranges and predicted ranges are used to compute differential corrections, which then are provided to GPS users. DGPS accuracy is approximately 1 m with some newer technology approaching 10 cm accuracy.

Each transportation department's use of DGPS is based on its own particular needs; however, there is a common thread among transportation applications. This technology is used to improve public and personnel safety as well as efficiency.

There are wide ranges of transportation applications using DGPS, including:

- Geographic databases for use in emergency 911 systems.
- Highway inventory (i.e., signs, milepost markers, right of way, guardrail, and bridges).
- Emergency response services (e.g., police, fire, and rescue).
- Automatic vehicle location for public transit and other fleets.
- Snowplow guidance for low-visibility situations.
- Inventory of highway-railroad crossings and road centerline.
- Tracking hazardous materials from origin to destination.
- Mapping pavement condition, safety, accident, and traffic data.

The integration of DGPS with GIS is one of the most widely used applications. GIS allows the association of statistics of any kind with a specific geographic location and the displaying of the data on an interactive map. The GIS provides a "base map" on which to display geocoded information and the DGPS provides an accurate physical location for each data point. For a list of States that are using DGPS, see the FHWA summary report *An Investigation of the Use of Differential Global Positioning System (GPS) Technology and its Augmentations within State and Local Transportation Departments*, see *http://www.fhwa.dot.gov/publications/research/operations/its/00093.pdf*. Agencies within many States are considering or are already planning to incorporate GPS and DGPS technologies into a variety of applications.

Aerial Photography and Satellite Imagery

The term "remote sensing" was first used in 1960, but aerial photography was first used 100 years earlier with cameras hung from balloons, kites, and pigeons, for a variety of tasks relating to planning and land use. Remote sensing is:

> "…the measurement or acquisition of information of an object by a recording device that is not in physical or intimate contact with the object. In practice, remote sensing is the use at a distance (as from aircraft, spacecraft, satellite, or ship) of any device for gathering information about the environment. Satellite and aerial photography images are the two most commonly used forms of remote sensing in transportation."

The value of remote sensing from aerial photography and satellite is evident from the availability and usage of internet mapping websites that allow map and photographic views of the Earth. These maps allow for simple identification of specific addresses, businesses, type of business, and directions (including mileage and travel time) from one specific location to another. Some of the commercial mapping sites illustrate the significance of mapping and imagery combined in the 'hybrid' views, and the coordination of aerial views with ground images via DGPS correlated Videologs.

Aerial photograph showing a four-way intersection
(*http://www.ops.fhwa.dot.gov*)

Within GIS, agencies can analyze digital satellite images generated through remote sensing to produce a map-like layer of digital information about a transportation network and the surrounding land use. The quality of remote sensing data depends on its spatial resolution. Spatial resolution refers to the size of a pixel that is recorded in a raster image (e.g. photograph image). Typically, pixels may correspond to square areas. Most digital ortho photos are based on 3 inch, 6 inch, or 12 inch pixels. This precision has been improved by decreasing data collection cost due to the widespread use of digital cameras. Remote-sensed data can be used in a similar manner to the Videolog to identify elements of the transportation network. With advancements in technology, Videolog and satellite imagery are increasingly being geo-referenced; therefore, elements are inherently being tied to spatial coordinates and are readily placed on a map or integrated into a GIS.

Traffic Counters

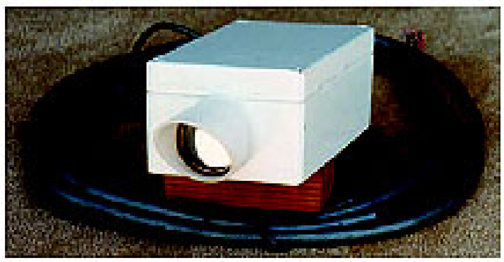
Collecting traffic information is vital to the daily operations of many groups within a DOT, and safety is no exception. Traffic volumes can be used for many purposes in safety analysis. For example, traffic volumes allow agencies to determine expected crash rates for a road segment or intersection based on similar operational characteristics. Agencies can use this information to rank roadways and intersections by priority for receiving safety funding.

Traffic counters are classified into two basic groups: (1) intrusive and (2) non-intrusive. Intrusive detectors are the most common and are embedded into the road surface or lay exposed on the surface. These include the loop detector, the pneumatic road tubes, bending plates, and piezoelectric sensors. These devices usually last three or more years, require detouring traffic to install the device, and involve compromising the road surface for installation (except for temporary pneumatic tube counters). These devices also are not always accurate, as they usually do not accurately count motorcycles and bicycles. Current methods for automatically collecting pedestrian volumes can be expensive and inaccurate.

Non-intrusive detectors usually do not include road surface installation. Examples include:

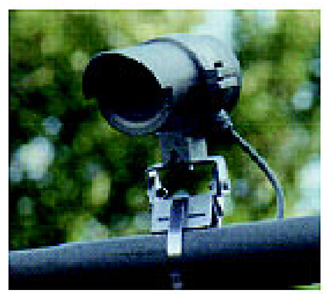
Passive Infrared

Passive infrared devices detect the presence of vehicles by comparing the infrared energy naturally emanating from the road surface with the change in energy caused by the presence of a vehicle. Figure 22 provides an example of passive infrared devices. Since the roadway may generate either more or less radiation than a vehicle depending on the season, the contrast in heat energy is what is detected.



Eltec 842 passive infrared vehicle presence sensor. [*Source*: L.A. Klein, *Sensor Technologies and Data Requirements for ITS* (Norwood, MA: Artech House, 2001)].

ASIM IR 250 series passive infrared sensor. This multizone sensor performs vehicle counting, speed measurement, classification by length, and presence detection. (Photograph courtesy of ASIM Technologies, Uznach, Switzerland).



Siemens Eagle PIR-1 sensor. Performs vehicle counting, stop line presence detection, occupancy detection, and queue detection. (Photograph courtesy of Siemens ITS, Austin, TX).

Figure 22. Examples of passive infrared devices.

Active Infrared

Active infrared devices detect the presence of vehicles by emitting a low-energy laser beam(s) at the road surface and measuring the time for the reflected signal to return to the device. The presence of a vehicle is measured by the corresponding reduction in time for the signal return.

Magnetic -- Passive and Active

Passive magnetic devices measure the change in the earth's magnetic flux created when a vehicle passes through a detection zone. Active magnetic devices, such as inductive loops, apply a small electric current to a coil of wires and detect the change in inductance caused by the passage of a vehicle. Figure 23 illustrates how one passive magnetic device works.

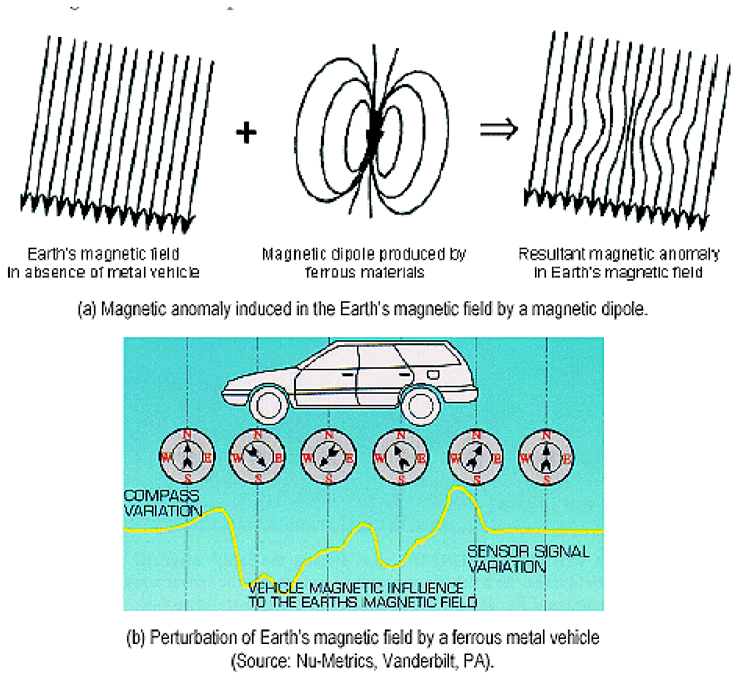


Figure 23. Example of passive magnetic device.

Microwave -- Doppler, Radar, and Passive Millimeter

Doppler microwave devices transmit low-energy microwave radiation at a target area on the pavement and then analyze the signal reflected back to the detector. According to the Doppler principle, the motion of a vehicle in the detection zone causes a shift in the frequency of the reflected signal. This can be used to detect moving vehicles and to determine their speed. Radar devices use a pulsed, frequency-modulated, or phase-modulated signal to determine the time delay of the return signal, thereby calculating the distance to the detected vehicle. Radar devices have the additional ability to sense the presence of stationary vehicles and to sense multiple zones through their range finding ability. A third type of microwave detector, passive millimeter, operates at a shorter wavelength than other microwave devices. It detects the electromagnetic energy in the millimeter radiation frequencies from all objects in the target area.

