# U.S. DOT Roadway Transportation Data Business Plan (Phase 1)

## Data Business Plan

www.its.dot.gov/index.htm
**Final Report — January 2013**
**FHWA-JPO-13-084**

U.S. Department of Transportation

Produced by Contract No. DTFH61-06-D-00004, Task CA04-075M2, U.S. DOT
U.S. DOT Roadway Transportation Data Business Plan (Phase 1)
Data Business Plan
U.S. Department of Transportation
FHWA Office of Operations

## Notice



## 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 for continuous quality improvement.

**Technical Report Documentation Page**

| 1. Report No. | 2. Government Accession No. | 3. Recipient's Catalog No. |
|---|---|---|
| FHWA-JPO-13-084 | | |

| 4. Title and Subtitle | 5. Report Date |
|---|---|
| U.S. DOT Roadway Transportation Data Business Plan (Phase 1)<br>Data Business Plan | January 2013 |
| | 6. Performing Organization Code |

| 7. Author(s) | 8. Performing Organization Report No. |
|---|---|
| Anita Vandervalk, Dena Snyder, and Judith K. Hajek | |

| 9. Performing Organization Name And Address | 10. Work Unit No. (TRAIS) |
|---|---|
| Cambridge Systematics, Inc.<br>1566 Village Square Boulevard, Suite 2<br>Tallahassee, FL 32309 | 11. Contract or Grant No.<br>DTFH61-06-D-00004 |

| 12. Sponsoring Agency Name and Address | 13. Type of Report and Period Covered |
|---|---|
| U.S. Department of Transportation<br>ITS Joint Program Office-HOIT<br>1200 New Jersey Avenue, SE<br>Washington, DC 20590 | Final Report<br>2011-2013 |
| | 14. Sponsoring Agency Code<br>HOTM |

**15. Supplementary Notes**

FHWA COTM: Walter During

**16. Abstract**

In 2010 the FHWA Office of Operations, Office of Transportation Management (HOTM) commissioned the development of a white paper, Data Capture and Management: Needs and Gaps in the Operation and Coordination of U.S. DOT Data Capture and Management Programs. The white paper examined current data capture and management activities across various U.S. DOT program areas, identified gaps and potential opportunities for filling the gaps to effectively and efficiently coordinate and manage the programs' activities. The primary recommendation from the white paper was that the HOTM develop a Data Business Plan to address the gaps identified in the paper. The Data Business Plan would serve to improve coordination among real-time data capture programs within U.S. DOT by clearly defining U.S. DOT needs for real-time data, address overlaps and gaps in program needs with respect to stakeholders and ultimately result in cost savings for U.S. DOT.

In 2011, the FHWA Office of Operations, Office of Transportation Management (HOTM) commissioned the Data Business Plan study. The purpose of the Data Business Plan is to:

- Address gaps identified in the Data Gaps paper (e.g., improve the coordination and communication mechanism across U.S. DOT offices to avoid investing resources on the same or similar types of data related programs);
- Serve as prototype for other U.S. DOT offices; and
- Provide leadership by suggesting or offering best data collection/management practices to primary data collectors.

Applying the concepts in this Data Business Plan will ensure the stakeholders are in line with FHWA priorities and immediately allow them to benefit from the improvement of coordination of data programs. These benefits include improved availability of data to support planning, operations and performance measures activities; possible reduction in expenditures due to the elimination of duplicate data; more rapid, targeted data acquisitions; broader sharing of data resources; systematic coordination and clarification of data-related federal policy and reduced data collection and management costs. These benefits are particularly relevant as it relates to the connected vehicle initiative. In summary, application of Data Business planning concepts will ensure the ability to better serve the needs of the U.S. DOT and its customers.

| 17. Key Words | 18. Distribution Statement |
|---|---|
| Data business plan, stakeholder coordination, outreach, data capture and management, data governance, connected vehicle, roadway mobility data | No restrictions. This document is available to the public through the National Technical information Service, Springfield, VA 22161 |

| 19. Security Classif. (of this report) | 20. Security Classif. (of this page) | 21. No. of Pages | 22. Price |
|---|---|---|---|
| Unclassified | Unclassified | **231** | N/A |

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

# Table of Contents


**Executive Summary ..... 1**

Background ..... 1
Purpose ..... 2
Summary of Deliverables ..... 2
Data Coordination Framework ..... 3
Conclusion ..... 4
Next Steps ..... 5

**1.0 Introduction 6**

1.1 Background ..... 6
1.2 Purpose ..... 7
1.3 Benefits of Data Business Plan ..... 8
1.4 Summary of Deliverables ..... 8
1.5 Scope of the Data Business Plan ..... 10
1.6 Plan Organization ..... 11

**2.0 Stakeholder Needs and Gap Analysis ..... 13**

2.1 Data Business Plan Stakeholders ..... 13
Internal Stakeholders ..... 13
External Stakeholders ..... 14
2.2 Stakeholder Needs for Data Programs ..... 18
2.3 Stakeholder Needs for Data Business Planning ..... 22
2.4 Gap Analysis ..... 26

**3.0 Data Coordination Framework ..... 30**

3.1 Mobility Data Coordination Group ..... 30
3.2 Data Coordination Model Diagram ..... 31
3.3 Roles and Responsibilities ..... 33
3.4 Rules of Engagement ..... 36
3.5 Outreach Mechanisms ..... 36

**4.0 Data Management Practices ..... 38**

4.1 Data Governance ..... 38
4.2 Data Acquisition ..... 39
4.3 Data Quality ..... 39
4.4 Data Standards ..... 39
4.5 Business Analysis Tools ..... 40
4.6 Data Privacy and Security ..... 40
4.7 Data Storage and Access ..... 40
4.8 Traceability ..... 41
4.9 Performance Measures ..... 41
4.10 Risk Assessment ..... 41
4.11 Value of Data Programs ..... 42
4.12 Knowledge Management ..... 42

**5.0 Conclusion and Next Steps ..... 43**

5.1 Conclusion ..... 43
5.2 Next Steps ..... 44

**APPENDIX A. Glossary of Terms ..... 45**
**APPENDIX B. Best Practices in Data Management ..... 48**
B.1 National Best Practices in Data Management ..... 48
U.S. Department of Education – National Center for Education Statistics ..... 48
Oak Ridge National Laboratory: Best Practices for Preparing Environmental Data Sets to Share and Archive ..... 49
National Oceanographic Data Center (NODC) ..... 50
Applicability to U.S. DOT ..... 51
Data Management Checklist ..... 51
Real-Time Data Capture and Management State of the Practice Assessment and Innovations Scan ..... 52
Policy Analysis and Recommendations for the Data Capture and Management Program: Implementation of Open Data Policies and System Policies for the Research Data Exchange and Data Environments ..... 57
Industry Options: State-of-the-Practice Policies and Lessons Learned on Open Data and Open Source ..... 58
B.2 State Best Practices in Data Management ..... 59
Alaska Department of Transportation and Public Facilities (ADOT&PF) ..... 59
Minnesota Department of Transportation (MnDOT) ..... 62
B.3 Internal Best Practices ..... 66
Data Quality ..... 66
**APPENDIX C. Best Practices in Data Governance ..... 69**
C.1 Overview of Data Governance ..... 69
Data Governance Benefits ..... 71
Data Governance Challenges ..... 72
C.2 Best Practices in Data Governance ..... 73
National Best Practices ..... 73
State Best Practices ..... 75
C.3 State of Data Governance at U.S. DOT ..... 79
FHWA Data Governance Advisory Council ..... 79
FHWA Visualization Working Group ..... 80
Connected Vehicle Senior Policy Working Group ..... 80
RDE/OSADP Policy Group ..... 82
**APPENDIX D. Example Data Governance Framework ..... 83**
D.1 Map Data Programs to Business Objectives ..... 83
D.2 Define Stakeholder Roles and Responsibilities ..... 85
D.3 Develop Data Governance Model ..... 86
D.4 Develop Data Governance Charter ..... 88
Example Data Governance Charter ..... 88
D.5 Develop Data Catalog ..... 90
D.6 Assess Data Governance Maturity ..... 99

**APPENDIX E. FHWA Data Governance Advisory Council Charter .. 102**
Article I. Purpose, Authority, and Duration ..... 103
Article II. Scope and Responsibilities ..... 103
Article III. Membership ..... 104
Article IV. Schedule ..... 104
Article V. Effective Date and Review ..... 105
**APPENDIX F. Data Mapping ..... 106**
**APPENDIX G. U.S. DOT Mobility Data Coordination Group – Memorandum of Understanding ..... 119**
G.1 PARTICIPANTS ..... 119
G.2 RULES OF ENGAGEMENT ..... 120
G.3 OPERATIONAL GUIDELINES ..... 122
**APPENDIX H. Data Inventory ..... 124**
H.1 INTRODUCTION ..... 124
Background ..... 124
Purpose ..... 125
H.2 DATA INVENTORY ..... 125
Methodology ..... 125
High Level Data Inventory ..... 126
Data Source Hierarchies ..... 143
Detailed Inventory – Infrastructure Data ..... 151
Detailed Inventory – Travel Data ..... 157
Detailed Inventory – Climate Data ..... 171
Detailed Inventory – Modal Data ..... 174
Detailed Inventory – Travel Behavior Data ..... 201
Detailed Inventory – Other Data Programs ..... 207
Detailed Inventory – Real-Time Data Capture and Management Program ..... 215
H.3 RECOMMENDATIONS AND NEXT STEPS ..... 221

## List of Tables


Table 2-1. Stakeholder List ..... 14
Table 2-2. External Stakeholder Matrix ..... 16
Table 2-3. Stakeholder Needs for Data Programs ..... 18
Table 2-4. Stakeholder Needs for Data Business Planning ..... 22
Table 3-1. Stakeholder Roles and Responsibilities ..... 34
Table D-1. Example Stakeholder Roles and Responsibilities ..... 85
Table D-2. Example Data Catalog ..... 91
Table D-3. Example Data Management Maturity Model ..... 100
Table G-1. Stakeholder Roles and Responsibilities ..... 121
Table H-1. Data Programs ..... 128
Table H-2. POC Probe Message Header Data Fields ..... 214


## List of Figures


Figure 1-1. Data Business Plan Deliverable Summary ..... 10
Figure B-1. Potential Applications of Current Industry Practices ..... 53
Figure B-2. ADOT&PF's Mission and Core Business Programs ..... 60
Figure B-3. ADOT&PF's Phased DBP Development ..... 61
Figure B-4. VDOT Data Business Plan Framework ..... 64
Figure D-1. Example Mapping of Mission, Business Processes, and Relationships to Core Data Systems ..... 84
Figure D 2. Example Data Governance Model Diagram ..... 87
Figure F-1. Highway Performance Monitoring System (HPMS) Data Mapping ..... 107
Figure F-2. National Speed Data Collection and Reporting Program (Future) Data Mapping ..... 108
Figure F-3. Travel Monitoring and Analysis System (TMAS) Data Mapping ..... 109
Figure F-4. Vehicle Travel Information System (VTRIS) Data Mapping ..... 110
Figure F-5. Clarus Data Mapping ..... 111
Figure F-6. North American Transborder Freight Data Mapping ..... 112
Figure F-7. Cross-Town Improvement Project (C-TIP) (Future) Data Mapping ..... 113
Figure F-8. Freight Analysis Framework Data Mapping ..... 114
Figure F-9. Freight Performance Measures Initiative Data Mapping ..... 115
Figure F-10. International Freight Data System (IFDS) (Future) Data Mapping ..... 116
Figure F-11. Data.gov Data Mapping ..... 117
Figure F-12. DCM Program Research Data Exchange Data Mapping ..... 118
Figure H.1 Basic Data Source Hierarchy ..... 144
Figure H.2 Data Hierarchy for Infrastructure Data ..... 145
Figure H.3 Data Hierarchy for Volume Data – Office of Highway Policy Information ..... 146

Figure H.4 Data Hierarchy for Speed Data – Office of Highway Policy Information .....147
Figure H.5 Data Hierarchy for Volume and Speed Data – Office of Operations .....148
Figure H.6 Data Hierarchy for Vehicle Classification and Weight Data ...149
Figure H.7 Data Hierarchy for Climate Data.....150
Figure H.8 FPMS II Database Structure.....188
Figure H.9 FAF Geographic Coverage.....192
Figure H.10 FAF Data Products .....193
Figure H-11 2009 NHTS Summary of Content .....205
Figure H.12 DUAP Evaluation Framework.....212
Figure H.13 Overview of the Research Data Exchange .....217

## List of Attributes

Figure 1-1. Cambridge Systematics, Inc., January 2013

Figure 3-1. Cambridge Systematics, Inc., January 2013

Figure B-1. Real-Time Data Capture and Management State of the Practice Assessment and Innovations Scan: Lessons from Scan of Current Practices, Science Applications International Corp., April 2011

Figure B-2. Alaska Data Business Plan: Data Governance, Standards, and Knowledge Management White Paper. Cambridge Systematics, Inc., September 2009

Figure B-3. Alaska Data Business Plan: Data Action Plan White Paper. Cambridge Systematics, Inc., September 2009

Figure B-4. Virginia Department of Transportation System Operations Directorate Data Business Plan, Virginia Department of Transportation, June 26, 2008

Figure C-1. Cambridge Systematics, Inc., January 2013

Figure C-2. NCHRP Report 666: Target Setting Methods and Data Management to Support Performance-Based Resource Allocation by Transportation Agencies, Transportation Research Board, 2010

Figure D-1. Cambridge Systematics, Inc., January 2013

Figure D-2. Cambridge Systematics, Inc., January 2013

Figure F-1. Cambridge Systematics, Inc., January 2013

Figure F-2. Cambridge Systematics, Inc., January 2013

Figure F-3. Cambridge Systematics, Inc., January 2013

Figure F-4. Cambridge Systematics, Inc., January 2013

Figure F-5. Cambridge Systematics, Inc., January 2013

Figure F-6. Cambridge Systematics, Inc., January 2013

Figure F-7. Cambridge Systematics, Inc., January 2013

Figure F-8. Cambridge Systematics, Inc., January 2013

Figure F-9. Cambridge Systematics, Inc., January 2013

Figure F-10. Cambridge Systematics, Inc., January 2013

Figure F-11. Cambridge Systematics, Inc., January 2013

Figure F-12. Cambridge Systematics, Inc., January 2013

Figure H-1. Cambridge Systematics, Inc., June 2012

Figure H-2. Cambridge Systematics, Inc., June 2012

Figure H-3. Cambridge Systematics, Inc., June 2012

Figure H-4. Cambridge Systematics, Inc., June 2012

Figure H-5. Cambridge Systematics, Inc., June 2012

Figure H-6. Cambridge Systematics, Inc., June 2012

Figure H-7. Cambridge Systematics, Inc., June 2012

Figure H-8. Liao, C. Freight Performance Measure Systems (FMS) System Evaluation and Data Analysis: Final Report. University of Minnesota Center for Transportation Studies Report No. CTS-08-01, January 2008, http://conservancy.umn.edu/bitstream/5767/1/CTS08_01.pdf

Figure H-9. Southworth, F., D. Davidson, H. Hwang, B. Peterson, and S. Chin. The Freight Analysis Framework, Version 3 (FAF3): A Description of the FAF3 Regional Database and How It Is Constructed. Prepared for Federal Highway Administration, June 16, 2011, pg. 4.

Figure H-10. Southworth, F., D. Davidson, H. Hwang, B. Peterson, and S. Chin. The Freight Analysis Framework, Version 3 (FAF3): A Description of the FAF3 Regional Database and How It Is Constructed. Prepared for Federal Highway Administration, June 16, 2011, pg. 2.

Figure H-11. FHWA, 2009 National Household Travel Survey User's Guide, Version 2, Updated October 2011. http://nhts.ornl.gov/2009/pub/UsersGuideV2.pdf

Figure H-12. University of Michigan Transportation Research institute, VII Data Use Analysis and Processing (DUAP), Final Project Report (Phase II), October 2010

Figure H-13. McGurrin, M., A. Corrigan, J. Larkin, and K. Wunderlich. Concept of Operations: Data Capture and Management Research Data Exchange (Draft Report). RITA ITS Joint Program Office, August 1, 2011

# Executive Summary

## Background

In 2010 the FHWA Office of Operations, Office of Transportation Management (HOTM) commissioned the development of a white paper, *Data Capture and Management: Needs and Gaps in the Operation and Coordination of U.S. DOT Data Capture and Management Programs.*[1] The white paper examined current data capture and management activities across various U.S. DOT program areas, identified gaps and potential opportunities for filling the gaps to effectively and efficiently coordinate and manage the programs' activities.

The primary recommendation from the white paper was that the FHWA Office of Operations, Office of Transportation Management (HOTM) develop a Data Business Plan to address the gaps identified in the paper. The Data Business Plan would serve to improve coordination among real-time data capture programs within U.S. DOT by clearly defining U.S. DOT needs for real-time data, address overlaps and gaps in program needs with respect to stakeholders and ultimately result in cost savings for U.S. DOT.

In 2011, the FHWA Office of Operations, Office of Transportation Management (HOTM) commissioned the Data Business Plan study. The purpose of the Data Business Plan is to:

- Address gaps identified in the Data Gaps paper (e.g., improve the coordination and communication mechanism across U.S. DOT offices to avoid investing resources to the same or similar types of data related programs);
- Serve as prototype for other U.S. DOT offices; and
- Provide leadership by suggesting or offering best data collection/management practices to primary data collectors.

The timing of this Data Business Plan is ideal given the recently signed charter for the FHWA Data Governance Council (see Appendix E). The charter states that the Council will perform the following functions: Develop a proposed plan for corporately managing FHWA data; review existing data collection efforts for need, consistency and efficiency annually; and review and approve all new data collection effort and coordinate FHWA data collection efforts with other modes within the Department. The formal adoption of this Council signifies a need for improved data management within FHWA. Applying the concepts in this Data Business Plan will ensure the stakeholders are in line with FHWA priorities and allow them to benefit from the improvement of coordination of data programs immediately. These benefits include improved availability of data to support planning, operations and performance measures activities; decrease in possible expenditures for duplicate data; more rapid, targeted data acquisitions; broader sharing of data resources; systematic coordination and clarification of data-related federal policy and reduced data collection and management costs. These benefits are particularly relevant as it relates to the Connected Vehicle Initiative. In summary, application of Data Business planning concepts will ensure the ability to better serve the needs of customers of FHWA.