Passive Acoustic

Passive acoustic devices consist of an array of microphones aimed at the traffic stream. The devices are passive in that they are listening for the sound energy of passing vehicles.

Ultrasonic -- Pulse and Doppler

Pulse devices emit pulses of ultrasonic sound energy and measure the time for the signal to return to the device. Doppler devices emit a continuous ultrasonic signal and utilize the Doppler principle to measure the shift in the reflected signal.

Video

Video devices use a microprocessor to analyze the video image input from a video camera. Two basic analysis techniques are used: tripline and tracking. Tripline techniques monitor specific zones on the video image to detect the presence of a vehicle. Video tracking techniques employ algorithms to identify and track vehicles as they pass through the field of view. The video devices use one or both of these techniques.

In field tests, some of these detectors did not function well at intersections and accuracy may be compromised by weather and road conditions as they depend on detected heat, sound, or radio waves. The video image camera detectors may suffer from accuracy problems and require installment expertise and setup time, but offer the capability of collecting additional traffic data beyond standard counts. For detail about these capabilities, see the FHWA report: *Field Test of Monitoring of Urban Vehicle Operations Using Non-Intrusive Technologies* (*http://ntl.bts.gov/lib/jpodocs/repts_te/6665.pdf*).

ADDITIONAL DATA COLLECTION SYSTEMS FOR ROADWAY INFORMATION

Agencies can collect some of the data for roadway databases using automated systems. For example, this report discusses various types of roadway data collection sensors that may be installed on a data collection van. Agencies collect traffic volume data with a combination of automated, semi-automated, and manual systems. The most advanced systems use permanent detectors (typically embedded in the roadway, but some systems rely on external sensors) tied to software that aggregates data for determining hourly and daily traffic counts as well as speeds and vehicle classifications. Agencies still widely use older technology (such as temporary installations of road tubes with digital or even pneumatic counters), especially at locations where the counts are collected for a short duration on an annual or even longer-term schedule.

Agencies still collect some roadway data manually. Agencies may perform these efforts using paper-and-pencil forms or using a variety of field data collection devices including software implemented on a personal digital assistant (PDA), tablet, handheld data collection device, or a field data collection computer. The software products used in capturing these data, either remotely or by transcribing manual data into a system, share some common features. These features include:

- Databases designed specifically to store a particular type of data - roadway inventory, traffic volume, pedestrian/bicycle counts, sign inventories, etc.
- Tools for managing the data - record entry/acceptance, validation checks, sorting, editing, duplication detection, record deletion, etc.

- Output generators to provide data extracts, as well data aggregation or analytic reports.

The most advanced of these tools and software are typically capable of collecting GPS coordinates for each record entered. The coordinates may be tied directly to a statewide or departmental GIS in order to support automatic interface and linkage with other relevant sources of roadway data through standard location coding and mapping methods.

With roadway data collection technology comes the opportunity, in comparison to older methods, to collect more data at substantially lower costs, usually in a more timely manner and often with improved precision and accuracy. It is true, however, that while these methods offer savings over options that would collect the same data manually, the choice facing most departments is not between two alternative methods of obtaining the data but rather whether or not to collect the data at all. It is usually much easier to document the cost of data than it is to prove its benefit. The FHWA Office of Safety conducted a study to develop guidance on methodologies that State and local DOTs could implement to make the case for investing in data collection, data systems and processes. A Guidebook was developed to assist States in justifying the decision to invest in additional data collection efforts related to safety. If a State is uncertain of the value of data collection, or if a State is having difficulty justifying the allocation of resources to data collection projects, the Guidebook provides instruction in how States may assess the potential impact of investment in safety data improvement. The Guidebook may be accessed at http://safety.fhwa.dot.gov/rsdp/downloads/guidebook.pdf.

It is important to recognize the central role that data quality plays in determining both the cost and benefit of data for decision-making. Not all data collection methods are equal in terms of guaranteeing that the resulting dataset is reliable. In addition, when a State wants to justify the move to more advanced (and perhaps initially expensive) methods of data collection, it is important to be able to quantify the improvements that will result. Data quality improvements are important to quantify for decision-makers because they demonstrate the reliability of the data. The benefits of more reliable data are not lost on decision-makers even if those benefits are sometimes difficult to quantify as precisely as one would like.

CHAPTER 7—ROADWAY DATA QUALITY MEASUREMENT

The management of the roadway systems development process to ensure data quality can be organized into three parts:

- Quality planning: selecting standards and how to meet them.
- Quality assurance: evaluations during project development.
- Quality control: monitoring project results and improving performance as needed.

QUALITY PLANNING

The data quality decisions may vary from State to State depending on their priorities. Quality planning involves deciding what level of quality to expect in the resulting system. Examples of these data standards are provided in the final section of this document (Roadway Safety Data Quality Measures). These standards for the system should be planned into the design of the system from the earliest possible point in the design cycle.

As part of the quality planning effort, the design team would also conduct a benefit/cost analysis for achieving the desired level of quality. Assessing the downside of increased quality can be a difficult task for designers unfamiliar with roadway data reporting and analysis in general, or for those system users who are unfamiliar with system development. It is important, therefore, that the system design team include personnel from both the information technology and the user management functions so that decisions of how high to set the quality bar can be made in an informed manner and the resulting system is sustainable over time.

FHWA Office of Safety developed methodologies to estimate the costs and benefits of investing in data and data systems for safety. These methodologies are documented in *Benefit-Cost Analysis of Investing in Data Systems and Processes for Data Driven Safety Programs Decision-Making Guidebook* (http://safety.fhwa.dot.gov/rsdp/downloads/guidebook.pdf).

QUALITY ASSURANCE

The quality assurance process includes those techniques used by the system development team to ensure that the system meets expectations. At this point in development, the design is complete and the system is being developed and tested before it is implemented. Quality assurance continues after implementation as well. The measurements that apply to the system's performance may not seem familiar to those used to managing data quality, but they are very important for making sure that the system itself is not causing data problems.

For example, does the roadway system design allow access by all users who have a responsibility of analyzing safety data and setting project priorities? A measure of what percentage of the critical users has flexible access for analysis is a quality assurance metric. It tells system managers if there is a problem with the system accessibility. Such "down time" may not have a measurable impact on more traditional data quality measures (e.g., timeliness, accuracy, completeness, etc.) but can have a major impact on how well the system is perceived by users and thus how willing they are to continue using it. Following system implementation, the system developers and managers are the primary users of the quality assurance measures. These measures allow tracking of how well the hardware and software are performing and lets managers know if there is a problem that needs to be addressed. They also can serve as an early warning of problems that could affect data quality – the bottom line for most of the users of a system.

Quality Control

For data managers and users, quality control is the most familiar part of the project planning process. Quality control is the set of measurements and procedures put in place to ensure that the data quality is meeting expectations. The measures of data quality can cover a wide variety of issues at a wide range of "levels" – from global indicators of overall quality to micro-level indicators of the validity of data in one particular field of the database. Quality control processes are the responses of the system (the software and the people working with it) to quality problems that arise. For example, data quality metrics would show if location data were not meeting expectations (e.g., location coding is not precise enough or features cannot be matched to a location on the roadway network). The quality control processes are the response to these questions – what do the data managers and collectors do about the problems.

Additional detailed information on quality measures are provided in *Chapter 8, Roadway Safety Data Quality.*

Benefit/Cost

The two most desired types of system-level performance measures for the roadway systems are those that would relate the availability of the data to lives saved; and those that would accurately track the costs maintaining the system. In other words, measurements of the benefits and costs associated with maintaining a complete and accurate record of the roadway-related data used for safety analyses. Operational departments (e.g. operations, infrastructure etc.) can also realize benefits in reduced costs and less duplication of data.

On the surface, it would seem that the cost of collecting and maintaining roadway-related data would be an easy metric to define and obtain. However, the recent trend has been to distribute the responsibility of various roadway-related data to those users most closely involved in the collection and use of specific data. For example, the planning division may establish the base network from design plans, change orders, and as-built plans, while inventory updates may be made in the maintenance or other divisions. These distributed costs make it more difficult to obtain a reliable cost estimate.

Much of the roadway-related safety data is initially collected for other uses, which makes the definition of costs and benefits more difficult to distinguish from the State DOT's routine operational roadway databases. A standard method for explicitly tracking the system cost might include the following components:

- Time spent collecting initial data (if not already collected for existing databases).
- Transmission of imports of data from existing systems.
- Initial software purchase and implementation or in-house development.
- Annual maintenance including any licensing and support.
- Separate line items for life-cycle costs of the hardware and software.
- Total number of roadway segments and features obtained electronically and the percentage that represents of total data collection for the entire system.

FHWA's *Benefit-Cost Analysis of Investing in Data Systems and Processes for Data Driven Safety Programs: Decision-Making Guidebook* (http://safety.fhwa.dot.gov/rsdp/downloads/guidebook.pdf) provides a methodology for such analyses.

Customer Service

States may wish to measure the impact of their delivery of service to the users of the roadway and other safety-related data. Customer service, in particular, may be affected by the use of a decision support system that consolidates safety-related data and a State may wish to measure how much better able it is to meet customers' needs (both internal to the agency and external customers) once data are available electronically. For example, the proportion of data requests met via a web portal (thus requiring little or no direct staff time) could be measured and the time it takes to deliver data following a customer request could be measured. States wishing to show the full impact and utility of their system will measure customer service as well as costs.

Michigan provides online maps that are integrated with the Web and available to local jurisdictions. Figure 24 provides examples of data and maps.

Figure 24. Michigan DOT sharing maps on the web.

CHAPTER 8—ROADWAY SAFETY DATA QUALITY

The key to roadway-related safety analysis is location-based linkage of three basic data sources:

- Roadway inventory (to provide detailed descriptions of locations and features).
- Traffic volume (to provide level-of-use statistics and measures of exposure).
- Crash data (for estimates of safety performance of locations and attributes).

Until recently, the standard for safety data analysis was a simple rank ordering of locations based on their crash frequency, crash rate (frequency of crashes divided by AADT), or a weighted ranking using both measures. This analysis still is useful to provide engineers and other safety professionals with a good sense of where to look in the roadway network. With access to more detailed data describing locations and an increased ability to merge data from multiple data sources, more effective safety analyses can be conducted.

States can create separate rank orderings of roadways that belong to a particular type. For example, rather than identify the "top 10" roadways overall, engineers might look at intersections, interchanges, highways, and county routes separately and compare among similar roadway types. In this way, they could examine a series of "top 10" lists for each of the roadway location types of interest. With the development of CMFs, engineers have started to focus on the attributes of roadway locations including:

- Type of intersection/interchange.
- Number of lanes and lane width.
- Presence or absence of traffic control, turn lanes, and other features.

The availability of detailed data describing locations means that engineers can look at roadway location types on several dimensions and determine the most likely safety concerns and the predicted crash frequency for each. Modern network screening techniques, as described in the HSM, take advantage of richer data sources and increased ability to merge inventory, traffic, and crash data.

To assess a State's ability to conduct safety analyses, data quality performance measures are needed for each of the key components: roadway inventory, traffic volume, and crash. For the crash data elements, there are several examples of data quality measurements including the FHWA *Crash Data Improvement Program Guide* (http://safety.fhwa.dot.gov/cdip/finalrpt04122010/). There are, unfortunately, few sources of comprehensive data quality measurements for roadway inventory and traffic volume. In the discussion that follows, measures are presented that have been derived from the available resources including:

- FHWA's *Performance Measures for Roadway Inventory Data*.

- NHTSA's Program Advisory on Traffic Records Systems.
- NHTSA's 2011 *Model Performance Measures for State Traffic Records Systems.*
- NHTSA's Traffic Records 101 web-based training.
- Expert panel suggestions.

Data Quality Characteristics

The NHTSA Program Advisory for Traffic Records suggests six data quality characteristics that should be measured. These are defined below.

Timeliness

Information should be available within a specific timeframe to allow for meaningful analysis of the status of the issue under investigation (e.g., as soon as a roadway is open to the public, data about that roadway is available on the roadway inventory database; traffic data are updated prior to compilation of annual statistics in support of problem identification).

Accuracy

Information within the database should be correct and reliable in describing the data element it purports to describe. Accuracy typically is enhanced through the practice of conducting consistency checks and validations on the data being entered into the database. For roadway data, this may involve field audits, validations against other data sources (including an agency GIS, as-built plans, etc.), as well as consistency checks to identify unexpected changes in crash frequency or severity distributions for specific locations.

Completeness

Information within the database should be complete for all roadways that are open to the public, whether designated State routes, county roads, or city streets. For any selected roadway location, each relevant data element should be available within the database and all required data elements within the record should be completed with appropriate responses.

Consistency/Uniformity

Information collected should be consistent among all reporting jurisdictions with all reporting jurisdictions using valid data collection methods (e.g., traffic counts, lane measurements, right-of-way definitions). Ideally, information will be reported using nationally accepted and published guidelines and standards (e.g., MIRE Version 1.0) to define each data element.

Integration

Information in one database should be capable of being interfaced or joined with information from other databases through use of common data elements called linking variables. An example of integration is the linkage of information from a statewide crash database with data in a roadway inventory file to support problem identification, countermeasure selection/design, and program evaluation.

Accessibility

Information within the database should be readily available to all eligible users. States differ in how much and what types of roadway information they make available and to whom. A well-functioning system will support the State DOT as well as its local engineering partners with the data needed to perform their engineering-related duties. Access for other partners depends on a variety of factors including cost, as well as concerns over tort liability, data complexity, and analytic competence of the users.

Data Quality Metrics

States are recommended to develop performance measures to assess the quality of the data being entered and maintained in their roadway, traffic, and crash databases. Developing performance measures to assess the timeliness, accuracy, completeness, and consistency of the data within the databases can help to identify deficiencies and allow states to develop and implement actions to improve the data quality. Developing additional performance measures for the roadway, traffic and crash databases can permit a state to assess how well and completely the information in these databases integrates with other databases and how accessible the data are to eligible users.

The following tables provide examples of potential data quality metrics for each of the three key databases for safety analyses: roadway inventory (Table 2), traffic volume (Table 3), and crash data (Table 4). The source of each metric is indicated in the columns, along with a brief definition of the measurement. The metrics are organized according to the data quality attributes defined above. Metrics for integration may apply to all of the systems. These metrics expand upon those provided in the FHWA *Performance Metrics for Roadway Inventory Data* (http://safety.fhwa.dot.gov/rsdp/downloads/performancemeasures.pdf) that incorporates NHTSA performance metrics, to give practitioners with additional examples of potential metrics. Additional information on crash data metrics can be found in the FHWA CDIP Guide (*http://safety.fhwa.dot.gov/cdip/finalrpt04122010/*). It is up to each agency to decide what metrics are most appropriate for them.

Table 2. Roadway inventory data quality metrics.

Metric	Description/Examples
Timeliness Metrics	
Update Cycle	Measures the timing of updates to the file (continuous, annual, and other). · Percent of records updated in a year. · Percent of records reviewed in a year.
Update Delays	Delay from change in "ground truth" to posting of the change in the inventory file. · Number of days from roadway change is complete to posting update. · Number of days from completion of a periodic review to posting update. · Percentage of records updated electronically.
Accuracy Metrics	
Features Locations	Measures variance between ground truth and the coded location of each feature. · Centerline location is within 10 feet of actual. · Cross-section data elements are accurate w/in +/- 1 ft.
Source of Data	Measures the proportion of the database derived from accurate sources. · Percent of records based on "as built" plans. · Percent of records based on field measurements.
Error Rate	Measures accuracy of data in individual records and data elements within a record. · Percentage of records with zero errors in predefined "critical" data elements.
Completeness Metrics	
Roadways Included	Measures the percentage of defined locations that are included in the database. · Percentage of State-maintained roadway locations included in the inventory. · Percentage of local roadway locations included in the inventory. · Percentage of intersections/interchanges included in the intersection inventory. · Percentage of ramps included in the inventory.
Data Fields Collected	Measures the percentage of location records with data present in key fields. · Percentage of State roadway records with zero missing data in key fields. · Percentage of local roadway records with zero missing data in key fields. · Percentage of key fields coded "unknown." · Percentage of roadway locations with defined geographic coordinates.

Table 2. Roadway inventory data quality metrics (cont'd).