[1] Data Capture and Management, Needs and Gaps in the Operation and Coordination of U.S. DOT Data Capture and Management Programs, Report No. FHWA-HOP-11-004, Cambridge Systematics, November 2010.

# Purpose

The following are the goals and objectives of the Data Business Plan:

- GOAL 1. Improve the coordination and communication mechanism across U.S. DOT and FHWA offices involved with **roadway travel mobility data** to avoid investing resources in the same or similar types of data related programs.
  - Objective 1.1. Identify gaps and redundancies in roadway travel mobility data programs.
  - Objective 1.2. Devise "rules of engagement" regarding collaboration of the data functions for roadway travel mobility data.

*Roadway travel mobility data* includes data from roadway travel modes, including vehicle, truck freight, bicycle/pedestrian, and transit, as well as roadway inventory data.

*Involvement* means the office is performing one or several of the following functions related to data: collection, analysis, reporting, dissemination, or providing guidance to other stakeholders related to those functions.

- GOAL 2. Improve the coordination of the data capture activities associated with ITS JPO-sponsored research in wirelessly connected vehicle technologies with roadway travel mobility data programs within U.S. DOT and FHWA.
  - Objective 2.1. Identify how current and planned data from the connected vehicle initiative can support existing roadway travel mobility data programs.
  - Objective 2.2. Identify how data from roadway travel mobility data programs within U.S. DOT and FHWA can support the connected vehicle initiative.
  - Objective 2.3. Identify existing/future data inventory and data structures/ policies/governance practices that could be applicable to the Research Data Exchange.

*Roadway travel mobility data programs* are formally recognized data programs within U.S. DOT that involve the collection, analysis or reporting of roadway travel mobility data.

# Summary of Deliverables

The following documents have been produced under Phase 1 of the Data Business Plan project:

- The *Stakeholder Needs Description* documents stakeholder needs related to specific data programs and business processes that support the goals and objectives of the Data Business Plan.
- The *Data Inventory* provides a high level and detailed inventory of relevant national, State, and regional data programs that support the goals and objectives of the Data Business Plan. It is included in Appendix H.
- The *Data Governance Framework* represents the third element of the Data Business Plan. It is included in Appendix D.

The DBP is designed to be a "living document" to be updated and refined during implementation. Implementation of the DBP will include outreach with stakeholders through meetings in the spring of 2013, survey instruments with the stakeholders for each community of interest for the roadway travel mobility data programs, and forums for gathering information from stakeholders regarding business needs for data and information for each of these programs.

If successful, the DBP will create an implementable coordination process that results in time and cost savings to the U.S. DOT through improved efficiency in business operations and work processes through use of data sharing technology, elimination of redundant data collection efforts, and consensus in the use of streamlined data sources across organizational business units.

# Data Coordination Framework

It is recommended that U.S. DOT formally designate a U.S. DOT Mobility Data Coordination Group to coordinate on data issues impacting specific types of roadway mobility data (e.g., infrastructure/ inventory, roadway travel mobility, climate data, modal data, and data capture activities associated with ITS JPO-sponsored research in wirelessly connected vehicle technologies), as well as cross-cutting data management issues (e.g., data quality, standards, data privacy and security, etc.) that impact all data types. The goals for the group would be guided by the goals and objectives articulated for the Data Business Plan.

Once the Mobility Data Coordination Group and the individual Data Working Groups have been formally established, the groups should develop and approve a Memorandum of Understanding (MOU) to set forth the purpose, goals, membership, roles and responsibilities, data management policies, meeting schedule, and "rules of engagement" regarding collaboration and coordination for the group. Rules of engagement could include the following:

- Share RFP's for current and upcoming initiatives.
- Share current initiatives, activities, and best practices related to specific types of roadway travel mobility data.
- Identify needs and opportunities to create links between existing data sets and connected vehicle data sets in the future.
- Identify needs and opportunities to integrate national data sets to support performance measurement and asset management purposes.
- Share current activities and best practices related to data strategies, policies, standards, metadata, architecture, procedures and metrics.
- Develop data standards and stewardship recommendations for consideration by the FHWA Data Governance Advisory Council. Data standards are very important to achieve harmonization across stakeholders and should be a prominent data management practice in the rules of engagement.
- Identify opportunities to coordinate resources to reduce data redundancy and implement cost sharing strategies for the collection, management, and maintenance of roadway travel mobility data. Redundancy could be addressed through data standardization and an annual review of data programs to identify where duplicate data collection and storage can/should be eliminated in and replaced with a single

source of data for specific data programs. This will help to ensure that data is collected once and used many times.

- Identify opportunities to reduce redundancy in the development and maintenance of duplicate data systems, promote efficiency in system maintenance, and promote open source initiatives.
- Explore methods to enhance access to information and data for the roadway travel mobility data programs. This includes developing web-portals easily accessible by internal and external stakeholders for each of these programs to obtain data and information as needed. This will facilitate sharing of data with internal/external stakeholders, thereby reducing costs associated with data collection.
- Understand and promote the value of data as a U.S. DOT-wide asset.

# Conclusion

The implementation process will be important in finalizing and activating the plan:

- GOAL 1. Improve the coordination and communication mechanism across U.S. DOT and FHWA offices involved with *roadway travel mobility data* to avoid investing resources in the same or similar types of data related programs.

  *This will be accomplished through implementation of the Data Coordination Plan and development of a web portal to facilitate coordination.*

  - Objective 1.1. Identify gaps and redundancies in roadway travel mobility data programs.

    *This was accomplished in Section 2.0.*

  - Objective 1.2. Devise "rules of engagement" regarding collaboration of the data functions for roadway travel mobility data.

    *This will be accomplished in implementation of the Data Coordination Plan.*

- GOAL 2. Improve the coordination of the data capture activities associated with ITS JPO-sponsored research in wirelessly connected vehicle technologies with roadway travel mobility data programs within U.S. DOT and FHWA.

  The Data Coordination Plan will be the vehicle for accomplishing this goal.

  - Objective 2.1. Identify how current and planned data from the connected vehicle initiative can support roadway travel mobility data programs.

    *This was identified in Section 4.2 of the Stakeholder Needs Description.*

  - Objective 2.2. Identify how data from roadway travel mobility data programs within U.S. *DOT and FHWA can support the connected vehicle initiative.*

    *This was identified in Section 4.2 of the Stakeholder Needs Description.*

- Objective 2.3. Identify existing/future data inventory and data structures/policies/ governance practices that could be applicable to the Research Data Exchange.

  *This was accomplished in the Data Inventory and review of national best practices in Data Governance (Appendix C).*

# Next Steps

The next step for the Data Business Plan project is to begin Phase 2, which includes the following tasks:

- Execution of Data Business Plan Coordination. This would include implementation of the Data Coordination Framework, including formalization of the U.S. DOT Mobility Data Coordination Group and implementation of the data management practices recommended in this document. The first step to implementation should be to establish the Travel Data Working Group and apply the Data Coordination Framework described in Section 3 to this smaller group in order to test the various framework components (e.g., develop a MOU, rules of engagement, outreach mechanisms, etc.). An example MOU is provided in Appendix G.
- Data Integration Test Pilots. The test pilots will demonstrate the benefits and value of the Data Business Plan by integrating connected vehicle data with national data sets; integrating proposed and existing national or project data sets; and eliminate gaps and overcome redundancies in national and project data sets associated with mobility data.
- Execution of data governance within the FHWA Office of Operations. Data governance should be executed within the Office of Operations based on the example Data Governance Framework in Appendix D. The Office of Operations should also meet with the Chairperson of the FHWA Data Governance Advisory Council to identify opportunities for coordination and/or collaboration with that group. Upon implementation, the Data Business Plan should transition to a U.S. DOT-owned document with a federal perspective and include a data governance structure, data governance charter, and data governance manual.

# 1.0 Introduction

## 1.1 Background

In 2010 the FHWA Office of Operations, Office of Transportation Management (HOTM) commissioned the development of a white paper, *Data Capture and Management: Needs and Gaps in the Operation and Coordination of U.S. DOT Data Capture and Management Programs.*[2] The white paper examined current data capture and management activities across various U.S. DOT program areas, identified gaps and potential opportunities for filling the gaps to effectively and efficiently coordinate and manage the programs' activities. The white paper focused on three primary types of data – infrastructure, travel, and climate – described as follows:

1. Infrastructure (Inventory) Data – Included roadway geometry, roadway inventory, intersection characteristics, and the state of system controls;
2. Travel (Speed and Volume) Data – Travel data included vehicle location, presence, and speed within the system, internal vehicle status such as fuel consumption rate, or externally measured data such as recorded external temperature; and
3. Climate (Weather) Data – Climate data included data on weather and pavement surface conditions collected from road weather information systems (RWIS).

Data sources were identified, stakeholder interviews conducted, gaps identified and recommendations made related to the "real time data programs." The paper suggested that the definition of real-time should be two-fold depending on the use of the data:

- Real-time data for real-time decision-making; and
- Analysis using archived data where timeliness is less important than having a large enough and representative sample to support analysis.

It was proposed that the timeframes be defined specifically based on data use and stakeholder needs. For example, regarding speed data, the need for real-time would depend on the actual mobility or speed application for which it is being used. If the speed were required for national reporting, the actual collection period could be 15 minutes, while the required time for the data to be archived and reported could be weekly or monthly.

The gaps identified in the white paper included the operations, maintenance and management of data programs, stakeholder and program cooperation/ coordination, and the role of U.S. DOT in data integration.

The primary recommendation from the white paper was that the FHWA Office of Operations, Office of Transportation Management (HOTM) develop a Data Business Plan to address the gaps identified in the paper. The Data Business Plan would serve to improve coordination among real-time data capture programs within U.S. DOT by clearly defining U.S. DOT needs for real-time data,

[2] Data Capture and Management, Needs and Gaps in the Operation and Coordination of U.S. DOT Data Capture and Management Programs, Report No. FHWA-HOP-11-004, Cambridge Systematics, November 2010.

address overlaps and gaps in program needs with respect to stakeholders and ultimately result in cost savings for U.S. DOT.

In 2011, the FHWA Office of Operations, Office of Transportation Management (HOTM) commissioned the Data Business Plan study. The purpose of the Data Business Plan is to:

- Address gaps identified in the Data Gaps paper (e.g., improve the coordination and communication mechanism across U.S. DOT offices to avoid investing resources in the same or similar types of data related programs);
- Serve as prototype for other FHWA offices; and
- Provide leadership by suggesting or offering best data collection/management practices to primary data collectors.

The main tasks were as follows:

- Establish Data Business Plan team goals and objectives;
- Determine scope statement and stakeholder needs description;
- Conduct detailed inventory;
- Construct Data Governance Framework; and
- Develop Data Business Plan.

In mid-2012, Phase 2 was added to the project to execute the Data Business Plan and perform data integration test pilots.

## 1.2 Purpose

The following are the goals and objectives of the Data Business Plan:

- GOAL 1. Improve the coordination and communication mechanism across U.S. DOT and FHWA offices involved with **roadway travel mobility data** to avoid investing resources in the same or similar types of data related programs.
  - Objective 1.1. Identify gaps and redundancies in roadway travel mobility data programs.
  - Objective 1.2. Devise "rules of engagement" regarding collaboration of the data functions for roadway travel mobility data.

*Roadway travel mobility data* includes data from roadway travel modes, including vehicle, truck freight, bicycle/pedestrian, and transit, as well as roadway inventory data.

*Involvement* means the office is performing one or several of the following functions related to data: collection, analysis, reporting, dissemination, or providing guidance to other stakeholders related to those functions.

- GOAL 2. Improve the coordination of the data capture activities associated with ITS JPO-sponsored research in wirelessly connected vehicle technologies with roadway travel mobility data programs within U.S. DOT and FHWA.
    - Objective 2.1. Identify how current and planned data from the connected vehicle initiative can support existing roadway travel mobility data programs.
    - Objective 2.2. Identify how data from roadway travel mobility data programs within U.S. DOT and FHWA can support the connected vehicle initiative.
    - Objective 2.3. Identify existing/future data inventory and data structures/ policies/governance practices that could be applicable to the Research Data Exchange.

*Roadway travel mobility data programs* are formally recognized data programs within U.S. DOT that involve the collection, analysis or reporting of roadway travel mobility data.

## 1.3 Benefits of Data Business Plan

The timing of this Data Business Plan is ideal given the recently signed charter for the FHWA Data Governance Council (see Appendix E). The charter states that the Council will perform the following functions: Develop a proposed plan for corporately managing FHWA data; review existing data collection efforts for need, consistency and efficiency annually; and review and approve all new data collection effort and coordinate FHWA data collection efforts with other modes within the Department. The formal adoption of this Council signifies a need for improved data management within FHWA. Applying the concepts in this Data Business Plan will ensure the stakeholders are in line with FHWA priorities and allow them to benefit from the improvement of coordination of data programs immediately. These benefits include improved availability of data to support planning, operations and performance measures activities; decrease in possible expenditures for duplicate data; more rapid, targeted data acquisitions; broader sharing of data resources; systematic coordination and clarification of data-related federal policy and reduced data collection and management costs. These benefits are particularly relevant as it relates to the Connected Vehicle Initiative. In summary, application of Data Business planning concepts will ensure the ability to better serve the needs of customers of FHWA.

## 1.4 Summary of Deliverables

The following documents have been produced under Phase 1 of the Data Business Plan project:

- The *Stakeholder Needs Description* documents stakeholder needs related to specific data programs and business processes that support the goals and objectives of the Data Business Plan. The report defines internal stakeholders for each of the Federal data-related programs represented in the Data Business Plan and external stakeholders who are the customers of these programs. Data includes infrastructure, travel, climate and connected vehicle.
- The *Data Inventory* provides a high level and detailed inventory of relevant national, State, and regional data programs that support the goals and objectives of the Data Business Plan. The data inventory examines all stakeholders and documents the data programs they are using, their objectives, and how connected vehicle data

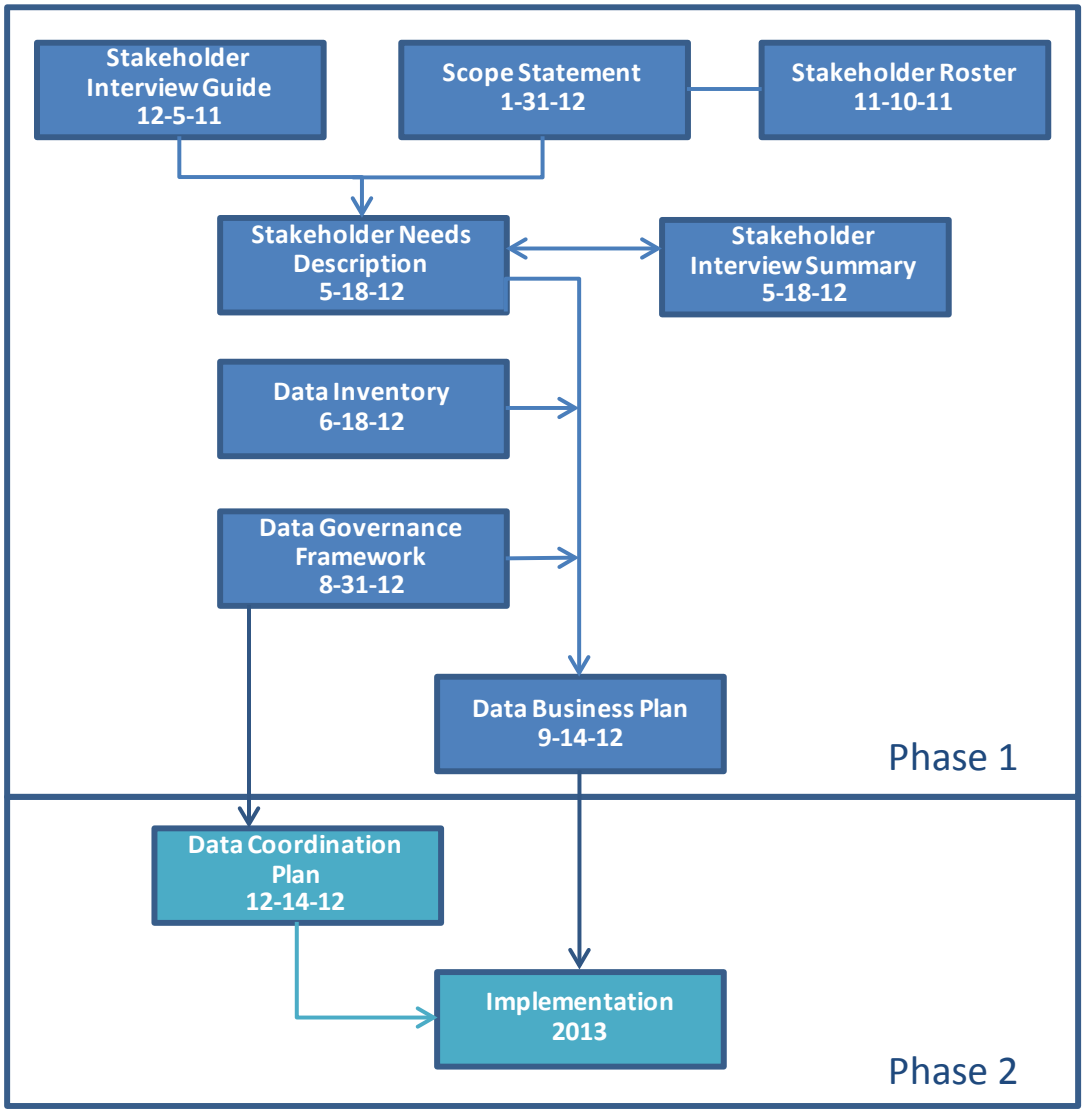
could benefit these programs. Data includes infrastructure, travel, climate and connected vehicle.

- The *Data Governance Framework* represents the third element of the Data Business Plan. It is included as Appendix D. Data includes roadway travel mobility (speed and volume) only. It includes the following components:
    - Mapping of Data Programs to Business Objectives;
    - Stakeholder Roles and Responsibilities;
    - Data Governance Model Diagram;
    - Draft Data Governance Charter;
    - Data Catalog; and
    - Assessment of Data Governance Maturity.

The second phase of the DBP project that began in October 2012 will develop the *Data Coordination Plan* based on the Data Business Plan document. Phase 2 will also include implementation of the DBP developed in Phase 1 (for the roadway travel mobility data programs.)

The following diagram shows the relationship between Phase 1 and Phase 2 of this project.