Metric	Description/Examples
Consistency/Uniformity Metrics	
Standards Compliance	Measures the number of data elements collected. · Percentage of data elements collected for segments. · Percentage of data elements collected for intersections. · Percentage of data elements collected for ramps. · Percentage of data elements collected for State & HPMS segments. · Percentage of data elements collected for local road segments.
Inter-jurisdiction Consistency	Measures the compliance with State standards for location inventory reporting across all reporting jurisdictions (DOT divisions, locals, MPOs, etc.). · Percentage of jurisdictions reporting. · Percentage of jurisdictions reporting complete data for key data elements. · Percentage of jurisdictions reporting within accuracy tolerances.
Year-to-Year Comparisons	Measures the percentage of records with unexplained/anomalous changes in the record from one year to the next. This could also be considered an accuracy measure.
Integration Metrics	
Macro-Level Integration	Measures the number of databases that are linked. This is an overall indication of the number of files that can be merged. Each linkage of two files counts as one integration. For example, linking crash & roadway inventory = 1 linkage.
Micro-Level Integration	Measures the strength of linkage between two or more files. This is a measure of the proportion of records that *should be linkable* that, in fact, are linked in the resulting file. · Percent of crash records with a corresponding roadway inventory record. · Percent of traffic count records with a corresponding roadway inventory record. · Percent of crash records actually linked to roadway inventory records. · Percent of traffic count records actually linked to roadway inventory records.

Table 2. Roadway inventory data quality metrics (cont'd).

Metric	Description/Examples
Automated versus Manual Linkage	Measures the ability to created merged datasets linking two or more data sources using completed automated and automated plus manual processes. Percent of crash and roadway inventory records linked without manual intervention. Percent of crash and roadway inventory records linked after all processing is completed. Percent of traffic volume and roadway inventory records linked without manual intervention. Percent of traffic volume and roadway inventory records linked after all processing is completed.
Accessibility Metrics	
External Agency Access	Measures the number of non-DOT users (and agencies) with access to the location-based safety data.
Internal DOT Users	Measures the number of internal users within DOT who have access to the location-based safety data.
Data Access and Analytic Requests	Measures the requests for data and analysis. This measure is a count of the number of requests that the data managers and analysts receive and complete in a period. · Number of requests for data extracts; and number fulfilled. · Number of requests for analytic results and number fulfilled. · Average time from request to completion.
User Satisfaction	Measures (via a survey) the number of users "satisfied" with the data and/or analyses available to them.

Table 3. Traffic volume data quality metrics.

Metric	Description/Examples
Timeliness Metrics	
Update Cycle	Measures the timing of updates to the file (continuous, annual, and other). · Percent of records updated in a year. · Percent of records reviewed in a year.
Update Delays	Delay from data collection to availability of volume data for analysis (i.e., the time it takes to post the raw data and any required calculated values such as AADT).
Accuracy Metrics	
Features locations	Percentage of records with "estimated" volume counts.
Source of Data	Measures the proportion of the database derived from various sources. · Percent of records based on permanent counters. · Percent of records based on temporary (3-day? Other?) counts. · Percent of records based on locally-provided data.
Completeness Metrics	
Roadways Included	Measures the percentage of defined locations that have traffic count data available. · Percentage of State-maintained roadway locations with actual counts. · Percentage of local roadway locations with actual counts. · Percentage of ramps with actual counts.
Consistency/Uniformity Metrics	
Interjurisdiction Consistency	Measures the compliance with State standards for traffic volume reporting across all reporting jurisdictions (DOT divisions, locals, MPOs, etc.). · Percentage of jurisdictions reporting. · Percentage of jurisdictions reporting complete data. · Percentage of jurisdictions in compliance with data collection standards.

Table 3. Traffic volume data quality metrics (cont'd).

Metric	Description/Examples
Integration Metrics	
Macro-level Integration	See the metrics discussed in Table 2: Roadway Inventory.
Accessibility Metrics	
External Agency Access	Measures the number of non-DOT users (and agencies) with access to the location-based traffic volume data.
Internal DOT users	Measures the number of internal users within DOT who have access to the location-based traffic volume data.
Data access and analytic requests	Measures the requests for data and analysis. This measure is a count of the requests that data managers and analysts receive and complete in a period. · Number of requests for data extracts; and number fulfilled. · Number of requests for analytic results and number fulfilled. · Average time from request to completion.
User Satisfaction	Measures (via a survey) the number of users "satisfied" with the data and/or analyses available to them.

Table 4. Crash data quality metrics.

Metric	Description/Examples
Timeliness Metrics	
Overall Timeliness	Measures the time from the crash event to posting of the data in a database ready for analysis. This is the sum of submission timeliness and processing timeliness components defined below.
Submission Timeliness	Measures the time from crash event to receipt of the report by the custodial agency. This is the "delay" contributed by law enforcement agency processes including data collection, report review, and report submission. It should be reported for each law enforcement agency, along with the statewide average.
Processing Timeliness	Delay from receipt by the custodial agency to posting the data in a database ready for analysis. This may be measured in individual components as well as overall. · Number of days for creation of an image archive record. · Number of days for data entry. · Number of days for location coding. · Number of days for post-processing quality control review.
Accuracy Metrics	
Edit Checks	Measures the initial and final accuracy of crash reports against the standards set for edit checks and quality validation. · Percent of crashes with 1 or more fatal errors at time of submission. · Percent of crashes with 1 or more "warning" errors at time of submission. · Percentage of crashes with 1 or more fatal errors when accepted as final. · Percentage of crashes with 1 or more "warning" errors accepted as final.
Location Accuracy	Measures the proportion of crashes that can be "landed" using the statewide location coding system. · Percent of crashes "landed" using automated processes only. · Percent of crashes "landed" when accepted as final.

Table 4. Crash data quality metrics (cont'd).

Metric	Description/Examples
Validation Against Driver and Vehicle Files	Measures the accuracy of data fields that can be validated against the driver and vehicle records. Percentage of in-State driver's license numbers with a corresponding driver history record. Percentage of in-State vehicle identification numbers (VINs) with a corresponding vehicle registration record. Percentage of VINs that decode accurately using standard VIN decoding software. Percentage of crashes where the decoded VIN information disagrees with the vehicle make/model/type coded on the crash report form.
Validation Against Injury Surveillance Data	Measures the accuracy of (in particular) injury codes collected on the crash report form. · Percentage of crash-involved persons with a mismatch between the injury severity score on the crash report and the injury severity score recorded in the trauma registry and/or Emergency Department or hospital discharge databases. · Percentage of crash reports coded with an incorrect overall crash severity based on a comparison of to the injury surveillance system injury codes.
Completeness Metrics	
Overall Completeness of Key Data Fields	Measures the percentage of reports with no missing data in key data fields. This usually is measured as an average statewide and separately for each individual law enforcement agency that submits crash reports.
Overall Missing Data	Measures the percentage of crash reports with no missing data. · Percent of crash reports with no (key) fields left blank. · Percent of crash reports with no unexplained or inappropriate use of "unknown" or "n/a" values.
Narrative and Diagram Completeness	Measures the percentage of reports with acceptably complete narratives and diagrams.

Table 4. Crash data quality metrics (cont'd).

Metric	Description/Examples
Consistency/Uniformity Metrics	
MMUCC Compliance	Measures the percentage of MMUCC data elements and attributes collected · Percentage of MMUCC data elements collected/defined on the crash report form or through linkage. · Percentage of MMUCC data attributes (values within data elements) collected/defined on the crash report form or through linkage.
Interjurisdiction Consistency	Measures the compliance with State standards for crash data collection. · Percentage of jurisdictions reporting all reportable crashes – this is based on a comparison of the fatal+injury/total crashes reported · Percentage of jurisdictions with no unexplained/unexpected drops in the number of crashes reported in comparison to prior years.
Integration Metrics	
Macro-Level Integration	Measures the number of databases that are linked. This is an overall indication of the number of files that can be merged. Each linkage of two files counts as one integration. For example, linking crash & roadway inventory = 1 linkage. Crash data may also be integrated with injury surveillance, driver, vehicle, and other databases.
Micro-Level Integration	Measures the strength of linkage between two or more files. This is a measure of the proportion of records that *should be linkable* that, in fact, are linked in the resulting file. · Percent of crash records with a corresponding roadway inventory record · Percent of crash-involved drivers with a corresponding driver history record. · Percent of crash-involved injured persons with a corresponding injury surveillance record. · Strength of linkage as measured using the Crash Outcome. Data Evaluation System (CODES) or other probabilistic linking software system.

Table 4. Crash data quality metrics (cont'd).

Metric	Description/Examples
Accessibility Metrics	
Overall Access	Overall Access
Data Access and Analytic Requests	Measures requests for data and analysis. This measure is a count of the number of requests that the data managers and analysts receive and complete in a period. Number of requests for data extracts; and number fulfilled. Number of requests for analytic results and number fulfilled. Average time from request to completion.
User Satisfaction	Measures (via a survey) the number of users "satisfied" with the data and/or analyses available to them.

CHAPTER 9—USE OF ROADWAY DATA IN SAFETY ANALYSIS

The FHWA provides and supports a wide range of data and safety analysis tools for State and local practitioners. These tools have been designed to assist practitioners in understanding safety problems on their roadways, link crashes to their roadway environments, and select and apply appropriate countermeasures. The tools' capabilities range from simple to complex. Some tools provide general information, while others allow more complex analysis of crashes under specific conditions and/or with specific roadway features.

Data Collection and Management Tools

Model Inventory of Roadway Elements (MIRE)

Safety data are key elements to sound decisions on the design and operation of roadways. Critical safety data include not only crash information but also roadway inventory and traffic data. MIRE is a data dictionary that includes a listing of roadway inventory and traffic elements deemed essential to safety management and proposes standardized coding for each. Additional information on MIRE is available at http://safety.fhwa.dot.gov/tools/data_tools/mirereport/.

Highway Safety Information System (HSIS)

HSIS is a multi-State database that contains crash, roadway inventory, and traffic volume data for a select group of States. The participating States were selected based on the quality of their roadway, traffic and crash data available and their ability to merge data from these various files. The HSIS is used to analyze a large number of safety problems, ranging from the more basic "problem identification" issues to modeling efforts that attempt to predict future accidents from roadway characteristics and traffic factors. The HSIS is used in support of the FHWA safety research program and as input to program and policy decisions. The HSIS is also available to analysts conducting research under the NCHRP, university researchers, and others involved in the study of highway safety. Additional information on HSIS is available at *http://www.hsisinfo.org/*.

Safety Data Analysis Tools

Highway Safety Manual (HSM)

The HSM provides information and tools to assist transportation professionals in making decisions that have a positive impact on highway safety. HSM data elements are provided in Table 5. Focusing on objective measures of safety with a primary emphasis on crash frequency and severity, the HSM includes analytical tools to quantify and predict the safety performance of a variety of elements considered in road planning, design, maintenance, construction, and operation. The HSM also includes a synthesis of validated highway research and procedures that are adapted and integrated into practice. Additional information on the HSM is available at *http://www.highwaysafetymanual.org/*.

HSM

Table 5. HSM data elements.

Crash Data

- Date (year)
- Location
- Type
- Severity level
- Relationship to intersection (at-intersection, intersection related, not intersection related)
  distance from the intersection

Roadway Characteristics and Traffic Volume Data

Roadway Segments

- Area type (rural/suburban/urban)
- Annual average daily traffic volume
- Length of roadway segment
- Number of through lanes
- Lane width
- Shoulder width
- Shoulder type
- Presence of median (divided/undivided)
- Median width
- Presence of concrete median barrier
- Presence of passing lane
- Presence of short four-lane section
- Presence of two-way left-turn lane
- Driveway density
- Number of major commercial driveways
- Number of minor commercial driveways
- Number of major residential driveways
- Number of minor residential driveways
- Number of major industrial/institutional driveways
- Number of minor industrial/institutional driveways
- Number of other driveways
- Horizontal curve length
- Horizontal curve radius
- Horizontal curve superelevation
- Presence of spiral transition
- Grade
- Roadside hazard rating
- Roadside slope
- Roadside fixed-object density
- Roadside fixed-object offset
- Percent of length with on-street parking
- Type of on-street parking
- Presence of lighting

Intersections

- Area type (rural/suburban/urban)
- Major-road average daily traffic volume
- Minor-road average daily traffic volume
- Number of intersection legs
- Type of intersection traffic control
- Left-turn signal phasing (if signalized)
- Presence of right turn on red (if signalized)
- Presence of red-light cameras
- Presence of median on major road
- Presence of major-road left-turn lane(s)
- Presence of major-road right-turn lane(s)
- Presence of minor-road left-turn lane(s)
- Presence of minor-road right-turn lane(s)
- Intersection skew angle
- Intersection sight distance
- Terrain (flat vs. level or rolling)
- Presence of lighting

Interactive Highway Safety Design Model (IHSDM)

IHSDM is a suite of software analysis tools for evaluating the safety and operational effects of geometric design decisions. Table 6 shows the IHSDM data requirements. The current version checks existing or proposed two-lane rural highway designs against relevant design policy values and provides estimates of a design's expected safety and operational performance. Future expansion of IHSDM crash prediction capabilities to include rural multilane highways and urban/suburban arterials is planned to match the *1st Edition Highway Safety Manual*. Intended users include highway project managers, designers, and traffic and safety reviewers in State and local highway agencies and engineering consulting firms. IHSDM currently includes five evaluation modules – crash prediction, design consistency, intersection review, policy review, and traffic analysis. Additional information on the IHSDM is available at http://www.ihsdm.org/.

Fundamental Data Elements (FDE)

In 2012, the FHWA Office of Safety issued MAP-21 Guidance on State Safety Data Systems (http://www.fhwa.dot.gov/map21/guidance/guidesafetydata.cfm). The guidance identifies a subset of MIRE, referred to as the fundamental data elements (FDE), that, "when integrated with crash data, enables States to conduct a sufficient safety analysis to identify safety problems and make more effective investment decisions." The guidance recommends that, "States should collect the FDEs on all public roads as soon as practicable in order to benefit from improved analyses as soon as possible."

Safety Analyst

Safety Analyst is an AASHTOWare™ product that can assist State and local highway agencies to program site-specific highway safety improvements for implementation. Table 7 shows the *SafetyAnalyst* data requirements. Additional information on *SafetyAnalyst* is available at http://www.safetyanalyst.org/.

Table 7. *Safety Analyst* data elements.

Safety Analyst

Roadway Segment Characteristics Data

- Segment number
- Segment location (in a form that is linkable to crash locations)
- Segment length (mi)
- Area type (rural/urban)
- Number of through traffic lanes (by direction of travel)
- Median type (divided/undivided)
- Access control (freeway/nonfreeway)
- Two-way vs. one-way operation
- Traffic volume (AADT)

Crash Data

- Crash location
- Date
- Collision type
- Severity
- Relationship to junction
- Maneuvers by involved vehicles (straight ahead/left turn/right turn/etc.)

Ramp Characteristics Data

- Ramp number
- Ramp location (in a form that is linkable to crash locations)
- Area type (rural/urban)
- Ramp length (mi)
- Ramp type (on-ramp/off-ramp/freeway-to-freeway ramp)
- Ramp configuration (diamond/loop/directional/etc.)
- Ramp traffic volume (AADT)

Intersection Characteristics Data

- Intersection number
- Intersection location (in a form that is linkable to crash locations)
- Area type (rural/urban)
- Number of intersection legs
- Type on intersection traffic control
- Major-road traffic volume (AADT)
- Minor-road traffic volume (AADT)

PEDESTRIAN AND BICYCLE SAFETY TOOLS

Pedestrian and Bicycle Crash Analysis Tool (PBCAT)

The Pedestrian and Bicycle Crash Analysis Tool (PBCAT) is a software application designed to assist State and local pedestrian and bicycle coordinators, planners, and engineers address pedestrian and bicyclist crash problems. PBCAT helps users create a database of details associated with crashes between motor vehicles and pedestrians or bicyclists, analyze the data, produce reports, and select countermeasures to address problems identified. Additional information on PBCAT is available at *http://www.bicyclinginfo.org/facts/pbcat/index.cfm.*

Pedestrian Safety Guide and Countermeasure Selection System (PEDSAFE)

The Pedestrian Safety Guide and Countermeasure Selection System (PEDSAFE) provides practitioners with the latest information available for improving the safety and mobility of those who walk. This online tool gives users a list of possible engineering, education, and/or enforcement treatments to improve pedestrian safety and/or mobility based on user input about a specific location. Additional information on PEDSAFE is available at *http://www.fhwa.dot.gov/research/deployment/pedsafe.cfm.*

Bicycle Countermeasure Selection System (BIKESAFE)

The Bicycle Countermeasure Selection System (BIKESAFE) provides practitioners with the latest information available for improving the safety and mobility of those who bicycle. BIKESAFE's resources include an overview of bicycling in today's transportation system and information about bicycle crash factors and analysis and selecting and implementing bicycling improvements. BIKESAFE's tools allow users to select appropriate countermeasures or treatments to address specific bicycling objectives or crash problems. Additional information on BIKESAFE is available at *http://www.bicyclinginfo.org/bikesafe/.*

INTERSECTION/INTERCHANGE SAFETY ANALYSIS TOOLS

Interchange Safety Analysis Tool (ISAT)

The Interchange Safety Analysis Tool (ISAT) provides design and safety engineers with an automated tool for assessing the safety effects of basic geometric design at typical existing interchanges and adjacent roadway network. ISAT also can be used to predict the safety performance of design alternatives for new interchanges and prior to reconstruction of existing interchanges. The primary outputs from an analysis include the number of predicted crashes for the entire interchange area, the number of predicted crashes by interchange element type, the number of predicted crashes by year, and the number of predicted crashes by collision type. The ISAT tool is available in Excel spreadsheet format. Additional information on ISAT is available at *http://www.fhwa.dot.gov/publications/research/safety/07045/07045.pdf.*

Surrogate Safety Assessment Module (SSAM)

The safety of intersections, interchanges, and other traffic facilities most often is assessed by tracking and analyzing police-reported motor vehicle crashes over time. Given the infrequent and random nature of crashes, this process is slow to reveal the need for remediation of either the roadway design or the flow-control strategy. This process is also not applicable to assess the safety of roadway designs that have yet to be built or flow-control strategies that have yet to be applied in the field.