**Figure 1-1. Data Business Plan Deliverable Summary** (Source: Cambridge Systematics, Inc., January 2013)




## 1.5 Scope of the Data Business Plan

The Data Business Plan includes the following components:

- Recommendations to address U.S. DOT needs in serving data end-users and address overlaps and gaps in program needs with respect to stakeholders as identified in the *Needs and Gaps* white paper and the *Stakeholder Needs Description*.

- Best practices and recommendations for stakeholder coordination regarding data management practices, including data standards, metadata standards, data dictionaries, standard naming conventions, and possible taxonomies to be used.
- Recommendations for coordinating data-related efforts within U.S. DOT with the goal of reducing redundant or uncoordinated efforts and to systematically explore the potential of new forms of data to meet federal needs.
- Recommendations for improved coordination among data-related programs and offices within U.S. DOT. One of the fundamental challenges of developing the DBP will be to define a strategy that encourages data collectors to adapt their existing data collection procedures to better address the needs of the data users.

The DBP is designed to be a "living document" to be updated and refined during implementation. Implementation of the DBP will include outreach with stakeholders through meetings held in the spring 2013, survey instruments with the stakeholders for each community of interest for the roadway travel mobility data programs, and forums for gathering information from stakeholders regarding business needs for data and information for each of these programs.

If successful, the DBP will create an implementable coordination process that results in time and cost savings to the U.S. DOT through improved efficiency in business operations and work processes through use of data sharing technology, elimination of redundant data collection efforts, and consensus in the use of streamlined data sources across organizational business units.

## 1.6 Plan Organization

The Data Business Plan is organized into the following sections:

- **Section 2.0: Stakeholder Needs and Gap Analysis.** This section identifies internal and external stakeholders for the Data Business Plan, summarizes their needs related to roadway travel mobility data and connected vehicle data capture activities as they relate to the Data Business Plan, and identifies gaps and overlaps in program activities.
- **Section 3.0: Data Coordination Framework.** This section provides a comprehensive data coordination framework, including a data coordination group, model diagram, roles and responsibilities, rules of engagement, and outreach mechanisms.
- **Section 4.0: Data Management Practices.** This section recommends data management practices necessary to acquire, describe, store and protect data to ensure it can be used.
- **Section 5.0: Next Steps.** This section includes next steps to be conducted in Phase 2 of the Data Business Plan.
- **Appendix A: Glossary of Terms.** This section provides a glossary of terms related to data coordination and data management.

- **Appendix B: Best Practices in Data Management.** This section describes best practices in data management at the national and state levels. It also includes best practices internal to U.S. DOT.
- **Appendix C: Best Practices in Data Governance.** This section describes best practices in data governance at the national and state levels. It also includes best practices and current data governance initiatives within U.S. DOT.
- **Appendix D: Example Data Governance Framework.** This section describes steps for implementing data governance and includes an example framework for the FHWA Office of Operations.
- **Appendix E: FHWA Data Governance Advisory Council Charter.** This section provides a copy of the signed charter for the FHWA Data Governance Advisory Council.
- **Appendix F: Data Mapping.** This section contains graphics illustrating the relationship between data sets, core data systems, and internal/external stakeholders for the data.
- **Appendix G: Example Memorandum of Understanding.** This section provides an example Memorandum of Understanding that could be used by the U.S. DOT Mobility Data Coordination Group.
- **Appendix H. Data Inventory:** The Data Inventory provides a high level and detailed inventory of relevant national, state, and regional data programs that support the goals and objectives of the Data Business Plan.

# 2.0 Stakeholder Needs and Gap Analysis

This section identifies the internal and external stakeholders for the Data Business Plan and summarizes their needs related to roadway travel mobility data and connected vehicle data capture as documented in the *Stakeholder Needs Description*. It also identifies the gaps and overlaps in the data programs that can impact the timely delivery of quality data for decision-making purposes. One of the key goals of the Data Business Plan project was to deliver a DBP to address these data needs, gaps and overlaps.

This section is organized as follows:

2.1 – Documents internal and external stakeholders;

2.2 – Summarizes stakeholder needs specifically related to data programs;

2.3 – Summarizes stakeholder needs specifically related to data business planning; and

2.4 – Summarizes gaps and overlaps related to data program activities.

## 2.1 Data Business Plan Stakeholders

The Data Business Plan Stakeholders are comprised of internal and external stakeholders as identified in the *Stakeholder Needs Description.*

### Internal Stakeholders

Internal stakeholders are broadly defined as data collectors and users for each of the roadway travel mobility data programs represented in the Data Business Plan. Internal stakeholders include offices internal to U.S. DOT, including the ITS Joint Program Office, FHWA Office of Operations, FHWA Office of Human Environment, RITA Bureau of Transportation Statistics, FHWA Office of Highway Policy Information, Turner-Fairbank Highway Research Center, FHWA Research and Development, FHWA Office of Performance Management, Federal Transit Administration, and Federal Motor Carrier Safety Association. The internal stakeholder list is shown in Table 2-1. The internal stakeholder list was used to define the Mobility Data Working Groups and Internal Community of Interest (COI) for Roadway Travel Mobility Data (see Section 3.0).

**Table 2-1. Stakeholder List**

| Stakeholder Name | U.S. DOT Office |
|---|---|
| **Brian Cronin** | ITS Joint Program Office (JPO) |
| **James Pol** | ITS Joint Program Office (JPO) |
| **Dale Thompson** | ITS Joint Program Office (JPO) |
| **Steven Beningo** | RITA Bureau of Transportation Statistics (BTS) |
| **Gabe Rousseau** | FHWA Office of Human Environment – Bicycle and Pedestrian Program (HEPH) |
| **Tianjia Tang** | FHWA Office of Highway Policy Information (HPPI) |
| **Paul Pisano** | FHWA Office of Operations – Road Weather and Work Zone Management |
| **Jimmy Chu** | FHWA Office of Operations |
| **Rich Taylor** | FHWA Office of Operations |
| **Bob Rupert** | FHWA Office of Transportation Management (HOTM) |
| **Gene McHale** | Turner-Fairbank Highway Research Center (TFHRC) |
| **Nazy Sobhi** | Volpe (Federal Transit Administration (FTA ) representative) |
| **Ed Strocko** | FHWA Office of Freight Management and Operations (HOFM) |
| **Tom Kearney** | FHWA Office of Freight Management and Operations (HOFM) |
| **Mike Sprung** | FHWA Office of Freight Management and Operations (HOFM) |
| **Randy Butler** | FHWA Office of Freight Management and Operations (HOFM) |
| **Bill Bannister** | Federal Motor Carrier Safety Association (FMCSA) |
| **Joe Moyer** | FHWA Research and Development |
| **Pete Stephanos** | FHWA Office of Performance Management |
| **Karl Wunderlich** | Noblis (Data Capture and Management (DCM) Program Technical Support) |

## External Stakeholders

External stakeholders are customers of the various roadway travel mobility data programs and include the following agencies and organizations:

- AASHTO;
- Application developers;
- Corridor Coalitions;
- Institute of Transportation Engineers (ITE);
- Researchers;

- Public;
- Transit Agencies;
- Federal, state, and local enforcement personnel;
- Connected Vehicle Steering Group;
- Motor carrier industry;
- ITS America Public Transportation Forum;
- Freight operators;
- American Public Transportation Association (APTA);
- Insurance companies;
- Community Transit Association of America (CTAA);
- Transportation system managers;
- State and local transportation agencies;
- Traffic Management Centers (TMCs);
- Private data providers; and
- National Weather Service.

Table 2-2 indicates which external stakeholders are customers to each of the major internal stakeholder offices. This table was developed through internal stakeholder outreach in order to help facilitate coordination among U.S. DOT stakeholder offices with common customers. The External Stakeholder Registry was used to define the external Community of Interest for the Data Coordination Framework (see Section 3.0).

**Table 2-2. External Stakeholder Matrix**

| U.S. DOT OFFICES | EXTERNAL CUSTOMERS OF DATA PROGRAMS MANAGED BY U.S. DOT OFFICES | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | AASHTO | Corridor Coalitions | Researchers | Transit Agencies | Connected Vehicle Steering Group | ITS America Public Transportation Forum | American Public Transportation Association (APTA) | Community Transit Association of America (CTAA) | State and local transportation agencies | Private data providers | Application developers | Institute of Transportation Engineers (TE) | Public | Federal, state, and local enforcement personnel | Motor carrier industry | Freight Operators | Insurance companies | Transportation system managers | Traffic Management Centers (TMCs) | National Weather Service |
| **ITS Joint Program Office** | ✓ | ✓ | ✓ | | ✓ | | | | ✓ | ✓ | ✓ | ✓ | ✓ | | | | | | ✓ | |
| **FHWA Office of Operations** | | ✓ | ✓ | ✓ | ✓ | | | | ✓ | ✓ | ✓ | ✓ | ✓ | | | ✓ | | ✓ | ✓ | |
| **FHWA Office of Operations – Road Weather Management** | | | ✓ | | ✓ | | | | ✓ | ✓ | ✓ | | ✓ | | ✓ | ✓ | | ✓ | ✓ | ✓ |
| **RITA Bureau of Transportation Statistics** | | | ✓ | | | | | | ✓ | ✓ | ✓ | ✓ | ✓ | | | | | | | |
| **FHWA Office of Human and Natural Environment** | | | ✓ | | | | | ✓ | | ✓ | ✓ | | ✓ | | | | ✓ | ✓ | ✓ | |
| **Office of Highway Policy Information** | | | ✓ | | | | | | ✓ | ✓ | | | ✓ | | | | | | | |
| **Turner-Fairbanks Highway Research Center** | | | ✓ | | | | | | | ✓ | | ✓ | ✓ | | | | | | | |
| **Federal Transit Administration** | | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | | ✓ | | ✓ | | | | | | | |

| U.S. DOT OFFICES | EXTERNAL CUSTOMERS OF DATA PROGRAMS MANAGED BY U.S. DOT OFFICES | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | AASHTO | Corridor Coalitions | Researchers | Transit Agencies | Connected Vehicle Steering Group | ITS America Public Transportation Forum | American Public Transportation Association (APTA) | Community Transit Association of America (CTAA) | State and local transportation agencies | Private data providers | Application developers | Institute of Transportation Engineers (TE) | Public | Federal, state, and local enforcement personnel | Motor carrier industry | Freight Operators | Insurance companies | Transportation system managers | Traffic Management Centers (TMCs) | National Weather Service |
| **FHWA Office of Freight Management and Operations** | | | ✓ | | | | | | ✓ | ✓ | ✓ | ✓ | ✓ | | | ✓ | | | | |
| **Federal Motor Carrier Safety Administration** | | | ✓ | | | | | | ✓ | | | | ✓ | ✓ | ✓ | ✓ | ✓ | | | |
| **FHWA Office of Performance Management** | | | ✓ | | | | | | ✓ | | | | | | | | | | | |
| **Data Capture and Management Research Data Exchange** | | | ✓ | | | | | | ✓ | | ✓ | | | | | | | | | |

## 2.2 Stakeholder Needs for Data Programs

The *Needs and Gaps* white paper documented initial stakeholder needs related to the operations, maintenance and management of data programs, stakeholder and program cooperation/coordination, and the role of U.S. DOT in data integration. The *Stakeholder Needs Description* identified stakeholder needs related to roadway travel mobility data and connected vehicle data capture activities as they relate to the Data Business Plan. Table 2-3 summarizes all stakeholder needs from both sources and is organized by data type (e.g., infrastructure, travel, climate, modal, connected vehicle) and area of improvement (e.g., data scope, data acquisition and updating, data quality, data standards, etc.). The last column describes the Data Business Plan components that address the stakeholder needs, all of which will be described in later sections of this report.

**Table 2-3. Stakeholder Needs for Data Programs**

| Type of Data | Area of Improvement | Stakeholder Need | Source | Data Business Plan Component |
|---|---|---|---|---|
| Infrastructure Data | Data Scope (Gaps) | Improved information on location and operability of transportation assets (e.g., ITS equipment, HPMS monitoring sites) | Stakeholder Needs Description | Data Coordination Framework – Infrastructure/ Inventory Data Working Group |
| | Data Standards | Improved consistency in state roadway inventory attribute definitions, domain values, positional accuracy and resolution, data collection methods, and frequency of updates | Data Gaps White Paper | |
| | | Improved state compliance with Federal guidelines for collecting and maintaining roadway geometry and inventory data items | Data Gaps White Paper | |
| Travel Data | Data Scope (Gaps) | National urban travel time/speed dataset | Stakeholder Needs Description | Data Coordination Framework – Travel Data Working Group |
| | | Lane occupancy data | Stakeholder Needs Description | |
| | | Volume data to accompany probe speed data | Data Gaps White Paper | |
| | | Supporting data on sources of congestion (e.g., bad weather, crashes, etc.) | Stakeholder Needs Description | Data Coordination Framework – Travel Data COI (Internal) |

| Type of Data | Area of Improvement | Stakeholder Need | Source | Data Business Plan Component |
|---|---|---|---|---|
| Travel Data (continued) | Data Quality | GPS capability to adequately identify slowdowns and distinguish between frontage roads and interstate locations | Data Gaps White Paper | Data Coordination Framework – Travel Data Working Group |
| | | Uniform data quality (e.g., most state agencies and TMCs don't have maintenance funding to maintain high levels of availability and accuracy) | Stakeholder Needs Description | |
| | Data Acquisition and Updating | Agreement on a core methodology for reporting on congested corridors at the national level | Stakeholder Needs Description | |
| | Data Standards | Consistency in VMT calculation methodologies (e.g., passenger VMT recently shifted from a 2- to 4-axle schema, which impacted BTS because it does not fit previously established methodologies) | Stakeholder Meeting – November 2012 | Data Coordination Framework – Travel Data COI (Internal) |
| | Internal Coordination | More coordination between the Office of Operations and Office of Highway Policy Information in collecting speed data | Data Gaps White Paper | |
| Climate Data | Data Scope (Gaps) | Weather condition data at or on the roadway | Stakeholder Needs Description | Data Coordination Framework – Climate Data Working Group |
| | Business Analysis Tools | Decision support tools to help operators make better decisions based on weather condition data | Stakeholder Needs Description | |
| | Value of Data Programs | Continued funding for *Clarus* | Stakeholder Needs Description | |
| | | Connect research to real-world implementation | Stakeholder Needs Description | |
| Modal Data | Data Scope (Gaps) | Real-time transit data including monitoring passenger loads and transit vehicle health | Data Gaps White Paper | Data Coordination Framework – Modal Data Working Group |
| | | Video or still pictures from onboard cameras as a supplemental data source for transit | Data Gaps White Paper | |
| | | Access to roadway pricing data for FTA to analyze impact on traffic flow and mode choice | Stakeholder Meeting – November 2012 | |
| | | Multimodal and intermodal data, including bicycle/ pedestrian, airport data to support trip planning, parking data, and port/cruise terminal information | Data Gaps White Paper | |
| | | International hazardous materials transport data | Data Gaps White Paper | |

| Type of Data | Area of Improvement | Stakeholder Need | Source | Data Business Plan Component |
|---|---|---|---|---|
| Modal Data (continued) | Data Scope (Gaps) (continued) | Truck volume, intensity of movement, mode choice, and number of shipments along roadways and between city pairs (could be obtained from connected vehicle data) | Data Gaps White Paper | Data Coordination Framework – Modal Data Working Group (continued) |
| | | Real-time truck freight data to support the Commodity Flow Survey | Data Gaps White Paper | |
| | | Truck parking availability data (FMCSA is partnering with FHWA to collect data on truck parking availability, but state DOTs need to broadcast this data to trucks) | Stakeholder Meeting – November 2012 | |
| | Data Quality | Validated HPMS truck volumes using additional sources such as the Freight Analysis Framework | Stakeholder Needs Description | |
| | | Truck and bus VMT for FHWA to set targets (e.g., truck and bus crashes per VMT). There is a need to support the accurate collection of truck VMT at the state level, since some states are not collecting it well | Stakeholder Meeting – November 2012 | |
| | Data Standards | Bicycle/pedestrian data integrated with other data programs | Stakeholder Needs Description | |
| | | Ability to integrate various freight data sources (currently problematic due to segmentation differences) | Data Gaps White Paper | |
| | | More cross-boundary coordination/standardization in collection of freight load data at the state level | Stakeholder Needs Description | |
| | | Statutory authority to make the Truck Parking Program uniform and seamless across state boundaries | Stakeholder Needs Description | |
| | | Improved state compliance with Federal WIM classification guidelines | Stakeholder Needs Description | |
| | | Integrated data from different freight programs | Stakeholder Needs Description | |
| | Data Acquisition and Updating | Improved willingness of transit agencies to share specific vehicle location and schedule adherence data with third party sources | Data Gaps White Paper | |
| | | Methods to measure travel data for bicycle/pedestrians | Stakeholder Needs Description | |

| Type of Data | Area of Improvement | Stakeholder Need | Source | Data Business Plan Component |
|---|---|---|---|---|
| Modal Data (continued) | Internal Coordination | Truck volume data to support analysis for the BTS Commodity Flow Survey | Data Gaps White Paper | Data Coordination Framework – Modal Data Working Group (continued) |
| | Value of Data Programs | Connect research with real-world implementation (e.g., Smart Roadside Program is research oriented, while the Truck Parking Program was authorized by Congress in response to a specific problem) | Stakeholder Needs Description | |
| Connected Vehicle Data Capture | Data Acquisition and Updating | Data requirements for connected vehicle applications | Stakeholder Needs Description | Data Coordination Framework – Connected Vehicle Data Capture Working Group |
| | | Connected vehicle applications to incorporate data from national data programs (asset location data, loop detector data from TTID/Traffic.com, TMAS lane specific data, 23 CFR 511, *Clarus*) | Stakeholder Needs Description | Data Coordination Framework – Connected Vehicle Data Capture COI (Internal) |
| | | Improved data provided through the DCM Program (e.g., travel data, freight data gaps, Freight Analysis Framework, freight applications, *Clarus* program, National Household Travel Survey) | Stakeholder Needs Description | |
| | | Include TMCs as a major stakeholder/potential data source for the DCM Program | Stakeholder Needs Description | Data Coordination Framework – COI (External) |
| | | Consolidated efforts through working groups in the acquisition and maintenance process | Review comment | |
| | Internal Coordination | More coordination of applications developed under the DCM/DMA programs | Stakeholder Needs Description | Data Coordination Framework – Connected Vehicle Data Capture Working Group |
| | | Defined plan for communicating DCM/DMA program developments | Stakeholder Needs Description | |
| | | Coordination of freight applications being developed under the connected vehicle data capture program | Stakeholder Needs Description | |
| | | Better understanding of the synergies between the DCM Program and other FHWA data programs (i.e., how DCM could benefit from data being collected for other programs and vice versa) | Data Gaps White Paper | Data Coordination Framework – Connected Vehicle Data Capture COI (Internal) |
| | | Articulate the value of connected vehicle test datasets | Stakeholder Needs Description | |

Note: DCM refers to the FHWA Real-Time Data Capture and Management (DCM) Program under the FHWA Office of Operations, Office of Transportation Management.