The Surrogate Safety Assessment Module (SSAM) is a technique combining microsimulation and automated conflict analysis, which analyzes the frequency and character of narrowly averted vehicle-to-vehicle collisions in traffic, to assess the safety of traffic facilities without waiting for a statistically above-normal number of crashes and injuries actually to occur. Additional information on SSAM is available at
http://www.fhwa.dot.gov/research/tfhrc/projects/projectsite/safety/ssam/index.cfm.

CHAPTER 10—SUMMARY

This *Roadway Data Improvement Program: Supplemental Information Resource* is intended to serve as resource document for practitioners in a State DOT and the allied agencies that provide or use roadway data. The primary focus of the RDIP is on the data systems used for safety decision making, and in particular, those that would support engineering efforts such as network screening (problem identification), diagnosis, countermeasure selection, and program evaluation. The RDIP is focused on the quality of data in these systems with particular attention to the six key attributes of timeliness, accuracy, completeness, consistency, integration, and accessibility of roadway inventory and traffic volume data. This *Resource* includes information on data quality as well as various programmatic uses of the data for safety analysis.

This RDIP *Supplemental Information Resource* synthesizes information from a broad range of topic areas on highway transportation. Roadway data systems are described along with their associated data collection tools and methods (including automated data collection). Uses of the data for planning, design, and safety improvement are discussed along with the various options for data management and overall system design. Examples of data quality metrics and safety analyses are provided as well.

Users of this *Resource* are encouraged to approach it as a convenient synopsis of the engineering highway safety field. Its focus on current practices provides assurance that every tool, technique, or technology presented here is in use and has produced results for at least one State. Moreover, the presentation of state-of-the-practice techniques for safety data analysis gives a State a comparison point for an internal review of its own systems.

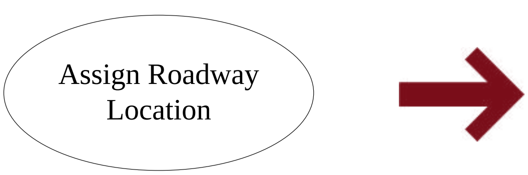
RDIP technical assistance team (TAT) site visits are available to help the State. FHWA selects subject matter experts who will travel to a state, conduct workshops, and develop recommendations tailored to the specific needs of that State. The TAT leaves a matrix of recommendations with the State at the end of the site visit, and then follows up with a full report containing descriptions of the status of a State's systems and recommendations for improving the quality of the data in those systems.

APPENDIX A—EXPLANATION OF DATA FLOW DIAGRAMS

A DFD is a graphical means of describing the flow of data through a system that produces, changes, or stores data. The DFD is a visual depiction of an automated or manual system from the perspective of the data. The DFD view uses the four graphic shapes defined below to describe the system. A DFD differs from a flowchart in its perspective: a flowchart describes a system of programs or manual processes that manipulate the data, while a DFD describes a system from the perspective of the flow of information. A DFD is drawn from most general to most specific in a series of levels of detail. Each page, or diagram, describes a part of the system with the first page DFD showing the most general picture of the data and its flow through the entire system. Any further pages (or "lower levels") of the DFD provide progressively more detail about the processes being described.

ENTITY - This box represents an entity that is involved in the flow of data. Entities are ordinarily a person, group of persons, an agency, etc.

DATA STORE - The rectangular box represents a data storage file(s) incorporating the data that are flowing into it. These file(s), or data stores, can be computerized or manual. The number is unique for each data store throughout all of the DFDs to ensure that each data store, or file, is specifically identified.



PROCESS - This oval represents a process or activity that is performed on the data flowing into it. The process is used when the data is actually changed or enhanced in such a way that it is important to show in the DFD. The change in the data can be effected by either an automated or a manual procedure. The asterisk signifies that more specific detail about this process may be found on a lower level page of the DFD.

DATA FLOW - This arrow represents a flow of data into, or out of, an entity, data store, or process. The designer labels the data flow arrow when necessary to describe clearly the specific type of data in the flow.

APPENDIX B—EXAMPLE SYSTEM BUSINESS FUNCTIONS

To assist States in developing and integrating the MIRE into a management information system structure that will provide greater utility in collecting, maintaining, and using MIRE data, FHWA has undertaken the MIRE Management Information System (MIRE MIS) project. It is important to be able to link statewide crash data to the MIRE data elements for use in safety improvement identification and analysis. A MIRE MIS will enable users to merge roadway and traffic data with crash data to enhance data analysis capabilities. The following table shows the high-level business requirements that a developer must satisfy when developing the MIRE MIS.

MIRE MIS Functions
1.0 General
1.1 Address key State-approved performance indicators. The suite of performance indicators will vary among States, but there are three main types of performance measures that can be supported: a. Software/system service levels: these are indicators of the function of the software itself, useful in benchmarking and management of users. Examples include tracking the number of records processed, proportion of records rejected, number of users, and system downtime. These measures would also be useful in tracking the cost of the system, including the costs associated with obtaining and storing the data. b. System performance: these are comprised of data gathered to support agencies' indicators of performance in key areas (such as safety, mobility, asset management, etc.). Indicators that could be useful in linking the availability of MIRE-compliant data to improved safety (lives saved, injury levels reduced, and crashes avoided) would be included in this set of measurements. c. Data quality: these are measures of the timeliness, accuracy, consistency, and completeness of the data in the MIRE MIS
1.2 Work within context and framework of State-specific approach. The MIRE MIS must fit within and support a State's existing business practices. For example, States define and combine data across segments in a variety of ways. The MIRE MIS will require flexibility to support multiple segmentation methods including those defined by changes in roadway features as well as those defined by distance, or a

combination of both. Varying methods of aggregating data across segments must be accommodated including aggregation based on similar features, by route, by contiguous segments, rolling averages, and others. In essence, the segmentation in the core roadway inventory file will be mirrored by MIRE MIS.

2.1 Using a modeling tool such as an Entity Relationship Diagram or other modeling tools, document State-specific information required for creation of the database.

Modeling tools are useful from the IT perspective in that they provide a succinct view of the data, data sources, and the interactions among those sources. State-specific entity relationships are incorporated into the data model to guide the database design and insure that all necessary inputs, outputs, and linkages are supported.

2.2 Develop a technical diagram that identifies the groupings of data and elements used by name and description.

The technical diagram complements the system's data dictionary and the data model described in the Entity Relationship Diagram by showing the flow of data through the system and the major groupings of data by sources and outputs.

2.3 Develop a technical diagram that defines the relationship between the data groups and correctly structures (normalizes) the data.

Data relationships are described as one-to-one; one-to-many; many-to-one; and many-to-many depending on how records from one source "map" to records in another source. To take a trivial example, crashes have a many-to-one relationship with any particular segment in the database. In a fully normalized database, each data element is present once, but can (by virtue of its specific relationships) be accessed through multiple paths.

2.4 Document the location and structure of system integration points both for import and export data relevant to the MIRE MIS.

As part of identifying the sources and outputs for the MIRE MIS, all of the points for importing into and exporting data out of the system are defined in terms of the interface requirements, the point in the system's process the import or export occurs, how the operations run (on demand, automated batch processing, as a step in the records management processes at the source, etc.), and how often.

3.0 Local Data Validation and Entry

3.1 Provide ability to accept both original data entry and import from other data sources.

In addition to data import points, the system must accommodate manual data entry both from a central location and via remote access (e.g., web-based users with permission to add records). The processes for both importing data electronically and manual data entry will be documented.

3.2 Provide ability to validate both electronic transmission and direct data entry based on defined business rules.

Systems evolve over time. It is anticipated that all implementations of MIRE MIS will begin with a set of data validations arising from the MIRE data dictionary as well as any State-specific data definitions that must be accommodated. Over time, as States identify errors that could be trapped through automated means, it is anticipated the number and complexity of validation checks will increase. The MIRE MIS will be designed with a basic set of error checks, but it will also accommodate the addition of new edit checks as defined by the users.