The *Stakeholder Needs for Data Business Plan* document stated the following: The connected vehicle policy program is researching and analyzing critical policy and institutional issues associated with data capture and management and dynamic mobility application programs, including issues related to security, ownership, and liability of data. The DBP should coordinate with these efforts as documented in the following reports:

- *Identification of Critical Policy Issues for the Data Capture and Management (DCM) and Dynamic Mobility Application (DMA) Programs;*
- *State-of-the Practice and Lessons Learned on Implementing Open Data and Open Source Policies;*
- *Policy and Institutional Issues Analysis for the Dynamic Mobility Applications (DMA) Open Source Application Development Portal (OSADP);*
- *Policy Analysis and Recommendations for Development of the Dynamic Mobility Applications; and*
- *Policy Analysis and Recommendations for the Data Capture and Management Program: Implementation of Open Data Policies and System Policies for the Research Data Exchange and Data Environments.*

## 2.3 Stakeholder Needs for Data Business Planning

The *Data Gaps* white paper and *Stakeholder Needs Description* also documented stakeholder needs related to the Data Business Plan. Table 2-4 summarizes these stakeholder needs and is organized by area of improvement (e.g., data acquisition and updating, data quality, data standards, etc.).

**Table 2-4. Stakeholder Needs for Data Business Planning**

| Area of Improvement | Stakeholder Need | Source | Data Business Plan Component |
|---|---|---|---|
| Data Acquisition and Updating | Policy to define who is responsible for collecting data | Stakeholder Needs Description | Data Management Practices – Data Acquisition and Updating |
| | Improved state compliance with Federal data collection guidelines | Stakeholder Needs Description | |
| | Ability of users to obtain access to updated data as it becomes available within the original data source. | Data Gaps White Paper | |
| | Lower data costs (i.e., if data is acquired from private sector companies) | Stakeholder Needs Description | |
| | Consolidated efforts through working groups in the acquisition and maintenance process | Review comment | |
| Data Quality | Data quality standards for the collection and processing of data | Data Gaps White Paper, Stakeholder Needs Description | Data Management Practices – Data Quality |
| | Policy to determine who is responsible for data quality | Stakeholder Needs Description | |

| Area of Improvement | Stakeholder Need | Source | Data Business Plan Component |
|---|---|---|---|
| Data Quality (continued) | Data quality standards should address missing or incomplete data | Data Gaps White Paper | Data Management Practices – Data Quality (continued) |
| | Consistent data quality performance measures | Data Gaps White Paper | |
| | Standards for data accuracy for national data programs. Some states use default values for data elements, and the data is not independently validated. | Data Gaps White Paper | |
| | Data governance program to identify resources required for implementing data quality standards | Data Gaps White Paper | |
| | Non-redundant data (no duplication in multiple datasets). Data quality requirements should be met within the original data source. | Data Gaps White Paper | |
| | Understanding the implications of the Data Quality Act of 2001 on data programs | Stakeholder Needs Description | |
| Data Standards | More information on geolocation and operational status of transportation assets (e.g., ITS, HPMS monitoring sites) | Stakeholder Needs Description | Data Management Practices – Data Standards |
| | Common location referencing system at a national scale | Data Gaps White Paper, Stakeholder Needs Description | |
| | Better integration of data from multiple sources | Stakeholder Needs Description | |
| | Processes to reduce data duplication, improve integration/compatibility, and address overlap issues | Stakeholder Needs Description | |
| | Processes/systems to reduce redundancy and promote consistency in data collection | Stakeholder Needs Description | |
| | Data aggregation to support users at all levels | Data Gaps White Paper | |
| | Data aggregation to support analysis at both the national and regional/corridor levels | Stakeholder Needs Description | |
| | Data standard to indicate update frequency and age of data | Data Gaps White Paper | |
| | Metadata to clarify how the data can be used or linked | Data Gaps White Paper | |
| | Specific metadata guidelines rather than vague requirements to submit “acceptable metadata” | Stakeholder Needs Description | |
| | Documentation and metadata supplied by source agencies | Stakeholder Needs Description | |
| | Policy to address the need for metadata | Stakeholder Meeting – November 2012 | |

| Area of Improvement | Stakeholder Need | Source | Data Business Plan Component |
|---|---|---|---|
| Data Standards (continued) | Consistency with open data principles documented in the FHWA report, *Policy Analysis and Recommendations for Implementing Open Data for the Mobility Program* | Stakeholder Needs Description | Data Management Practices – Data Standards |
| | Consistent definition of "real-time" across different U.S. DOT Offices | Stakeholder Needs Description | |
| Business Analysis Tools | Comprehensive analysis tools to support analysis at both the national and regional/corridor levels | Stakeholder Needs Description | Data Management Practices – Business Analysis Tools |
| Data Privacy and Security | Data sensitivity, privacy, and security policies to ensure data is usable at the state and local level. Data standards to specify what data elements should not be released due to privacy concerns. | Data Gaps White Paper | Data Management Practices – Data Privacy and Security |
| | Consistency with recommendations related to security, ownership and liability of data as documented in policy documents by the connected vehicle policy program | Stakeholder Needs Description | |
| Data Storage and Access | Improved FHWA access to detailed data to support analysis at both the national and regional/corridor levels (e.g., ability to evaluate the impacts of significant events such as a bridge failure) | Stakeholder Needs Description | Data Management Practices – Data Storage and Access |
| | Policy to define who is responsible for hosting data | Stakeholder Needs Description | |
| | Policy to address licensing and data ownership issues to determine whether data can be shared with others | Stakeholder Needs Description | |
| | Compliance with NHTSA's data retention policy | Stakeholder Needs Description | |
| Internal Coordination | Mechanism to communicate (internally and externally) what data is available and how to access it | Stakeholder Needs Description | Data Coordination Plan – Outreach Mechanisms |
| | More internal coordination in acquiring private sector data | Stakeholder Needs Description | |
| | Mechanism for states to share information on alternative types of analysis that are being done with existing data | Stakeholder Needs Description | |
| | Forum for internal stakeholders to discuss needs, gaps and opportunities for collaboration | Stakeholder Needs Description | |
| | Mechanism for better coordination of activities among different U.S. DOT offices. U.S. DOT staff are not aware of what projects others are working on. | Stakeholder Needs Description, Data Gaps White Paper | |
| | Symposium or other mechanism for groups to meet on a regular basis (e.g., every 6 months) to discuss data topics | Stakeholder Needs Description | |

| Area of Improvement | Stakeholder Need | Source | Data Business Plan Component |
|---|---|---|---|
| Internal Coordination | Regular ongoing communication between internal stakeholders. This process should be institutionalized. | Stakeholder Needs Description | Data Coordination Plan – Outreach Mechanisms |
| | Communication mediums such as periodic newsletters or website to outreach and provide updates to internal/external stakeholders | Stakeholder Needs Description | |
| | Engagement of stakeholders to actively participate in working groups. Stakeholders have other demand for their time and are often involved in multiple projects. | Data Gaps White Paper | |
| Traceability | Analysis of the use of real-time data for analytical purposes (in which real-time data is archived then analyzed) | Stakeholder Needs Description | Data Management Practices – Traceability |
| | Improved coordination across data programs and U.S. DOT offices to ensure everyone's business needs for data are being met | Stakeholder Needs Description | |
| | Guidelines for how programs should organize themselves to take advantage of the benefits derived from the connected vehicle data capture program | Stakeholder Needs Description | |
| | Improved coordination across data programs and U.S. DOT offices to communicate activities that impact shared external stakeholders | Data Gaps White Paper | |
| | Improved traceability, especially tracing back to original data sources and making sure U.S. DOT has permission to share/distribute the data | Stakeholder Meeting – November 2012 | |
| | Policy to deal with the intersection of traceability/data privacy | Stakeholder Meeting – November 2012 | |
| Performance Measures | Improved consistency across U.S. DOT in definitions of performance measures (e.g., travel time reliability) | Data Gaps White Paper | Data Management Practices – Performance Measures |
| Data Governance | Collaboration on connected vehicle data, including conducting assessments and providing guidance on how individual programs could be using the data. Need to develop an action plan and charter to define the goals and objectives for this group. | Stakeholder Needs Description | Data Management Practices – Data Governance |
| | Designated data owners in place | Stakeholder Needs Description | |
| | High-level champion for data governance within the agency | Stakeholder Needs Description | |
| | Improved coordination between U.S. DOT offices related to data collection and reporting functions | Data Gaps White Paper | |

| Area of Improvement | Stakeholder Need | Source | Data Business Plan Component |
|---|---|---|---|
| Data Catalog | Data catalog to clearly identify what data is being generated, owners of that data | Stakeholder Needs Description | Data Catalog |
| Value of Data Programs | Quantified benefits of data governance to U.S. DOT in terms of cost savings and leadership | Stakeholder Needs Description | Data Management Practices – Value of Data Programs |
| | Educate stakeholders about the value and benefits of data coordination and governance activities | Stakeholder Needs Description | |
| | Business tools to quantify the value of data, which can have monetary, environmental, or societal value. Truck data can also be monetized (e.g., parking availability, roadway pricing, etc.). | Stakeholder Meeting – November 2012 | |
| Knowledge Management System | Knowledge management practices to reduce the loss of subject matter experts in the event of staff turnover | Data Gaps White Paper | Data Management Practices – Knowledge Management |
| | Future plans for databases (e.g., the continuation of *Clarus* and availability of future databases such as the National Speed Database) | Stakeholder Needs Description | |
| | Stakeholder education to overcome reluctance to work in groups and share information internally due to "job protection" | Data Gaps White Paper | |

This section presented the stakeholder needs for integrated data and for data business planning from the perspective of internal and external stakeholders. All of these issues support the continued implementation of a Data Business Plan and Data Coordination Framework for U.S. DOT.

## 2.4 Gap Analysis

Several gaps and overlaps were identified in the *Data Gaps* white paper and *Stakeholder Needs Description*. Additional research confirms that the following gaps and overlaps exist in program activities, and that there are opportunities for collaboration to address these issues. The gaps/overlaps are summarized as follows:

- There is currently a gap in asset information for HPMS traffic monitoring sites. Since HPMS data is now reported in GIS format (starting in 2011), it may be possible to obtain asset information for traffic monitoring sites.
- There is overlap between HPMS, the National Speed Data Collection and Reporting Program, TMAS, and the Urban Congestion Reporting Program as it relates to the use of speed and volume data.
- There is overlap in the data sets that support the Travel Monitoring and Analysis System (TMAS) and Vehicle Travel Information System (VTRIS).

- The Freight Analysis Framework (FAF) integrates data from a variety of sources such as the Commodity Flow Survey, the Freight Performance Measures (FPM) Program, Transborder Freight Data, and other data sources. The Office of Freight Management and Operations fuses the data to estimate the number of trucks, and then HPMS data is used with modeling to distribute volumes along the highway network. A major product is the freight network volume map. The Office of Highway Policy Information is trying to produce a similar map with vehicle volumes, and the two offices may benefit from improved coordination.
- The ITS Joint Program Office is conducting a project to geolocate ITS assets using Google StreetView. As part of the project, they are working to collect GIS base maps around the country. They currently have base maps for 30 states. There is potential overlap with a similar GIS location referencing project being conducted by the Office of Highway Policy Information.
- Currently, no cross referencing is being done between the Motor Carrier Management Information System (MCMIS) and the freight data being collected by ATRI through the Freight Performance Measures project. FMCSA uses data for day-to-day operations, while volume and speed data are more relevant to planning and analysis. FMCSA is more interested in the safety aspects of data.
- *Clarus* has some overlap with other data programs (e.g., MADIS, which is a NOAA program, and MesoNet, which is a regional collective for sharing weather data).
- Limited datasets such as the Transportation Technology and Innovation Demonstration (TTID) data, ITS data used for Urban Congestion Reporting, or truck speed data from the Freight Performance Measure program could be used as a surrogate for a national dataset.

Objectives 2.1 and 2.2 of the Data Business Plan are to identify how connected vehicle data could be useful to stakeholders. The following documents stakeholder needs and potential use of connected vehicle data.

- Connected vehicle applications could incorporate data from national data programs. Stakeholders suggested that the following programs could serve as a source of data for connected vehicle applications:
    - Asset location data could be tied into the connected vehicle effort in terms of knowing where equipment is deployed. Inventory data is critical in assessing data gaps and where there is coverage.
    - TTID/Traffic.com data could provide loop detector data. Traffic.com is the primary source of loop detector data installed on major interstate highways, and the program has consistently maintained high levels of data quality. Traffic.com data is made available for all FHWA research. Access to the data is provided through a data warehouse, which can be accessed by FHWA and the 27 state agencies. Each FHWA Program Office is provided with a user ID and password. The DCM Program provides an opportunity to identify data gaps and redundancy and to link to Traffic.com data. There is also a need to investigate whether the test data sets are using TTID/Traffic.com data.

- TMAS could provide lane specific information to support connected vehicles. Reporting on traffic monitoring sites is done through the Traffic Monitoring & Analysis System (TMAS) reports, not HPMS. TMAS could have a direct impact on V2V and V2I by providing lane specific information. This will be a need in the future. Even the freight performance measures initiative does not have lane specific monitoring.
- 23 CFR 511 could be used as a data source. As part of 23 CFR 511, there is a need to have the processes in place to ensure real-time information is being entered into some system so it could be used as a data source.
- *Clarus* needs to be linked to the DCM Program. *Clarus* is in its last year of funding. The Dynamic Mobility Applications program will rely on *Clarus* data, so it will need to continue to be funded to support connected vehicle applications. Fixed sensor data could be used as "ground truth data" to validate data generated by individual vehicles.

- National data programs could benefit from improved data provided through the DCM Program. Stakeholders suggested that the following programs could benefit from improved data:
  - Travel Data Programs: The DCM Program could be a source of travel data for existing data programs such as HPMS and the Urban Congestion Report program.
  - Freight Data Gaps: Coordination with the DCM program could fill the gaps in obtaining origins and destinations concerning moving tonnage between large urban areas. There has not been work done to coordinate with the connected vehicle initiative yet.
  - Freight Analysis Framework: There is a need to define the implications of connected vehicles on the Freight Analysis Framework. When vehicles start reporting themselves through connected vehicle applications, there will not be a need for ATRI data. The Freight Analysis Framework has not fully considered the implications of the connected vehicle data, and they do not have a plan in place for staying in touch with the developments of the program.
  - Freight Applications: Connected vehicle data could benefit freight applications. There are three freight applications that would benefit from connected vehicle data:
    - ATIS for Freight. Data needs for the ATIS for Freight (FRATIS) application will depend on what comes out of the ConOps. They are looking for information to support more efficient freight movements, including truck parking, weather information, and route guidance that incorporate bridge height/weight restrictions.
    - Dynamic Route Guidance. Dynamic Route Guidance gives drivers specific information on routes. If a route changes due to congestion, it provides an alternate route. The program focuses on gathering data for routes that are approved for freight movement.
    - Drayage Optimization. Drayage Optimization uses data available through the above programs, plus private sector data, and also connects to a carrier's work management system. This application allows work flows to be aligned with specific drivers and specific routes. Bringing in traveler information would allow operators to identify their next best move to improve efficiency. The application will be driven by private data sources.
  - Road Weather Management Program: The Road Weather Management Program would be interested in data they could get from the vehicle (e.g., temperature, pressure, windshield wiper activity to tell if it is raining at a location, etc.). They would like the ability to combine yaw with traction control, speed, braking, etc. to tell what the pavement condition is.

- National Household Travel Survey (NHTS): The DCM Program may benefit but not fully meet the needs of surveys such as the NHTS. It will depend on what connected vehicle data the Federal government will have access to. If it is not a Federal source, there is no control over data quality. However, probe data might not be sufficient to meet all data needs for these surveys (i.e., it would not provide information on movement of commodities, purpose of trip, demographic data, etc.).

The next section describes a detailed Data Coordination Framework for meeting the data and business process needs and for addressing the gaps/overlaps in data programs at U.S. DOT.

# 3.0 Data Coordination Framework

This section proposes a comprehensive Data Coordination Framework for U.S. DOT that includes the following components:

- Mobility Data Coordination Group
- Data Coordination Model Diagram
- Roles and Responsibilities
- Rules of Engagement
- Outreach Mechanisms

## 3.1 Mobility Data Coordination Group

It is recommended that U.S. DOT formally designate a Mobility Data Coordination Group to coordinate on data issues impacting specific types of roadway mobility data (e.g., infrastructure/inventory, roadway travel mobility, climate data, modal data, and data capture activities associated with ITS JPO-sponsored research in wirelessly connected vehicle technologies), as well as cross-cutting data management issues (e.g., data quality, standards, data privacy and security, etc.) that impact all data types. The goals for the group would be guided by the goals and objectives articulated for the Data Business Plan:

- GOAL 1. Improve the coordination and communication mechanism across U.S. DOT and FHWA offices involved with **roadway travel mobility data** to avoid investing resources in the same or similar types of data related programs.
  - Objective 1.1. Identify gaps and redundancies in roadway travel mobility data programs.
  - Objective 1.2. Devise "rules of engagement" regarding collaboration of the data functions for roadway travel mobility data.
- GOAL 2. Improve the coordination of the data capture activities associated with ITS JPO-sponsored research in wirelessly connected vehicle technologies with other related data programs within U.S. DOT and FHWA.
  - Objective 2.1. Identify how current and planned data from the connected vehicle initiative can support existing roadway travel mobility data programs.
  - Objective 2.2. Identify how data from roadway travel mobility data programs within U.S. DOT and FHWA can support the connected vehicle initiative.
  - Objective 2.3. Identify existing/future data inventory and data structures/ policies/governance practices that could be applicable to the Research Data Exchange.

## 3.2 Data Coordination Model Diagram

The model diagram shown in Figure 3-1 proposes a formal structure for the U.S. DOT Mobility Data Coordination Group. The following describes the components of the Data Coordination Model diagram:

**A.** U.S. DOT Mobility Data Coordination Group;
**B.** Roadway Travel Mobility Data Working Groups (Internal);
**C.** Community of Interest (Internal); and
**D.** Community of Interest (External).

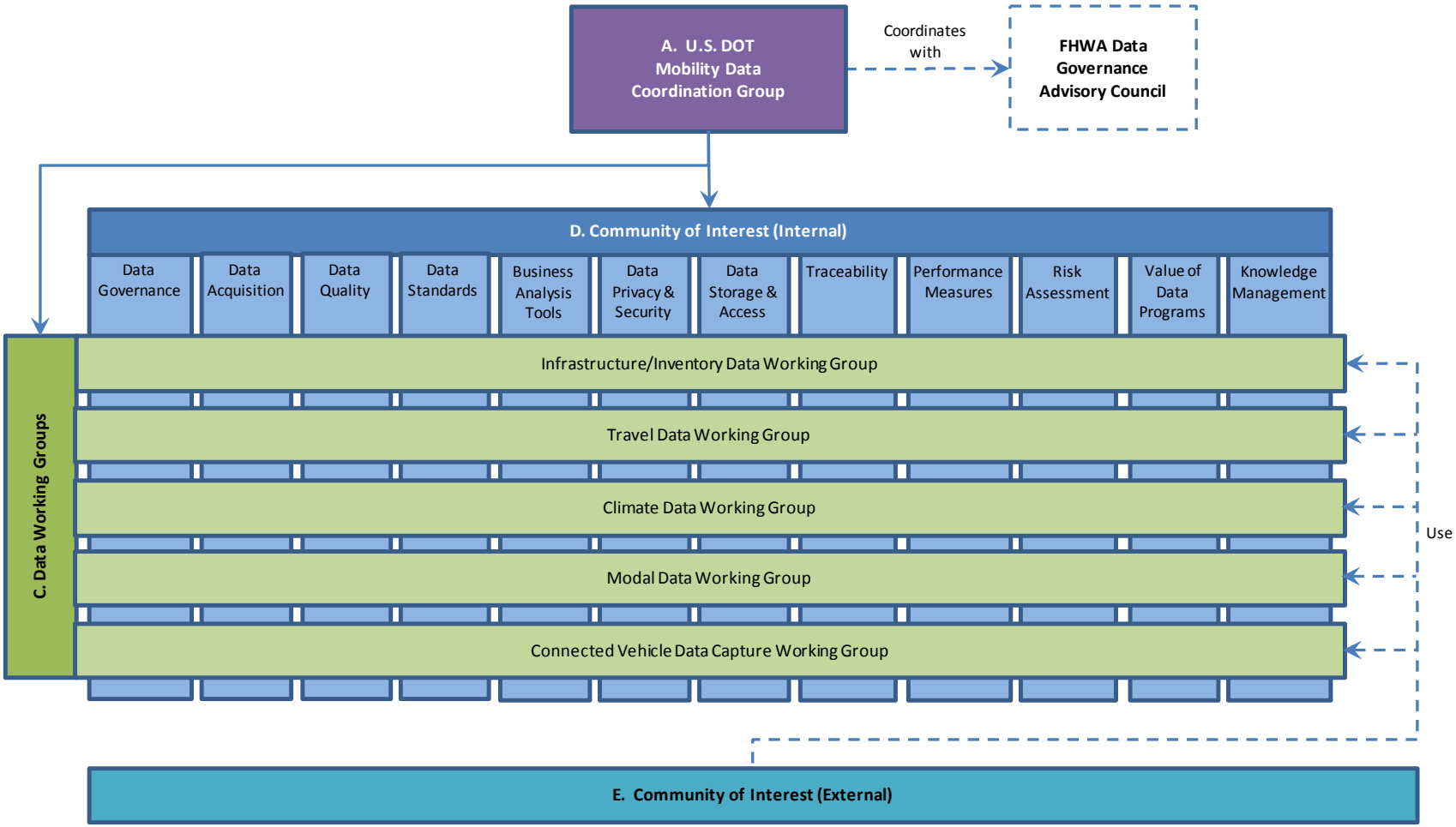
The U.S. DOT Mobility Data Coordination Group (A) would include designated individuals within U.S. DOT responsible for the oversight of roadway travel mobility data programs to support the business functions of their office. The Mobility Data Coordination Group would be chaired by a data coordination champion from within the FHWA Office of Operations; this individual would also liaison with the FHWA Data Governance Advisory Council.

The U.S. DOT Mobility Data Coordination Group would serve as an umbrella structure for smaller groups who would meet to coordinate on specific data issues as follows:

- Data Working Groups (B) would coordinate on data issues impacting a specific type of roadway mobility data (e.g., infrastructure/ inventory, travel data, climate data, modal data, and data capture activities associated with ITS JPO-sponsored research in wirelessly connected vehicle technologies).
- Internal Community of Interest (C) would coordinate on cross-cutting data management issues, such as data quality, standards, data privacy and security, etc., that impact all data types.

As users of roadway travel mobility data, the external Community of Interest (D) would not be involved in data coordination activities but would benefit from the improved coordination and data management practices resulting from the U.S. DOT Mobility Data Coordination Group.

**Figure 3-1. Proposed Data Coordination Model Diagram** (Cambridge Systematics, Inc., January 2013)



## 3.3 Roles and Responsibilities

The following data coordination roles are defined for U.S. DOT:

- U.S. DOT Mobility Data Coordination Group – The designated individuals within U.S. DOT responsible for the oversight of data programs to support the business functions of their offices.
- U.S. DOT Mobility Data Coordination Group Team Leader – A designated individual from within the FHWA Office of Operations who would chair the Mobility Data Coordination Group and liaison with the FHWA Data Governance Advisory Council.
- Data Working Groups – Any persons or offices internal to U.S. DOT that collect or manage roadway travel mobility data. This group dictates the policies, procedures, and business practices associated with roadway travel mobility data programs. Separate Data Working Groups would be defined for each data type (infrastructure/inventory, travel data, climate data, modal data, and connected vehicle data capture activities).
- Community of Interest (Internal) – Any persons or offices internal to U.S. DOT that collect, own, maintain, use or interface with, access, benefit from, or are otherwise affected by roadway travel mobility data. This group would coordinate and share best practices on data management issues such as data quality, standards, data privacy and security, etc.
- Community of Interest (External) – Any persons or offices external to U.S. DOT that use or interface with, access, benefit from, or are otherwise affected by roadway travel mobility data.

The recommended roles and responsibilities for supporting the Data Coordination Framework are listed in Table 3-1. The table corresponds to the components in the Proposed Data Coordination Model Diagram (Figure 3-1). These roles/responsibilities will be vetted with stakeholders during a Mobility Data Forum to be held in spring 2013.

**Table 3-1. Stakeholder Roles and Responsibilities**

| Role | Membership | Responsibility |
|---|---|---|
| U.S. DOT Mobility Data Coordination Group | • ITS Joint Program Office (JPO)<br>• RITA Bureau of Transportation Statistics (BTS)<br>• FHWA Office of Human Environment – Bicycle and Pedestrian Program (HEPH)<br>• FHWA Office of Highway Policy Information (HPPI)<br>• FHWA Office of Operations (HOTO)<br>• FHWA Office of Operations – Road Weather Management<br>• FHWA Office of Transportation Management (HOTM)<br>• Turner-Fairbank Highway Research Center (TFHRC)<br>• FHWA Office of Freight Management and Operations (HOFM)<br>• FHWA Office of Performance Management (TPM)<br>• Volpe (Federal Transit Administration (FTA ) representative)<br>• Federal Motor Carrier Safety Association (FMCSA)<br>• Noblis (DCM Program Technical Support) | Finalize Data Coordination Framework with input from Data Working Groups and Internal Community of interest<br>Develop and approve U.S. Mobility Data Coordination Group charter |
| Infrastructure/Inventory Data Working Group | • FHWA Office of Highway Policy Information (HPPI)<br>• RITA Bureau of Transportation Statistics (BTS)<br>• FHWA Office of Performance Management (TPM) | Address stakeholder needs related to infrastructure/inventory data<br>Identify and address gaps and redundancies in infrastructure data programs<br>Devise ‘‘rules of engagement’’ regarding collaboration and coordination<br>Develop data standards and stewardship recommendations for consideration by the U.S. DOT Mobility Data Coordination Group |
| Travel Data Working Group | • FHWA Office of Highway Policy Information (HPPI)<br>• RITA Bureau of Transportation Statistics (BTS)<br>• FHWA Office of Operations (HOTO) | Address stakeholder needs related to travel data<br>Identify and address gaps and redundancies in travel data programs<br>Devise ‘‘rules of engagement’’ regarding collaboration and coordination<br>Develop data standards and stewardship recommendations for consideration by the U.S. DOT Mobility Data Coordination Group |
| Climate (Weather) Data Working Group | • FHWA Office of Operations – Road Weather Management | Address stakeholder needs related to climate data<br>Identify and address gaps and redundancies in climate data programs<br>Devise ‘‘rules of engagement’’ regarding collaboration and coordination<br>Develop data standards and stewardship recommendations for consideration by the U.S. DOT Mobility Data Coordination Group |

| Role | Membership | Responsibility |
|---|---|---|
| Modal Data Working Group | • RITA Bureau of Transportation Statistics (BTS)<br>• Volpe (Federal Transit Administration (FTA ) representative)<br>• FHWA Office of Human Environment – Bicycle and Pedestrian Program (HEPH)<br>• FHWA Office of Freight Management and Operations (HOFM)<br>• Federal Motor Carrier Safety Association (FMCSA) | Address stakeholder needs related to modal data<br>Identify and address gaps and redundancies in modal data programs<br>Devise ‘‘rules of engagement’’ regarding collaboration and coordination<br>Develop data standards and stewardship recommendations for consideration by the U.S. DOT Mobility Data Coordination Group |
| Connected Vehicle Data Capture Working Group | • ITS Joint Program Office (JPO)<br>• Turner-Fairbank Highway Research Center (TFHRC)<br>• FHWA Research and Development<br>• Noblis (DCM Program Technical Support) | Address stakeholder needs related to connected vehicle data capture activities<br>Identify and address gaps and redundancies in connected vehicle data capture activities<br>Devise ‘‘rules of engagement’’ regarding collaboration and coordination<br>Develop data standards and stewardship recommendations for consideration by the U.S. DOT Mobility Data Coordination Group |
| Community of Interest – Internal | • ITS Joint Program Office (JPO)<br>• RITA Bureau of Transportation Statistics (BTS)<br>• FHWA Office of Human Environment – Bicycle and Pedestrian Program (HEPH)<br>• FHWA Office of Highway Policy Information (HPPI)<br>• FHWA Office of Operations (HOTO)<br>• FHWA Office of Operations – Road Weather Management<br>• FHWA Office of Transportation Management (HOTM)<br>• Turner-Fairbank Highway Research Center (TFHRC)<br>• FHWA Office of Freight Management and Operations (HOFM)<br>• FHWA Office of Performance Management (TPM)<br>• Volpe (Federal Transit Administration (FTA ) representative)<br>• Federal Motor Carrier Safety Association (FMCSA)<br>• Noblis (DCM Program Technical Support) | Coordinate with the Data Working Groups to:<br>• Address data gaps and overlaps<br>• Share current activities and best practices in data management<br>• Coordinate resources and cost sharing strategies to reduce redundancy in data collection, integration, and data systems<br>• Facilitate sharing of data with internal/external stakeholders<br>• Identify how current and planned data from the connected vehicle initiative can support existing roadway travel mobility data programs<br>• Identify how data from roadway travel mobility data programs within U.S. DOT and FHWA can support the connected vehicle initiative<br>• Identify existing/future data inventory and data structures/policies/ governance practices that could be applicable to the Research Data Exchange |
| Community of Interest – External | External stakeholders for infrastructure/inventory, travel, climate, modal, and connected vehicle data capture (listed in Table 2.2) | TBD |

## 3.4 Rules of Engagement

Once the Mobility Data Coordination Group and the individual Data Working Groups have been formally established, the groups should develop and approve a Memorandum of Understanding (MOU) to set forth the purpose, goals, membership, roles and responsibilities, data management policies, meeting schedule, and "rules of engagement" regarding collaboration and coordination for the group. Rules of engagement could include the following:

- Share RFP's for current and upcoming initiatives.
- Share current initiatives, activities, and best practices related to specific types of roadway travel mobility data.
- Identify needs and opportunities to create links between existing data sets and connected vehicle data sets in the future.
- Identify needs and opportunities to integrate national data sets to support performance measurement and asset management purposes.
- Share current activities and best practices related to data strategies, policies, standards, metadata, architecture, procedures and metrics.
- Develop data standards and stewardship recommendations for consideration by the FHWA Data Governance Advisory Council. Data standards are very important to achieve harmonization across stakeholders and should be a prominent data management practice in the rules of engagement.
- Identify opportunities to coordinate resources to reduce data redundancy and implement cost sharing strategies for the collection, management, and maintenance of roadway travel mobility data. Redundancy could be addressed through data standardization and an annual review of data programs to identify where duplicate data collection and storage can/should be eliminated in and replaced with a single source of data for specific data programs. This will help to ensure that data is collected once and used many times.
- Identify opportunities to reduce redundancy in the development and maintenance of duplicate data systems, promote efficiency in system maintenance, and promote open source initiatives.
- Explore methods to enhance access to information and data for the roadway travel mobility data programs. This includes developing web-portals easily accessible by internal and external stakeholders for each of these programs to obtain data and information as needed. This will facilitate sharing of data with internal/external stakeholders, thereby reducing costs associated with data collection.
- Understand and promote the value of data as a U.S. DOT-wide asset.

## 3.5 Outreach Mechanisms

The Data Coordination Framework should incorporate mechanisms for ongoing outreach to identify the internal/external Communities of Interest (COI) for data products and systems; make sure COIs know how to access and use available data and understand the value it can provide; communication

protocols to establish clear lines of communication between the U.S. DOT Mobility Data Coordination Group and data business owners; and ongoing, regular meetings for internal stakeholders to discuss needs, gaps, and opportunities for collaboration. The following outreach mechanisms are recommended to enhance the willingness, understanding, and commitment of stakeholders to embrace data coordination/data management practices:

1. Establish an education plan to gain full understanding and support for data coordination/data management and its intended outcomes at the executive and senior management levels.
2. Establish an education plan or resource documents (e.g., Data Inventory, Data Catalog) to ensure that potential data users understand how to access and make effective use of data products and programs.
3. Establish communication mediums such as periodic newsletters and a full website with collaboration features (e.g., wiki, forums, document repositories, etc.) to outreach and provide updates to the Data Working Groups and Internal Community of Interest on current initiatives, activities and best practices related to roadway travel mobility data and data management practices.
4. Establish a communication protocol and plan for communicating performance measure results to members of the U.S. DOT Mobility Data Coordination Group.
5. If stakeholders agree to formalize the U.S. DOT Mobility Data Coordination Group, it is recommended that the Data Working Groups and internal Community of Interest meet on a regular basis (e.g., semi-annually) to ensure continued coordination on current initiatives, activities, and best practices related to roadway travel mobility data and data management practices. These meetings will serve as a forum for internal stakeholders to discuss needs, gaps and opportunities for collaboration.

# 4.0 Data Management Practices

Data management is defined as the development, execution, and oversight of architectures, policies, practices, and procedures to manage the information lifecycle needs of an enterprise in an effective manner as it pertains to data collection, storage, security, data inventory, analysis, quality control, reporting, and visualization. Data management practices are necessary to acquire, update, describe, standardize, analyze, store and protect data to ensure it can be used.

This section recommends data management practices for U.S. DOT based on experiences of other national and state agencies in data management to support business needs and decisions, as well as the stakeholder needs for data business planning documented in Section 2.0. Appendix B provides a comprehensive review of national and state practices. The recommendations also follow best practices recommended in *NCHRP 666, Volume II: Guide for Target-Setting and Data Management*, Chapter 2: Guide for Data Management.[3]

Each core data system and data product will require the data practices described below, and coordination on these cross-cutting issues should take place through meetings of the Internal Community of Interest. Much of the work needed to perform each of these practices can be completed during the spring of 2013 through the Mobility Data Forums, survey instruments with the Data Working Groups and Internal Community of Interest for roadway travel mobility data programs, and forums for gathering information from stakeholders regarding business needs for data and information for each of these programs.

## 4.1 Data Governance

Data governance is defined as the execution and enforcement of authority over the management of data assets and the performance of data functions. The management of data assets is usually accomplished through a data governance board or council. This role is critical in successfully managing data programs that meet business needs and in supporting a comprehensive data business plan for the organization. Support for a data governance program is performed by data stewards, who ensure data is managed according to the policies established by the data governance board or council. An overview and national best practices in data governance is provided in Appendix C.

It is **recommended** that U.S. DOT Mobility Data Coordination Group stay in close coordination with the FHWA Office of Highway Policy Information's Data Governance Advisory Council as they formulate their charter. It is anticipated that the Data Business Plan will be referenced or used by the FHWA Data Governance Advisory Council in the implementation of an FHWA-wide data governance initiative. Other ongoing data governance activities within U.S. DOT are the FHWA Visualization Working Group, the Connected Vehicle Senior Policy Working Group, and the RDE/OSADP Policy Group. More information on these initiatives is provided in Appendix C.

[3] Cambridge Systematics, Inc. NCHRP 666: Target-Setting Methods and Data Management to Support Performance-Based Resource Allocation by Transportation Agencies, Volume II: Guide for Target-Setting and Data Management. Transportation Research Board, Washington, D.C., 2010.

It is also recommended that individual stakeholder offices consider implementing data governance for the roadway travel mobility data programs they have direct control over. For example, the FHWA Office of Operations, as data business owners for climate data programs such as *Clarus*, could implement a data governance program to oversee the governance of climate data to support these programs. Steps for implementing data governance and an example data governance framework for the FHWA Office of Operations are provided in Appendix D.

## 4.2 Data Acquisition

Data acquisition involves direct collection of data in the field (by state agencies or contractors), assembly of a data product from raw data sources (e.g., Traffic Volume Trend (TVT) reports and Urban Congestion Reports), or acquisition of data from a private data vendor. Collection frequency involves the specific intervals in which data is updated. The entire lifecycle management of data must be considered (e.g., under what condition the data is updated, archived, or deleted).