3.3 Provide electronic notification for critical and non-critical data errors.

Individual records will be flagged whenever any validation rule is violated. These rules are generally grouped by severity into critical (often called "fatal errors") and non-critical ("warnings"). Most commonly, fatal errors cause a record to be rejected—such records are held in a pending database awaiting correction and resubmittal. Errors generating a warning message may be allowed to remain uncorrected—the record is passed into the production database and updated if a correction is received. All errors are logged so that they may be analyzed and used for aggregate reporting of data quality measurements of accuracy, completeness, and consistency.

3.4 Establish communication protocol or use existing secure protocols based on industry standard mechanisms.

The MIRE MIS must be capable of one- and two-way communication with various system users, including data collectors, as well as interface with a variety of other systems. These communications typically require a secure connection between the user/source and the centralized system both to insure data integrity and to protect sensitive information from release. MIRE MIS will be designed to accommodate existing secure communications protocols in States where these exist, but it will also require some new communication protocols to be established in some States.

3.5 Provide standard templates for data entry and validation through browser interface.

The MIRE MIS data entry function will be web-based. Screens that open in a browser on a secure website will be built around a standardized template for all MIRE MIS implementations (customizable to meet a State's specific needs). Validation (including edit checks and interfaces with other systems) will take place during data entry and immediately upon submission of the data through the web. Data quality feedback (edit check results) will be provided to the user via the same web interface.

3.6 Provide notifications to log and track data input, updates, and deletions.

System management reports available to the custodian of the MIRE MIS records and IT support staff will track user's inputs, updates, and deletions. Error logs will also be stored to support management reporting of data quality.

3.7 Provide data collection tools via browser, such as GIS, mobile mapping.

The user interface for data collectors will be web-based. This interface will include standard screens for data entry as well as screens for users to access assistive technologies such as the GIS or map-based location validation and input.

3.8 Provide electronic submission for added digital or scanned attachments.

Images may be appended to any record during data entry or updating. Users will have the option of directly uploading scanned or digital images into an image archive in the MIRE MIS. Because the data and images are associated, data managers and users will have the ability to search images in the database using a variety of search terms limited only by their access to the data elements in the database.

3.9 Provide the ability to accept data from other local or State databases

MIRE MIS will accept data from local or other State roadway inventory files for which an interface has been defined or provided. This serves to reduce the amount of data that must be entered manually into the system and reduces the cost of data collection to the State by taking advantage of already existing sources—thus reducing redundancy.

3.10 Document and publish standards for data collection and transmission.

In order to ensure consistent and error-free data, all entries into the system must be validated prior to acceptance into the production database. A single unified set of validation rules and data definitions will be supplied to all data collectors. In addition, a package for those wishing to implement electronic data sharing (uploads and downloads) will also be provided with data transmission guidelines and require file structures. A testing protocol for acceptance of electronic data will also be developed and shared with these external sources of data.

4.1 Provide an imaging system that contains a database that can provide access to all supporting documentation associated with any roadway segment.

The imaging system supports users by making visual evidence available for review. Because each image is tied to the database record for a particular segment, the MIRE MIS also serves as an indexing system for sorting and retrieving of the images as well. While not intended to replace or function as a Photolog system, the image system in MIRE MIS can fill the need for storing of both scanned documents as well as digital photos captured in the field. This can support multiple uses including in-depth engineering analysis (e.g., condition diagramming) as well as asset management.

4.2 Receive and validate quality of submitted digital images.

Digital images can be most simply validated by checking the final image record size against the original (as submitted). Other methods may also be used including checksums, and use of image “wrapper data” which can be examined for both a match to the original submission and already existing records in the MIRE MIS database. In rare cases (such as poor initial image quality), human intervention—a manual review of the images—may be required. These are more a matter of State agency policy and procedures than a function of MIRE MIS; however, the MIRE MIS will support a review of images by the staff.

4.3 Scan and index paper documents received from local agencies into a direct-access storage medium, creating digital images of reports and all supporting documentation provided.

In addition to digital photographs, the MIRE MIS image archive will be able to collect and store scanned images (of documents, hard-copy photos, or other materials obtained from local agencies) through digital scanning. As this process will be managed centrally by the MIRE MIS data steward agency staff, it assumed that indexing information will be added (i.e., image "wrapper" information) at the time the materials are scanned. This action will allow the scanned images to be added to the image archive and treated in the same manner, from that point forward, as are the digital images received from external sources (i.e., electronically transmitted to MIRE MIS).

4.4 Provide a unique identifier for each record, including for records received from local agencies.

There are two types of unique identifiers that may be implemented in MIRE MIS: system-assigned and State assigned. MIRE MIS will assign a unique identifier to each record received as a function of adding the record to the production database. In addition, each State may have its own record identifiers that it can assign (through manual or automated processes) to the records as they arrive. MIRE MIS could track local agencies' record numbers but these are not assumed to be "unique" in that multiple agencies could conceivably assign the same record numbers to data in their own stand-alone systems. These record numbers are retained so that the State may use them in future contacts with the local agencies (e.g., to discuss a particular record or roadway segment).

4.5 Establish a process for returning manual or electronic data to local agencies for correction and track receipt of corrected submissions.

The MIRE MIS error logging and tracking function is designed to support State data quality management processes fully. Each error in each record is recorded as part of the "as submitted" data. Records with a critical (fatal) error are automatically rejected (i.e., not added to the production database) and are held in a pending status until those serious errors are corrected or (at the State's option in some cases) over-ridden. Non-critical errors (i.e., those that generate a warning only) may be handled in a number of ways (at the State's option). Typically, non-critical errors are noted in the record but the record itself is not barred from addition to the production database. States may choose to hold these records in a pending status awaiting correction or not. In any case, the MIRE MIS supports notification of the submitting agency/staff of both critical and non-critical errors through case-level and aggregate reports. Case-level reports are specific to an individual record and typically are used to identify a record that contains a critical error that must be corrected. Aggregate reports summarize the errors noted from each data submitter over some defined period (e.g., monthly, quarterly, annually). The latter are a useful source of feedback to the submitting agencies and may be helpful in designing training content or revising data collection instructions/manuals.

Finally, MIRE MIS retains a log of all errors and all reports that contained an error. A tracking system is established that uses this information to generate reminder notices for any errors that have not been corrected by a specified time (e.g., 2 weeks after notification). These notices are provided to the system's administrators automatically and may be sent electronically to the submitting agencies responsible for correcting errors.

4.6 Establish automated workflow routings and work queues for data transmitted from other systems (e.g., PMS, bridge, crashes).

MIRE MIS supports agency workflows by notifying designated staff when a record matching an established routing definition is added to the system. These may be useful in a number of situations. One example is to establish work queues for internal staff (e.g., location coders, data quality managers, etc.). Another example may be to provide data from the roadway inventory files for external users such as the FARS analysts who may need accurate roadway segment descriptions to add to their records of fatal crashes.

4.7 Develop MIRE MIS application and relational database.

This is the main system development task and takes place after all the functions and State-specific modifications have been described.

4.8 Capture Graphical User Interface (GUI) input screens that enable key from image or paper processes if needed; including split screen.

This task results in the main data entry system and process control screens for centralized scanning. It supports the paper-management process at the agency serving as the steward of statewide roadway inventory data. The same (or similar) data entry screens may be made available to external users in a form suitable for web-based remote entry of records.

4.9 Establish an intuitive flow for data entry and include features such as highlighting, table driven drip down lists, pre-populated fields, capturing system dates, etc.

The MIRE MIS data entry process will support high quality data entry through user assistance and constraints. Pick lists, for example, serve to constrain users' inputs in comparison to entries made into a free text field so that only allowable information (data that will pass basic system edits) may be entered. Highlighting of specific field on the data entry helps users to identify errors and mandatory data fields so that these may be addressed during initial data entry rather than waiting until the report is submitted.

4.10 Provide ability to flag records and data elements based on Federal and State rules, close out year rules, etc. (e.g., data used for HPMS, data used for *SafetyAnalyst*, data on-system/off-system).

These flags maybe treated as a specific type of validation rule which results in mandatory corrections (if the data as entered do not meet the reporting requirements) and for work-flow routing (e.g., flagging relevant records for the staff responsible for HPMS, HSIS, etc.)

4.11 Provide ability to search/access any field or digitized documents.

In the search/access function, the entire coded portion (all database contents will the exception of images) of the MIRE MIS can (potentially) serve as the indexing system. This means that users with the appropriate access permissions could search on any field and obtain records matching a specified search criterion. The search function will support both simple and complex queries. Simple queries are those that make use of one or a limited number of filters. Complex queries are those that may require the use of a query builder tool or a query model in order to ensure that all the tables in the MIRE MIS are accessed correctly. Queries result in a standardized list report that lists all records matching the search criteria. These may be fed into a reporting system that can generate output that is more advanced.

The search of the digital image archive works in a similar way with the addition of the capability to specify a particular document type of interest (photo, scanned report, all images, etc.).

In addition to this powerful query function designed for internal users of the MIRE MIS, it is also possible to support external users (and less-technical internal users) through a constrained query tool made available online. This tool is designed to allow users to build a query by selecting from among allowed data fields, establishing sort and filtering options, date ranges, and other criteria.

4.12 Support the designated number of internal and external users.

MIRE MIS stewards may establish limits on access to the system. Each user must establish an account that is secured with a login identifier and password. The system's administrators can establish user access levels by class (Administrator, database manager, general user, data entry, etc.) and for each individual (i.e., users may be granted selective access to some portions of a higher level class without being given the full permissions associated with that pre-defined class. Permissions may be established at the record-level and for each database field independently to allow individual users (and classes of users) permission to view, edit, and/or delete information as separate permissions established in the user profile.

4.13 Provide an on-line, point-and-click, mapping location tool for data entry.

As a proven way to increase location data accuracy, users responsible for entering records will have access to a "smart map" which gives them a point-and-click interface for entering location data. Users can call up the map, zoom, pan in any direction, and click on the precise location of interest. The map will then supply all required location information and auto-populate the relevant fields on the data entry form. This may include latitude/longitude coordinates, roadway name(s), route number(s), segment identification number, milepoint/milepost, and any other information that is available from within the department's or State's GIS. If a statewide GIS does not exist, it is possible to implement the smart map capability using generic mapping tools (e.g., Google Earth).

4.14 Support a multi-linear referencing system.

MIRE MIS can incorporate translation tables to support multiple location coding methods including latitude/longitude, milepoint, milepost, and others. Ideally, the information is available within a GIS and the act of clicking on the smart map generates the relevant location information for entry into MIRE MIS. Where this is not feasible, MIRE MIS will still have the capability to store multiple location codes and the State can use the information entered into MIRE MIS to develop the desired translation tables.

Database

5.1 Provide for MIRE MIS retention and archiving as determined by State.

Each State may establish its own record retention and archival process for MIRE MIS. The system can generate an archive file organized by year, which the State can then store separately. Archived records may be accessed through the MIRE MIS by the reverse process of "reattaching" the archive to the system at which the archived records become another database source file that can be searched and used in analyses.

5.2 Provide a scalable and secure relational database with backup and recovery procedures.

The MIRE MIS is designed to work well regardless of the number of records. The database is secure with administrator control over user access and permissions. Backup and recovery are supported through automated and administrator-controlled processes.

5.3 Provide web enabled analysis tool for authorized user groups.

Authorized users may access a set of web-based analysis tools include search and report generating functions. The search tool will be designed to enable users quickly specify criteria such as date ranges and record selection/filtering criteria based on any field in the database to which they are allowed access. The search feature results in a list of matching records that is then available for use in analyses. From the user's perspective, the selection criteria and report specification are just two steps in the process of generating a desired output from the system. Multiple user-generate output types are supported including a simple list report, one-way and multi-way cross tabulations, graphic displays, and data extracts. The State may specifically limit users as to which types of report they may obtain.

5.4 Support export of data for HPMS.

5.5 Support export of data for *SafetyAnalyst* and other tools as approved.

With any known standard for data submissions, MIRE MIS can be programmed to generate a compliant data extract. The system will be developed with the capability to provide data extracts matching the HPMS and *SafetyAnalyst* data input requirements. System Administrators will have access to a tool that will allow them to update these output definitions and create new ones as the need arises.

5.6 Allow retrieval of roadway and crash data for proposed safety projects.

The MIRE MIS search and reporting functions support the intended use of the data in support of safety projects by providing authorized users with a way to select relevant records and generate output reports. The search feature's output of a simple list of matching records also supports users' access to individual records including inventory, crash, traffic, and any other data available through MIRE MIS. One possibility is that the data could be accessible through a GIS providing users with a map-based interface that would allow them to click on a location, call up crash reports for that location along with the related inventory and traffic data. This functionality is supported by MIRE MIS as it can provide data and records to an external GIS. MIRE MIS's web-based spatial analysis tool supports this capability for authorized users.

5.7 Provide analysis tools to authorized users in the form of queries and ad hoc reports from MIRE MIS.

The analysis functions within MIRE MIS are not constrained—every field in the database can be used in a query or report. User permissions established by the Administrator control that fields are available to a particular user. The internal query and reporting capabilities are akin to those described for a general user audience via the web. In fact, the same tools are available to both users, but the one intended for general user access is constrained in that selected fields are not available for query or analysis. Internal users can access any field as long as they have the appropriate permissions as established by the System Administrator.

5.8 Provide a web enabled spatial analysis tool for authorized user groups.

Authorized users will have access to a spatial analysis tool that can generate cross tabulation reports and support map-based output of results.

5.9 Provide ad hoc reporting ability for spatial analysis tool.

The ad hoc reporting capabilities of the spatial analysis tool include one-way and multi-way cross tabulations, and reporting of frequency counts and proportions (percentages).

5.10 Provide graphical depiction of analyses based on, for example, road sections, bridges, and intersections for specific timeframe and weather conditions.

Users of the web-based spatial analysis tool will have the option of producing graphical output in addition to the tabular reports. The graphical output may include support for condition diagrams, intersection/site diagrams of crashes, frequency histograms, pie charts, line graphs, and others.

6.1 Provide reports, queries, and/or inquires defined by the State, such as proposed safety project information.

The analytic tools designed for MIRE MIS support user-generate ad hoc query and analysis functions. Any query or analysis defined in the system may be saved for future use and shared with other users.

6.2 Allow authorized users to view and generate pre-defined reports, queries, and/or inquires via a browser.

State-defined standard reports are established in the system in the same way as ad hoc reports. They are simply saved and shared with the user community in general. The web-based analysis tool will allow users to access any pre-defined query or report accessible to them based on their permissions established by the System Administrator. Users will be blocked from generating output from a standard/pre-defined report for which they do not have the required user permissions as established by the System Administrator.

6.3 Provide an interface for submitting requests for additional standard reports via browser.

Users may request database searches and reports (i.e., analytic assistance) by completing a form available on line. The MIRE MIS system will log and track all of these user requests. It will also support workflow routing of requests to the appropriate agency staff.

Government Reports

6.4 Generate mandated and standard reports for federal agencies.
MIRE MIS will be programmed with a series of standard reports meeting the requirements established under federal programs. These report definitions will be accessible to the relevant agency staff so that they can be modified in the future as the reporting requirements change.

6.5 Provide aggregate information to other agencies and organizations via various media or through direct file transmission.

MIRE MIS will support electronic sharing of reports through creation of .pdf output. The system will also support data transmission by allowing users to define data extracts, which can then be shared electronically. The System Administrator in the user profile will establish user permissions for the creation of data extracts.

7.1 Embedded help file & help file index.

MIRE MIS documentation will be used to generate a help file that will be available to users from within the system interface (by clicking on the "Help" button. The embedded help will document all user-accessible system features and the functioning of each option in the user interface.

7.2 Reporting and standards instruction manual.

MIRE MIS data definitions and relevant data validation checks will be documented in an instruction manual for data collectors. The same manual can be shared with analysts using the system so that they can be made aware of the conditions for data acceptance into the production database.

7.3 Context-sensitive help on data fields.

For data fields on any data entry form, users may access help in the form of information derived from the data collector's instruction manual. This will include the data definition for that field, instructions for completing data entry into the field, and relevant validation rules for that field.

7.4 Step-by-step system function assistance (Show me how).

For user-accessible functions defined in the system, the help file will contain step-by-step instructions on how to access and complete a desired task using that function. This type of user assistance is akin to a "show-me-how" set of steps that take the user through every step in the process.

7.5 Function completion wizards (Do it for me).

For selected user-accessible functions, a wizard-based utility will provide users with expert-level automatic task completion. This is akin to a "do-it-for-me" command that will launch a default sequence of instructions and form completions by the system. Since it is impossible to anticipate every possible user need, this type of wizard-based approach effectively supports only the most standardized (routine) tasks; however, the wizards also serve an important training function within the system as they provide users with a complete and correct example upon which to base their own future uses of the system.

7.6 Interactive tutorials.

The MIRE MIS also includes a set of interactive tutorials. These are delivered as an animated, annotated slide show with multiple branching opportunities in which the user controls the forward progression based on choices made at each step in the process. This type of tutorial is a powerful user aid since it provides full step-by-step examples of task completion using the system. The ability to annotate key features of the presentation make this a good learning tool as well since users may be presented with options and explanations while they are working through the programmed examples.