**Recommendations** include developing a policy to define responsibilities for collecting and updating data; developing guidelines or federal mandates for the collection, update and maintenance of data; identifying opportunities for collaboration between connected vehicle data capture activities and existing data programs to leverage existing data acquisition efforts and processes; and specification of updating methods and responsibilities for updating. An annual review of data programs would identify where duplicate data collection and storage can/should be eliminated and replaced with a single source of data for specific data programs. This will ensure that data is collected once and used many times. Additional recommendations include developing a standard data template format that would foster joint usages and collaboration, minimize data duplications, and improve quality, standard and completeness of datasets. A standard data sharing agreement with third parties is also recommended.

## 4.3 Data Quality

Data quality ensures that the data and information provided by roadway travel mobility data programs meet the litmus test for data quality by providing data that has improved accuracy, timeliness, completeness, validity, coverage, and accessibility.

**Recommendations** include developing a policy to define responsibilities for data quality; adopting data quality standards for the collection, processing, use, and reporting of roadway travel mobility data. This includes requiring the use of metadata for each of these data systems that are used to support these programs. QA/QC procedures should also be documented for each of the data systems, with instructions on how to process data errors. Validation rules and allowable values for coded fields should be developed and incorporated into data systems and repositories. Established validation rules (e.g., the NTCIP 1204 for Environmental Sensor Station observations) should be used to the greatest extent possible.

## 4.4 Data Standards

This activity involves use of metadata standards, data dictionaries and descriptive information, all of which are keys to achieving consistency and integration of data from multiple data sources. The use of metadata standards and formats helps to facilitate the understanding, characteristics, and usage of

data. Metadata provides information such as data name, size, data type, where data is located, how it is associated and data ownership (source: http://en.wikipedia.org/wiki/Metadata). A data dictionary is a repository of metadata.

**Recommendations** include developing a policy to define responsibilities for supplying metadata standards for each type of data set (e.g., weather data, travel data, etc.), data dictionaries and descriptive information for data products; developing metadata guidelines to indicate update frequency, age of data, and specify how data can be used or integrated with other data sources; and coordination with applicable data standards, the Federal Enterprise Architecture Data Reference Model, the Data Quality Act, and open data principles documented in the FHWA report, *Policy Analysis and Recommendations for Implementing Open Data for the Mobility Program*.

## 4.5 Business Analysis Tools

Business analysis tools provide an efficient and cost-effective means for managing, analyzing and integrating roadway travel mobility data, and would support FHWA's needs for analysis at both the national and regional/corridor levels.

**Recommendations** include exploring the potential use of new and improved technology to support the roadway travel mobility data programs. This may include outreach to the vendor community who provides the data collection equipment for these programs. This also includes exploring the use of tools to integrate data from other systems and to enhance sharing of data from these programs with other systems and stakeholders. The types of tools recommended could include the use of XML formats for sharing of data, GPS technology for collection of data and identifying the location of data collected, and GIS tools for geographical display of data.

## 4.6 Data Privacy and Security

This service ensures that policies, standards, and procedures to secure data and protect the privacy of individuals and contributing agencies are clearly documented and distributed to all staff responsible for the collection, maintenance, and distribution of data for roadway travel mobility data programs.

**Recommendations** include coordinating with recommendations related to security, ownership, and liability of data as documented in policy documents completed by the Connected Vehicle Senior Policy Working Group.

## 4.7 Data Storage and Access

One of the primary concerns in initiating the Data Business Plan is to ensure that roadway travel mobility data will be managed in a way that facilitates sharing of data with each program's stakeholders.

**Recommendations** include defining business requirements for data access, analysis and reporting; developing a policy to define responsibilities for data storage, hosting, data retention (archive), and disposal; a policy to define data ownership and dissemination rights; and exploring methods to enhance access to information and data for the roadway travel mobility data programs. This includes

developing web-portals that are easily accessible by internal and external stakeholders for each of these programs to obtain data and information as needed.

## 4.8 Traceability

Traceability illustrates how the roadway travel mobility data programs align with FHWA's business needs for data. Traceability provides a tool for coordination to ensure that everyone's business needs for data are being met. Another dimension of traceability is data provenance. Since much of the same types of roadway travel mobility data can be generated by different systems, it is imperative that the source of each piece of data is known even as the data undergoes processing.

**Recommendations** include establishing internal Data Working Groups and COIs for the roadway travel mobility data programs to examine data and information needs in the common business areas for each program, on a regular basis.

## 4.9 Performance Measures

Performance Measures should be reflective of the business needs for roadway travel mobility data identified in the traceability exercise.

**Recommendations** include identifying measures of effectiveness (both qualitative and quantitative) and implementing a monitoring program to measure the success of data coordination/data management activities and provide confirmation that the program is necessary and is effectively delivering results. Example metrics include the timeframe for getting new forms of data integrated and executed, staff willingness to participate in data coordination/governance/management practices, and value of cost savings in generated data that is realized from having the Data Coordination Plan in place. Performance metrics for data management could include the areas of accuracy, accessibility, timeliness, completeness, etc.

It is also recommended to establish a communication protocol and plan for communicating performance measure results to executive level staff and internal Data Working Groups and Communities of Interest. The group should also coordinate with FTA, as they are establishing a new office in accordance with MAP-21 to determine how National Transit Data should be presented and measured.

## 4.10 Risk Assessment

Risk assessment documents: 1) how much data is needed; 2) how accurate the data should be; 3) what the refresh rate of the data should be; 4) who should have access to the data; and 5) many other questions which help to assess the risks associated with a particular data program.

**Recommendations** include performing a risk assessment using a risk assessment matrix which identifies each data set used to support the roadway travel mobility data programs and current and potential risks (e.g., loss of data) associated with these programs. The risk matrix should include the name of the datasets and the primary (data business owners) and secondary (data stewards) persons responsible for addressing any issues regarding potential or actual loss of data or data quality issues.

This risk assessment should be used to develop a Risk Management Plan to address risks if/when they occur. Risk management practices should include disaster/recovery procedures.

## 4.11 Value of Data Programs

The value of the data programs need to be demonstrated to the stakeholders and users and those who authorize investments in the data programs.

**Recommendations** include demonstrating the value of roadway travel mobility data programs by linking the data sets used to support these programs to business services and mission and goals of U.S. DOT. Visualization tools, including graphs/charts can be prepared to demonstrate this linkage. Charts can also be used to illustrate the percent of inquiries received on a daily, monthly, annual basis, which require data and information from these programs.

## 4.12 Knowledge Management

Knowledge Management is needed to organize, categorize, store, and archive knowledge about the data sets and workflow processes used to support the roadway travel mobility data programs. This is applicable to internal Data Working Groups and Communities of Interest only.

It is recommended that U.S. DOT develop and implement a Knowledge Management (KM) system as part of the data coordination framework to ensure that lessons learned and experiences pertaining to business operations within the organization are retained and archived as staff retire or leave the organization.

# 5.0 Conclusion and Next Steps

## 5.1 Conclusion

This section summarizes how the original goals and objectives were met through the development of the Data Business Plan. The implementation process will be important in finalizing and activating the plan.

- GOAL 1. Improve the coordination and communication mechanism across U.S. DOT and FHWA offices involved with *roadway travel mobility data* to avoid investing resources in the same or similar types of data related programs.

  This will be accomplished through implementation of the Data Coordination Plan and development of a web portal to facilitate coordination.

  - Objective 1.1. Identify gaps and redundancies in roadway travel mobility data programs.

    *This was accomplished in Section 2.0 above.*

  - Objective 1.2. Devise "rules of engagement" regarding collaboration of the data functions for roadway travel mobility data.

    *This will be accomplished in implementation of the Data Coordination Plan.*

- GOAL 2. Improve the coordination of the data capture activities associated with ITS JPO-sponsored research in wirelessly connected vehicle technologies with roadway travel mobility data programs within U.S. DOT and FHWA.

  The Data Coordination Plan will be the vehicle for accomplishing this goal.

  - Objective 2.1. Identify how current and planned data from the connected vehicle initiative can support roadway travel mobility data programs.

    *This was identified in Section 4.2 of the Stakeholder Needs Description.*

  - Objective 2.2. Identify how data from roadway travel mobility data programs within U.S. DOT and FHWA can support the connected vehicle initiative.

    *This was identified in Section 4.2 of the Stakeholder Needs Description.*

  - Objective 2.3. Identify existing/future data inventory and data structures/ policies/governance practices that could be applicable to the Research Data Exchange.

    *This was accomplished in the Data Inventory and review of national best practices in Data Governance (Appendix C).*

## 5.2 Next Steps

The next step for the Data Business Plan project is to begin Phase 2, which includes the following tasks:

- **Execution of Data Business Plan Coordination.** This would include implementation of the Data Coordination Framework, including formalization of the U.S. DOT Mobility Data Coordination Group and implementation of the data management practices recommended in this document. The first step to implementation should be to establish the Travel Data Working Group and apply the Data Coordination Framework described in Section 3.0 to this smaller group in order to test the various framework components (e.g., develop a MOU, rules of engagement, outreach mechanisms, etc.). An example MOU is provided in Appendix G.
- **Data Integration Test Pilots.** The data integration test pilots will demonstrate the benefits and value of the Data Business Plan by showing how a systematic data coordination process could yield tangible benefits in terms of reduced cost or improved federal capability; integrating proposed and existing national or project data sets; and eliminate gaps and overcome redundancies in national and project data sets associated with mobility data. The following outcomes could result from the data integration test pilot projects:
  - Elimination or enhancement of certain data elements in some national data sets,
  - Creation of a link between an existing data set and a connected vehicle data set in the future;
  - Maximize integration of national data sets for performance measurement and asset management purposes;
  - Investigation of how regional data sets can best be combined with national ones;
  - Recognition of how a project data set (such as TTID) could be leveraged in the future;
  - Creation of value added to the Research Data Exchange or other U.S. DOT research data sets; and
  - Documentation of lessons learned for the U.S. DOT, State DOTs and regions.
- **Execute data governance within the FHWA Office of Operations.** Data governance should be executed within the Office of Operations based on the example Data Governance Framework in Appendix D. The Office of Operations should also meet with the Chairperson of the FHWA Data Governance Advisory Council to identify opportunities for coordination and/or collaboration with that group. Upon implementation, the Data Business Plan should transition to a U.S. DOT-owned document with a federal perspective and include a data governance structure, data governance charter, and data governance manual.

# APPENDIX A. Glossary of Terms

This appendix provides a glossary of terms related to data coordination, management, and governance.

**Business Intelligence** – The ability for an organization to collect, maintain, and organize knowledge, thereby automating some current manual processes. BI technologies provide historical, current and predictive views of business operations. The goal of modern business intelligence deployments is to support better business decision-making.

**Climate (Weather) Data** – Climate data included data on weather and pavement surface conditions collected from road weather information systems (RWIS) and environmental sensor stations (ESS).

**Community of Interest (Internal)** – Any persons or offices internal to U.S. DOT that collect own, maintain, use or interface with, access, benefit from, or are otherwise affected by roadway travel mobility data. This group would coordinate and share best practices on data management issues such as data quality, standards, data privacy and security, etc.

**Community of Interest (External)** – Any persons or offices external to U.S. DOT that use or interface with, access, benefit from, or are otherwise affected by roadway travel mobility data.

**Connected Vehicle Data** – Data collected via a vehicle that has an independent onboard wireless capability to establish a two-way data linkage between a system onboard and another system not onboard, for the purpose of transferring information.

**Data Business Owners** – Individuals who manage the data and metadata for information systems within their area of responsibility for a business unit (roadway travel mobility data). Data business owners are responsible for maintaining the data dictionaries for the data systems and for establishing business requirements for the use of roadway travel mobility data.

**Data Business Plan** – Describes a systematic process for U.S. DOT to follow while conducting activities related to the collection, management, and maintenance of roadway travel mobility data.

**Data Catalog** – A catalog of information about the data used by stakeholders involved with roadway travel mobility data programs within U.S. DOT. The data catalog includes a list of relevant data programs, data business owners, data stewards, and instructions for accessing data standards and definitions with that program.

**Data Governance –** The execution and enforcement of authority over the management of data assets and the performance of data functions. The management of data assets is accomplished through the data governance team. This role is critical in successfully managing data programs that meet business needs and in supporting a comprehensive data business plan for the organization.

**Data Governance Charter** – Sets forth the purpose, mission, vision, goals and objectives, and data management policies for implementation of the Data Governance Framework.

**Data Governance Framework** – A formalized structure for organizing and communicating data governance and data management concepts.

**Data Governance Model Diagram** – A diagram illustrating the data governance framework specific to an organization.

**Data Governance Team** – The designated individuals within U.S. DOT stakeholder offices responsible for the oversight of data programs to support the business functions of the office.

**Data Inventory** – Provides a high level and detailed inventory of relevant national, State, and regional data programs that support the goals and objectives of the Data Business Plan. The Data Inventory documents all stakeholders and the data programs they are using, their objectives, and how connected vehicle data could benefit from these programs.

**Data Management** – The development, execution, and oversight of architectures, policies, practices, and procedures to manage the information lifecycle needs of an enterprise in an effective manner as it pertains to data collection, storage, security, data inventory, analysis, quality control, reporting, and visualization.

**Data Management Maturity Model** – A helpful tool for documenting levels and characteristics of maturity related to the development and implementation of data management and data governance principles as they relate to employees/processes, technology/tools, and institutional governance.

**Data Mapping** – A graphic that the illustrates the relationship between U.S. DOT's mission, U.S. DOT stakeholder offices' business objectives, program offices, and core data systems that support roadway travel mobility data.

**Data Management Practices** – Activities necessary to acquire, update, describe, standardize, analyze, store and protect data to ensure it can be used.

**Data Stewards** – Individuals who ensure data is managed according to policies established by the Data Governance Team.

**Data Stewardship –** The formalization of accountability for the management of data resources. Data stewardship is a role performed by individuals within an organization known as data stewards. The functions of data governance and data stewardship typically are part of an overall data management program within an organization.

**Data Working Group** – Any persons or offices internal to U.S. DOT that collect or manage roadway travel mobility data (infrastructure/ inventory, travel data, climate data, modal data, and connected vehicle data capture activities). This group dictates the policies, procedures, and business practices associated with roadway travel mobility data programs.

**External Stakeholders** – Any persons or offices external to U.S. DOT that are customers of the various roadway travel mobility data programs.

**FHWA Data Governance Advisory Council** – Advisory board formed by the Office of Highway Policy Information as part of the Data Integration Platform project. The goal of the Data Governance Advisory Council is to emphasize the corporate management of data, including reducing duplication, increasing compatibility, and making sure all offices have better policies in place when implementing new data programs. The initial emphasis of the Data Integration Platform will be on integrating data from internal programs such as the Federal Emergency Management Information System (FEMIS), Highway

Performance Monitoring System (HPMS), National Bridge Inventory (NBI), and the Recovery Act Database System (RADS). They will eventually incorporate other data programs such as the Fatality Analysis Reporting System (FARS), Federal Railroad Administration (FRA) databases, and the Freight Analysis Framework (FAF).

**Infrastructure (Inventory) Data** – Includes roadway geometry, roadway inventory, intersection characteristics, and the state of system controls.

**Internal Stakeholders** – Any persons or offices internal to U.S. DOT that are data collectors and users of roadway travel mobility data programs.

**Knowledge Management –** Systems, strategies, processes, or programs for capturing lessons learned from previous project experiences, dedicated to sharing and retaining critical organizational knowledge and used as a repository to store corporate knowledge about business process workflows and data and systems used to support such programs as the Highway Performance Monitoring System (HPMS).

**Maturity Model –** A helpful tool to assess where an organization stands with respect to implementing certain processes. A maturity model also can be used to benchmark for comparison or assist an agency in understanding common concepts. A typical maturity model identifies levels and characteristics of those levels. The model can be used to assess an agency's status and assist in identifying next steps to achieve success toward an ultimate goal state.

**Research Data Exchange** – A transportation data sharing system that promotes sharing of both archived and real-time data from multiple sources (including vehicle probes) and multiple modes.

**Roadway Travel Mobility Data** – Data from roadway travel modes, including vehicle, truck freight, bicycle/pedestrian, and transit, as well as roadway inventory data.

**Roadway Travel Mobility Data Programs** – Formally recognized data programs within U.S. DOT that involve the collection, analysis or reporting of roadway travel mobility (including vehicle, truck, freight, bicycle/pedestrian, and transit) and roadway inventory data.

**Rules of Engagement** – Practices followed or behavior displayed by the participants in situations of opposing interests such as negotiations. Unwritten rules of engagement determine what information is given, at what time, to whom, and in what manner; and what concession is granted and what is demanded in return. For work in a team, rules of engagement typically define the protocols of communication, conflict, decision-making and meetings.

**Travel (Speed and Volume) Data** – Travel data includes vehicle location, presence, and speed within the system, internal vehicle status such as fuel consumption rate, or externally measured data such as recorded external temperature.

**U.S. DOT Mobility Data Coordination Group** – The designated individuals within U.S. DOT responsible for the oversight of data programs to support the business functions of their offices.

**U.S. DOT Mobility Data Coordination Group Team Leader** – A designated individual from within the FHWA Office of Operations who would chair the Mobility Data Coordination Group and serve as a liaison for other U.S. DOT Data Governance initiatives.

# APPENDIX B. Best Practices in Data Management

Appendix B explores best practices in data management based on agency experiences at the national and state levels. These practices are presented to facilitate development and implementation of data services to support roadway travel mobility data programs.

## B.1 National Best Practices in Data Management

Many organizations at both the national and state level have some kind of program for managing data and information. At the national level, many of the data management practices at U.S. DOT, for instance, are focused on providing data to support particular programs, such as Asset Management, Pavement Management, Bridge Management, and assessing the performance of the nation's transportation system using the Highway Performance Monitoring System (HPMS). The state DOTs also have similar programs for managing data to support these type of systems at the state level.

This section explores some of the best practices at the national level, which can be used to improve data management practices at U.S. DOT as part of their Data Business Plan.

### U.S. Department of Education – National Center for Education Statistics

The source of the following information is *Data Systems Standards and Guidelines – Planning, Designing, Building Data Systems That Work* (http://nces.ed.gov/dataguidelines/guides.asp#planning).

Federal law established the Hawkins-Stafford Education Improvement Amendments of 1988 to "produce and maintain, with the cooperation of the States, comparable and uniform education statistics." In support of this law, the National Center for Education Statistics (NCES) provides guidance to states in the "collection, maintenance, and use of elementary and secondary education data."

Several best practice guides have been developed by the NCES to assist the states in this effort. Each of these guides can be categorized into 4 main areas:

- Planning the Data System
  - Decision support systems and how they can be implemented;
  - Metadata and its importance in promoting data quality; and
  - Technology guide to improve procurement of technology solutions for managing education data.
- Managing the Data
  - Disaster recovery guidance; and
  - Confidentiality issues.

- Data Quality Assurance
  - Improving data quality; and
  - Curriculum Guide – helps improve quality of data collected.
- Using Data
  - Data ethics for use of education data.

***Applicability to U.S. DOT***

Each of the four categories described above includes guidance commonly identified as best practices in data management, which should be adopted by U.S. DOT as part of their overall data management strategy. These include the use of:

- Guides which document policies, standards, and roles/responsibilities for data management;
- Standards for data quality;
- Risk Management practices, which include disaster/recovery procedures; and
- Metadata and its importance for ensuring the right data is used to respond to inquiries.

## Oak Ridge National Laboratory: Best Practices for Preparing Environmental Data Sets to Share and Archive[4]

This report discusses best data management practices that data collectors and providers should follow to improve the usability of their data sets. The report focuses on the preparations for sharing of data, preservation of data and archiving data. The report identifies seven best practices for preparing environmental data sets to share:

1. Define the Contents of Data Files;
2. Use Consistent Data Organization;
3. Use Consistent File Structure and Stable File Formats For Tabular and Image Data;
4. Assign Descriptive File Names;
5. Perform Basic Quality Assurance;
6. Assign Descriptive Data Set Titles; and
7. Provide Documentation.

Each of these practices should be included in any comprehensive data management program.

***Applicability to U.S. DOT***

One of the primary concerns of U.S. DOT in initiating the development of their Data Business Plan is to ensure that data for the roadway travel mobility data programs will be managed in a way that facilitates sharing of data with each program's stakeholders and which also allows for integration of

[4] Hook, L., S. Vannan, T. Beaty, R. Cook, and B. Wilson. Best Practices for Preparing Environmental Data Sets to Share and Archive. Oak Ridge National Laboratory, 2010, http://daac.ornl.gov/PI/BestPractices-2010.pdf.

other data sets (such as GIS and GPS data) to improve data quality (accuracy, timeliness, completeness, etc.) within each program.

Each of the seven best practices listed above (or some form of several of these) could be included in a Data Catalog. The data catalog provides a centralized location for identifying data systems; data elements, data definitions, and data file structures, all of which are needed by the data stewards, and data custodians responsible for managing data according to established policies and standards. A draft Data Catalog has been developed for U.S. DOT as identified in Appendix D of this Data Business Plan.

## National Oceanographic Data Center (NODC)

The following information describes how the National Oceanographic Data Center (NODC), is used to collect and share oceanographic data collected from the United States and other countries. The NODC operates as a component of the National Environmental Satellite, Data, and Information Service (NESDIS) of the National Oceanic and Atmospheric Administration (NOAA) of the U.S. Department of Commerce (DOC). The NODC provides data management services for physical, chemical, and biological data from the World ocean and adjacent waters.[5]

Data Management practices at NOAA include two basic components: data management services and data stewardship. Each of these components is recognized as critical to coordinating the "movement of data and information from the observing system sensors to the data user. This process includes the acquisition, quality control, metadata cataloging, validation, reprocessing, storage, retrieval, dissemination, and archival of data."

NOAA has also issued an Administrative Order (NAO 212-15), which establishes the scope and policy for managing environmental data and information.[6] Portions of this policy are listed below as an example for developing a similar data management policy for U.S. DOT:

***SECTION 2. SCOPE***

*This NAO applies to all NOAA environmental data and to the personnel and organizations that manage these data, unless exempted by statutory or regulatory authority.*

***SECTION 3. POLICY***

*01 Environmental data will be visible, accessible and independently understandable to users, except where limited by law, regulation, policy (such as those applicable to personally identifiable information or protected critical infrastructure information or proprietary trade information) or by security requirements.*

*02 Management of NOAA environmental data will be based upon an end-to-end data management lifecycle that includes:*

- *Determining what environmental data are required to be preserved for the long term and how preservation will be accomplished;*

[5] http://www.ncddc.noaa.gov/activities/science-technology/data-management/.
[6] http://www.ncddc.noaa.gov/activities/science-technology/data-management/.

- *Developing and maintaining metadata throughout the environmental data lifecycle that comply with standards;*
- *Obtaining user requirements and feedback;*
- *Developing and following data management plans that are coordinated with the appropriate NOAA archive for all observing and data management systems;*
- *Providing for the delivery to the archive and secure storage;*
- *Enabling integration and/or interoperability with other information and products; and*
- *Conducting scientific data stewardship to address data content, access, and user understanding.*

## Applicability to U.S. DOT

Each of the data management practices listed in NAO 212-15 Section 3, Subsection 02 above are similar to other agency practices; however, of particular importance is the use of obtaining user requirements and feedback. In the context of U.S. DOT, this would include establishing consistent methods for soliciting comments and feedback from the user groups (stakeholders) for the roadway travel mobility data programs. This could be accomplished through the use of quarterly, semi-annual, or annual meetings with the stakeholder communities of interest to ensure that these programs continue to meet the needs of all users, both internal and external to U.S. DOT.

## Data Management Checklist

The following Data Management Checklist was designed as a data **planning** checklist by MIT Libraries for data used in research projects. The checklist is part of a guide on "Data Management and Publishing" available from MIT Libraries.[7]

It provides examples of the types of questions that should be addressed by internal business data owners at U.S. DOT and also posed to internal and external stakeholders as part of the implementation of the Data Business Plan.

1. What type of data will be produced? Will it be reproducible? What would happen if it got lost or became unusable later?
2. Who will use it now, and later?
3. How long should it be retained? e.g., 3-5 years, 10-20 years, permanently
4. Are there tools or software needed to create/process/visualize the data?
5. Any special privacy or security requirements? e.g., personal data, high-security data
6. Any sharing requirements? e.g., funder data sharing policy
7. Is there good project and data documentation?
8. What directory and file naming convention will be used?
9. What file formats? Are they long-lived?
10. Storage and backup strategy?
11. Is there an ontology or other community standard for data sharing/integration?
12. Who in the research group will be responsible for data management?

[7] http://libraries.mit.edu/guides/subjects/data-management/checklist.html; http://libraries.mit.edu/guides/subjects/data-management/evaluate.html#retained.

***Applicability to U.S. DOT***

The questions listed above will help U.S. DOT to implement a comprehensive, well-designed data management plan by documenting the following important components of a successful data management program:

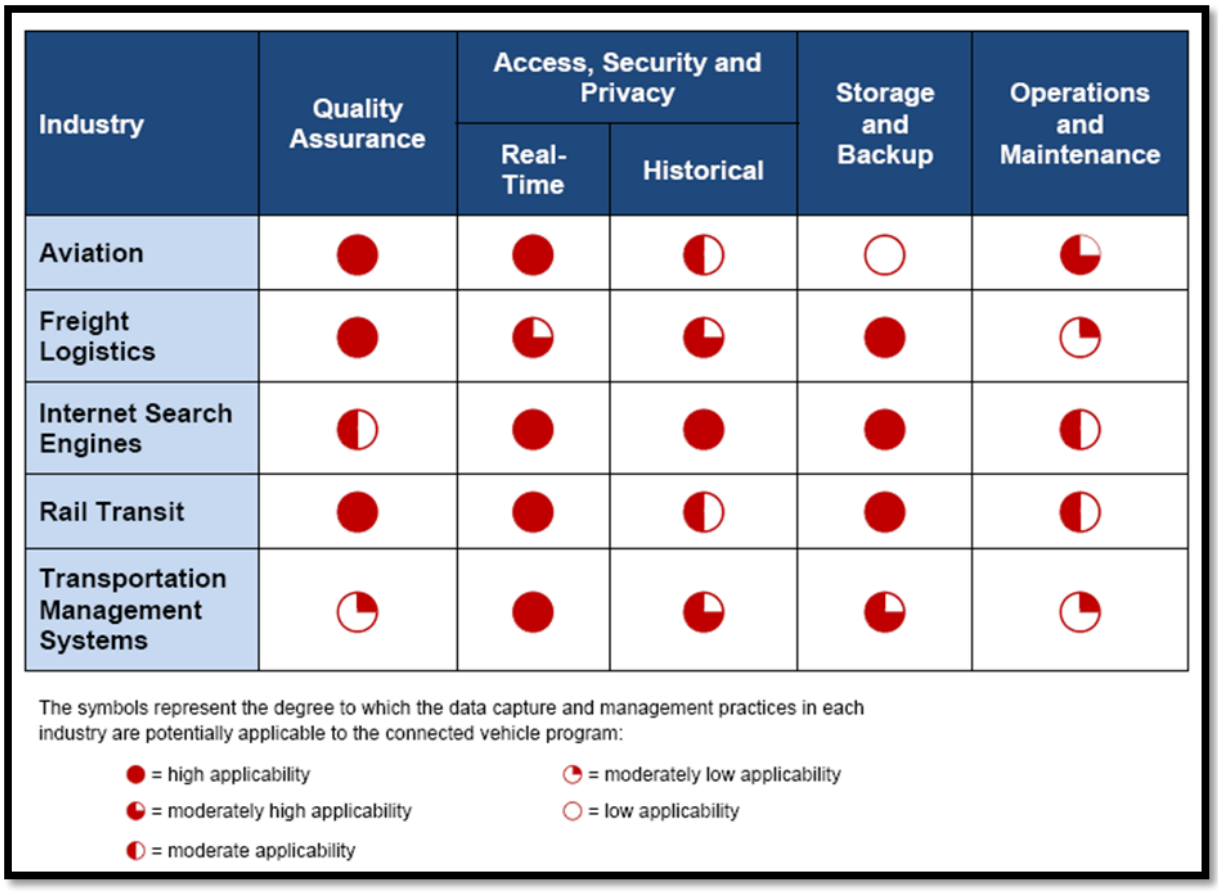
- Policies and procedures for sharing of data from the roadway travel mobility data programs;
- Roles/responsibilities of data collectors, data managers (data business owners, data stewards, data custodians) and data users (data stakeholders, communities of interest;
- Data standards for collection and reporting of roadway travel mobility data programs;
- Risk Management strategy to secure the data and limit risks to the FHWA Office of Operations for the roadway travel mobility data programs;
- Technology (hardware/software) needed to sustain and improve the roadway travel mobility data programs; and
- Data catalog with data definitions and file formats used to store roadway travel mobility data.

## Real-Time Data Capture and Management State of the Practice Assessment and Innovations Scan

The *Real-Time Data Capture and Management State of the Practice Assessment and Innovations Scan*[8] addresses issues related to data capture, data management, archiving, and sharing collected data to encourage collaboration, research and operational development and improvement. The scan covers five industries: Aviation, Freight Logistics, Internet Search Engines, Rail Transit Systems, and Transportation Management Systems. The potential relevance that each industry has for six relevant categories and for the connected vehicle program in general is captured in Figure B-12.

[8] Science Applications International Corporation. Real-Time Data Capture and Management State of the Practice Assessment and Innovations Scan: Lessons from Scan of Current Practices. April 2011.

**Figure B-1. Potential Applications of Current Industry Practices (**Source: Real-Time Data Capture and Management State of the Practice Assessment and Innovations Scan: *Lessons from Scan of Current Practices*, Science Applications International Corp., April 2011)




The following recommendations (excerpted from the source document) identify data management practices and considerations from the five industries that are applicable to the connected vehicle data capture program:

- Control the amount of information being transmitted. Information overload can cause critical messages to be overlooked.
- Start small with an implementation that addresses the most critical needs, either defined geographically or by category of information. Ensure that core competencies are answered before trying to do everything.
- Build for scalability. Avoid a situation where a system is built to perform very well for a test setup, but does not scale well in the real-world. One way to do this is to leverage technologies such as clustered databases, virtual warehousing, virtual servers, etc.
- Determine what data really needs to be stored and backed up. Once the system is out of test mode, there is no need to retain all information (e.g., once vehicles have safely traversed an intersection)? Rather than saving the entire transaction, only

critical data such as vehicle counts and other positional information could be saved, while the unneeded data is discarded.

- Consider allowing a third party to handle data storage needs. The cost of keeping all traffic data may be prohibitive for the government, but profitable for a third party. The pros and cons of using a third party, particularly in a cloud computing environment (which should be considered for data capture and management for the connected vehicle program), are as follows:
  - Pros: Reduced cost, low maintenance, redundancy, high availability, easy access, and speed.
  - Cons: Sacrifice of control – risk of hacking; could be difficult to get the data back; if web based, loss of internet is loss of access; and data may be mined for usage patterns.
- Determine who owns the data from the start and who can use the data.
  - Who pays for the data collection, storage, and dissemination? If the government pays for and owns the data, is it free to anyone? Who can sell the data?
  - Need to address privacy concerns. Applications that track and identify vehicles (commercial or private) through technology such as GPS have already come up against resistance because people do not want the government to have access to where they are driving.
  - Need to determine who will have access to the data and what they will have access to. Would a subset of the data allow the user to perform their function? This is another instance in which having the data collected and managed by a third party may mitigate some of the concerns. Having a third party collecting and aggregating the data, stripping it of identifying markers, and guaranteeing the security of the raw data, can help with privacy issues. The government would only receive the data they need without divulging the information for a single vehicle.
  - Determine if the data can/will be used by any other public sector agency for any other purpose than the originally stated reason to collect the data. One concern of the private sector or private citizens is that the data will be used for purposes to which they did not agree, such as for enforcement. For example, GPS data could reveal the speed at which the vehicle is traveling, or the length of time a truck driver has been on the road without a break (important for truck hours of service rules), etc.
- Make data available as soon as feasible. Even if more processing needs to occur in the background, providing a real-time, or pseudo-real-time, feed for current data is always an attractive option to users (and potential users) of the system.
- Segregate critical data-processing elements behind a set of filtering mechanisms. For example, if a set of algorithms are vital in terms of output, but can be overloaded by pumping too much data through them, provide intermediate filtering algorithms that can be leveraged to consolidate / prioritize the data that is passed to the critical system. If the critical system experiences an overload, the filter could be adjusted to reduce the load on the critical system.

The scan documented the following best practices for **quality assurance**:

- Data quality, particularly in the area of safety, can be improved by gathering redundant data from multiple sensors. Data can be combined so that false positives are filtered out. For example, the maritime and air industries real-time tracking applications are used primarily for safety applications. It is necessary to know the precise locations of the aircraft and vessels, not only for collision avoidance, but also for response to weather conditions. These data need to be very accurate and timely.
- Use standard industry reference files when possible. This helps in reducing the possibility of erroneous information. In the financial industry, error rates in the documents and transactions have fallen considerably with the advent of electronically transmitted documents, data and payments.
- Overall data quality is highly industry dependent. It is typical to provide fast, general data-quality analysis for real-time systems, and then thoroughly scrub and sanitize the overall data more diligently for historical and post real-time analysis and display.

The scan documented the following best practices for **access, security and privacy**:

- Access is most often controlled by the holder of the data. Access is carefully controlled within the transportation and logistics community.
- Systems have been designed so that the right people have access only to the data they need. Despite the large streams of data containing a considerable number of data elements, data within the transportation and logistics industry embed a trade and market sensitivity that would catastrophically impact shippers and carriers if the data went to the wrong hands. It would have serious implications on business and competitiveness.
- Access to the data is usually password protected.
    - Much of the data moving in the logistics systems is financial in nature and a high degree of encryption has been placed on the data as it moves.
    - There are several applications that capture tracking data for aircraft and vessels and display them, often along with other identifying information. These applications are available to almost anyone without any security check or password protection (unless users have purchased one of the systems that offer additional information or services).
- Protection of data source is highly important. In fact, within the search industry, it is so highly protected that there is no concrete evidence of exactly how it is protected.

The scan documented the following best practices for **data storage and backup**:

- Frequent backups and storage off site are typical.
- Perform preventative maintenance regularly.
- Considerable thought is needed to determine what data actually needs to be kept and for how long. This remains a problem in most industries. For example, in aviation, data are typically kept for a brief period of time before being discarded

(real-time data are paramount, historical data are not). In the case of an incident, however, data are spooled off for review before being destroyed.

The scan documented the following best practices for **operations and maintenance**:

- During development, start small, do not overload. Small could be defined geographically or by category of data.
- Use multiple servers to distribute the load for real time databases. Databases can grow quite large very quickly. Many technologies exist to achieve this, and the decision is easier once the approximate data sizes and elements have been defined.
- Considerable effort is needed to decide what level (granularity) of data are needed and how much and how often it should be transmitted. The possibilities of data overload (which will effectively negate the usefulness of the data) are high, and it is important to avoid an overreaction or early reaction based on small sample sizes.
    - In the logistics field, inventory data are available up to the minute in many retail stores. These are used not only for reordering to restock, but also to determine trends to choose items to sell in the future.
    - In the search engine industry, data are typically refreshed according to a 24-hour cycle; so once the data has been refreshed, it is 'static' until the next cycle begins.
    - In the aviation field, data are gathered (for the most part) as fast as possible in order to provide the ability to preemptively prevent incidents.
- Determine what is critical to communicate. For instance, the alert systems that are in operation for railroads and airlines do not collect continuous data, but rather data which are needed to alert an operator of a problem.

The scan documented the following best practices for **critical failures**:

- One major source of problems is that correcting a failure or issue is often dependent on a single person, making it difficult to rectify the problem quickly.
- If any system failure would be catastrophic, then it is necessary to keep round-the clock staff available to fix issues. Elevated labor costs are needed for the system to be highly available.

***Applicability to U.S. DOT***

The data management practices and considerations that are applicable to U.S. DOT (specifically to the connected vehicle data capture activities) can be listed as follows:

- Avoid information overload (Control the amount of information being transmitted).
- Start small during development and address critical needs first. "Small" could be defined geographically or by category of data.
- Build for scalability (Make sure that the system scales well in the real-world).
- Determine what data needs to be stored and backed up; and discard unneeded data.
- Determine the data users and owners from the start.

- Make data available as soon as feasible (provide real-time or pseudo real-time data); and
- Segregate critical data-processing elements behind a set of filtering mechanisms.

## Policy Analysis and Recommendations for the Data Capture and Management Program: Implementation of Open Data Policies and System Policies for the Research Data Exchange and Data Environments[9]

The case for adopting an open data policy is supported by current U.S. and international examples within the public sector. There are also several established licensing options that would facilitate implementation while addressing important liability questions pertaining to ownership and intellectual property. The primary advantages realized by an open data policy include increased access to information from taxpayer-funded systems, greater information sharing across organizations, and a readily available source of high-quality real-time data that encourages innovative applications and improved operational efficiency. In order to deliver these benefits, an effective open source policy must be widely accessible and cost-effective while addressing the risks concerning security, privacy, liability, and data quality.

Because the RDE architecture and technologies have not yet been chosen, there are a number of research actions/inputs that are needed to develop full policies in each area.

***Applicability to U.S. DOT***

Until further technical definition of the system architecture, technologies, and datasets are provided, there are two overall conclusions and one set of prospective analysis worth noting:

- **Implement based on an open data policy.** An open data policy is a viable option and is encouraged by the U.S. Government in general and is emerging as a trend with other governments around the Nation and around the world. The level of ‘‘openness’’ is highly dependent upon some of the technical inputs – the accessibility of the RDE to public users; the critical and minimum characteristics of the data that will be captured, used, stored, and archived; and the risks/trade-offs associated with the technical definition of what it means to be open. This paper and other related mobility policy reports attempt to put some definition to these open questions. There is a need to have the whole set of reports and definitions vetted by the technical team and stakeholders to ensure that the basis for recommending policies is solid.
- The RDE system policies can be based on proven solutions; however the federation policies require further analysis and development. The alternatives regarding the RDE architecture and set of technologies that are proposed for use in the construction and operation of the RDE appear synonymous with other

[9] Brecher, A., J. Hassol, and S. Sloan. Policy Analysis and Recommendations for the Data Capture and Management Program: Implementation of Open Data Policies and System Policies for the Research Data Exchange and Data Environments. Draft Report, March 2012.

portals in use with the Federal and State governments, academia, and industry. As a result, most of the RDE system policy can draw from existing models. The key differences, though, from a policy perspective include the wide-scale federation and the monitoring and enforcement of policies through such a dispersed system. Once decisions are made about the architecture and technologies, developing a set of alternative models of operation with supporting policies (also referred to as "scenarios") are a useful next step to determine how the technical, policy, and institutional recommendations align.

- Analysis: RDE Next Steps. It is not clear whether the RDE is available specifically as a research tool or if the RDE is meant as a prototype that demonstrates how real-time data capture and exchange and new data environments might work. If the latter, further analysis on technology transfer, steps and policies to support commercialization, and the viability of sustainable marketplaces will be needed.

## Industry Options: State-of-the-Practice Policies and Lessons Learned on Open Data and Open Source[10]

The document, *Industry Options: State-of-the-Practice Policies and Lessons Learned on Open Data and Open Source*, recommended policies for the Data Capture and Management (DCM) and Dynamic Mobility Applications (DMA) programs. The recommended policies can be summarized as follows:

- Metadata: Metadata standards minimize the potential for inconsistency caused by the creation or modification of metadata by numerous participants. Adoption of the ASTM International Standard Practice for Metadata to support Archived Data Management Systems (E 2468-05) is logical to support both the research and the operationalization of metadata to support DMA/DCM applications. Furthermore, the adoption of a modified Dublin Core[3] as the metadata schema is recommended. Since both these elements have widespread industry acceptance, they would ensure a smooth transition to private sector development of DCM/DMA applications.
- Data Security: A robust policy framework as well as guidelines exist to ensure security of data sets, data environments, and the hardware and software associated with both.
- Data Privacy: Developing a privacy policy requires development of Fair Information Practices (FIPs). Also, policies and guidelines for security are a key element of a privacy policy since data privacy is highly interrelated with data security.
- Intellectual Property: A thorough, well-documented and clearly communicated IP policy framework is necessary to provide a clear understanding of the rules of the game with respect to licensing, patents, and other aspects of intellectual property protection.

[10] Brecher, A., M. Cuddy, J. Hassol, and S. Sloan. Industry Options: State-of-the-Practice Policies and Lessons Learned on Open Data and Open Source. Draft Report, February 2011.

- Liability: With regards to risks with open data, data errors, and unintentional system problems, the U.S. DOT's connected vehicle legal policy team is exploring the concept of a shared risk environment, and analyzing the extent and applicability of the existing tort law. The recommendation is to apply the results of this analysis to the DCM and DMA programs.
- Governance: A thorough, well-documented policy needs to clearly state the reasons for implementing data governance, which includes ensuring that existing roadway travel mobility data programs are managed in a way that provides continued support to the U.S. DOT stakeholder offices in meeting their business needs. The policy also needs to identify the offices/persons responsible for overseeing data governance.

***Applicability to U.S. DOT***

It is recommended that the policies for metadata, data security, data privacy, intellectual property, liability, and governance be considered in the implementation of the Data Business Plan.

# B.2 State Best Practices in Data Management

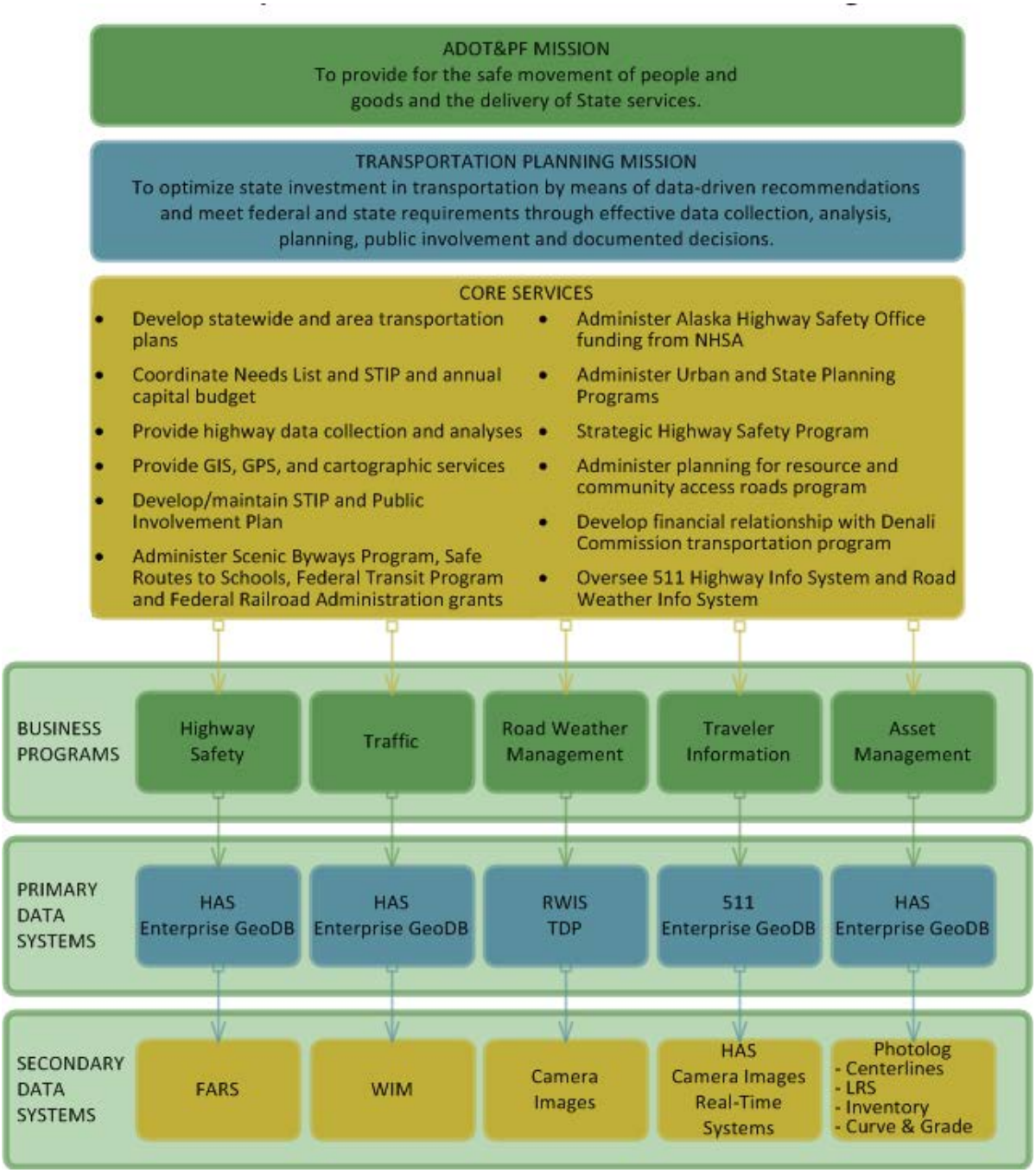
There are also several examples of excellence in data management practices demonstrated at state agencies including DOTs. These practices include the implementation of data business plans, establishing data management policies, implementing data governance frameworks, implementing the use of Business Intelligence tools like knowledge management systems and Geographic Information Systems (GIS), and developing data catalogs and data governance manuals to define the roles and responsibilities of offices in managing data and information to support business needs.

Some of these examples are described in the paragraphs below to demonstrate how these practices can be adopted at U.S. DOT as part of their Data Business Plan.

## Alaska Department of Transportation and Public Facilities (ADOT&PF)

The Alaska Department of Transportation and Public Facilities (ADOT&PF) is improving the management of data and information systems in support of the overall Agency and Planning Missions of the Department. ADOT&PF has identified the critical primary and secondary data system that are needed to support the core business functions of the Program Development Division as illustrated in Figure B-2.

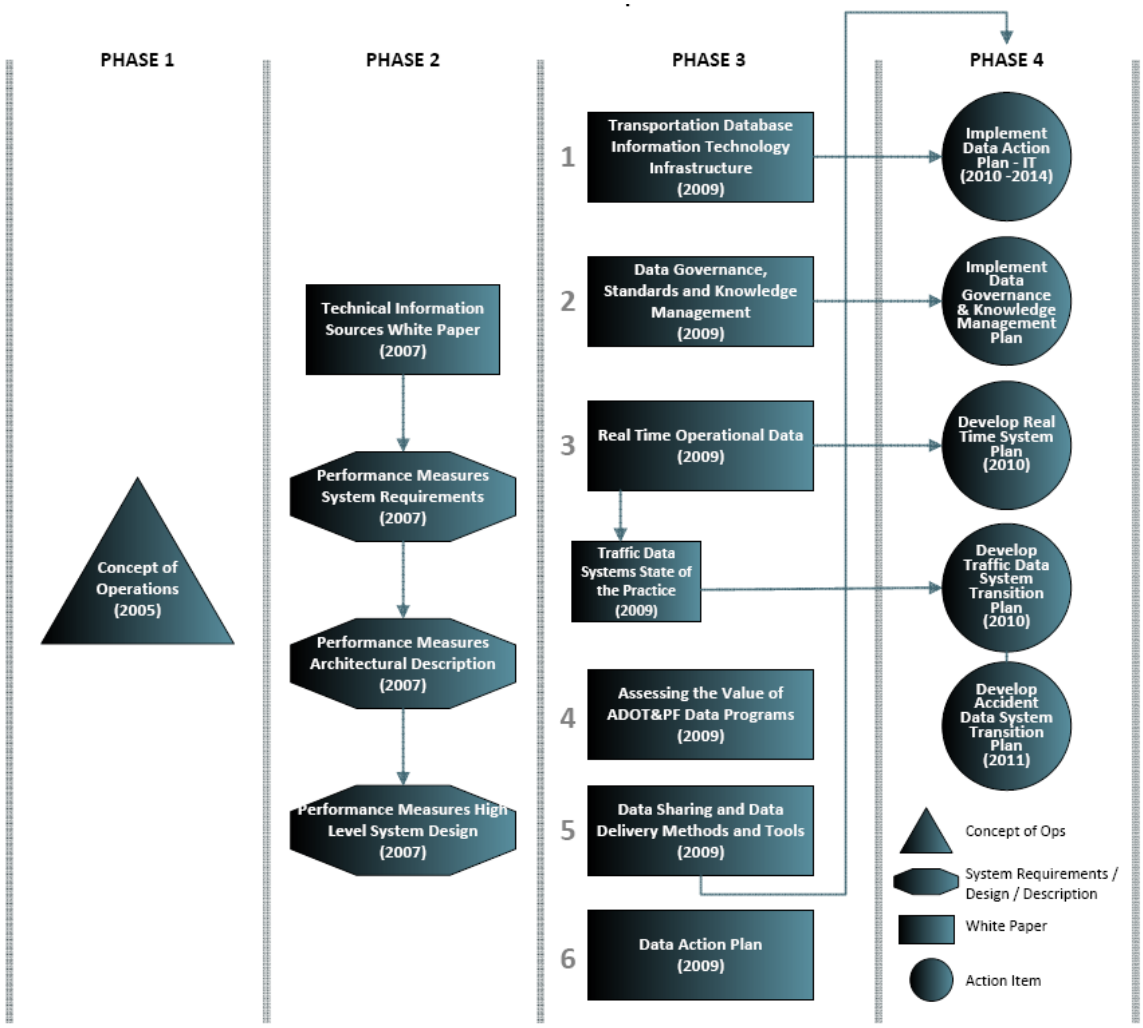
**Figure B-2. ADOT&PF's Mission and Core Business Programs** (Source: Alaska Data Business Plan: Data Governance, Standards, and Knowledge Management White Paper. Cambridge Systematics, Inc., September 2009)




The process to identify the relationship between data systems and business processes/decisions was part of the development of a Data Business Plan (DBP) over a multi-year, multi-phase effort, which began in 2005. Figure B-3 illustrates the phased approach for development of the DBP in which the

Department is now implementing practices in accordance with the Data Action Plan developed during Phase 3 of the project.

**Figure B-3. ADOT&PF's Phased DBP Development** (Source: Alaska Data Business Plan: Data Action Plan White Paper. Cambridge Systematics, Inc., September 2009)




Establishing the Data Action Plan allowed ADOT&PF to focus on the highest priority data and information needs first, as part of its data management program. The Data Action Plan includes recommendations for implementation of the following technology, processes, and standards:

1. Transportation Database Information Technology Infrastructure – implementation of an IT infrastructure that maximizes the advantages of an integrated Geodatabase and enterprise warehouse to support query, analysis, and reporting needs of the Division and the Department;
2. Data Governance, Standards, and Knowledge Management – establish a data governance framework, metadata standards and knowledge management system to strengthen data management at the Division;

3. Real-Time Operational Data and Traffic Data Systems – recommendations to meet the Section 1201 Real-Time System Management Information Program as part of SAFETEA-LU;
4. Assessing the Value of Data Programs – recommendations for assessing risks to the Department associated with the loss of critical data and information used for decision-making; and
5. Data Sharing Methods and Tools – recommendations for best practices in data sharing methods and tools that can be implemented with existing information systems, or planned as components in developing new systems.

***Applicability to U.S. DOT***

The approach used by ADOT&PF provides an example of a logical, well-designed approach for data management. This approach, as adapted for U.S. DOT, would include the following steps:

1. Define the Information Technology infrastructure improvements needed to support the data governance program;
2. Identify the methods and tools needed to share the roadway travel mobility data programs with stakeholders and the methods and tools needed to integrate data from other systems, with these programs; and
3. Define a risk management strategy or program to mitigate any potential risks to the U.S. DOT that may occur due to loss of data that supports the roadway travel mobility data programs.

Many of these steps have been completed through the development of this Data Business Plan. This includes developing the *Data Capture and Management: Needs and Gaps in the Operation and Coordination of U.S. DOT Data Capture and Management Programs* (Data Gaps) white paper, drafting a Data Governance Framework, identifying critical data and information systems as part of the Data Inventory, and identifying stakeholders (internal and external) for the working group and community of interest for roadway travel mobility data programs.

## Minnesota Department of Transportation (MnDOT)

The Minnesota Department of Transportation (MnDOT) has defined a clear vision and mission for managing data and information at the department, through the implementation of a Data Business Plan. MnDOT recognizes the importance of its data programs and considers them to be valuable assets that are needed to support strategic goals of the agency. These goals include investments in the improved safety, mobility, and infrastructure of MnDOT's transportation system. The MnDOT vision and mission statements are defined as follows:

***MnDOT Vision for managing data and information:*** *All MnDOT business decisions are supported by reliable data.*

***MnDOT Mission for managing data and information:*** *To provide reliable, timely data and information that is easily accessed, shared for analysis and integrated into MnDOT's decision-making process.*

MnDOT also identified seven key principles that will guide management of future MnDOT data and information systems investments:

1. Data will be managed as state assets;
2. Data quality will fit its purpose;
3. Data will be accessible and shared as permitted;
4. Data will include standard metadata;
5. Data definitions will be consistently used;
6. Data management is everybody's responsibility; and
7. Data shall not be duplicated.

MnDOT's data management program focuses on three key areas: (1) data governance, (2) Geographic Information Systems (GIS), and (3) data gaps and needs. MnDOT's leadership understands the value of a governance framework in establishing a structure for managing data programs used for business decision-making on a daily basis. After completing an assessment of the critical data and information needed to support MnDOT's business, the department identified the maintenance of GIS data to address needs for geospatial data, as a high priority. In support of this need, MnDOT designated a GIS Steering Team comprised of managers, who have decision making authority over GIS at MnDOT. This group is responsible for prioritizing investments in GIS technology, establishing an enterprise framework for GIS, identifying how GIS could and should fit into business processes, and advancing the state of the practice.

The use of metadata standards has also been implemented at MnDOT to ensure that the right data and information is used to respond to inquiries. MnDOT also uses an approach to Enterprise Architecture that addresses three major business challenges:

1. Managing the increasing *complexity* of information technology systems;
2. Delivering *business value* from those systems; and
3. Planning new IT solutions to meet business needs.

An assessment of the data gaps and needs regarding their data and information systems, helped MnDOT to identify asset management as one of their most critical needs, and therefore, the decision was made to move forward with procuring an asset management system.

Similar assessments of an agency's data programs should become part of an overall data management strategy. This assessment is typically related to risk assessment activities. The risk assessment also involves consideration of development of new systems to meet business needs and also whether to replace systems with newer applications, which can take advantage of new technologies and streamlined work-flow processes.

***Applicability to U.S. DOT***

The example from MnDOT provides excellent guidance for implementing a data management program at U.S. DOT. The basic guidance, tailored for U.S. DOT, includes the following steps:

1. Identify Vision and Mission for management of data and information for the roadway travel mobility data programs;
2. Define the principles that will guide the management of these programs at U.S. DOT;
3. Define the standards that need to be used for managing data and information;
4. Identify the gaps related to each of the roadway travel mobility data programs. Several gaps were identified in the Data Gaps white paper; and

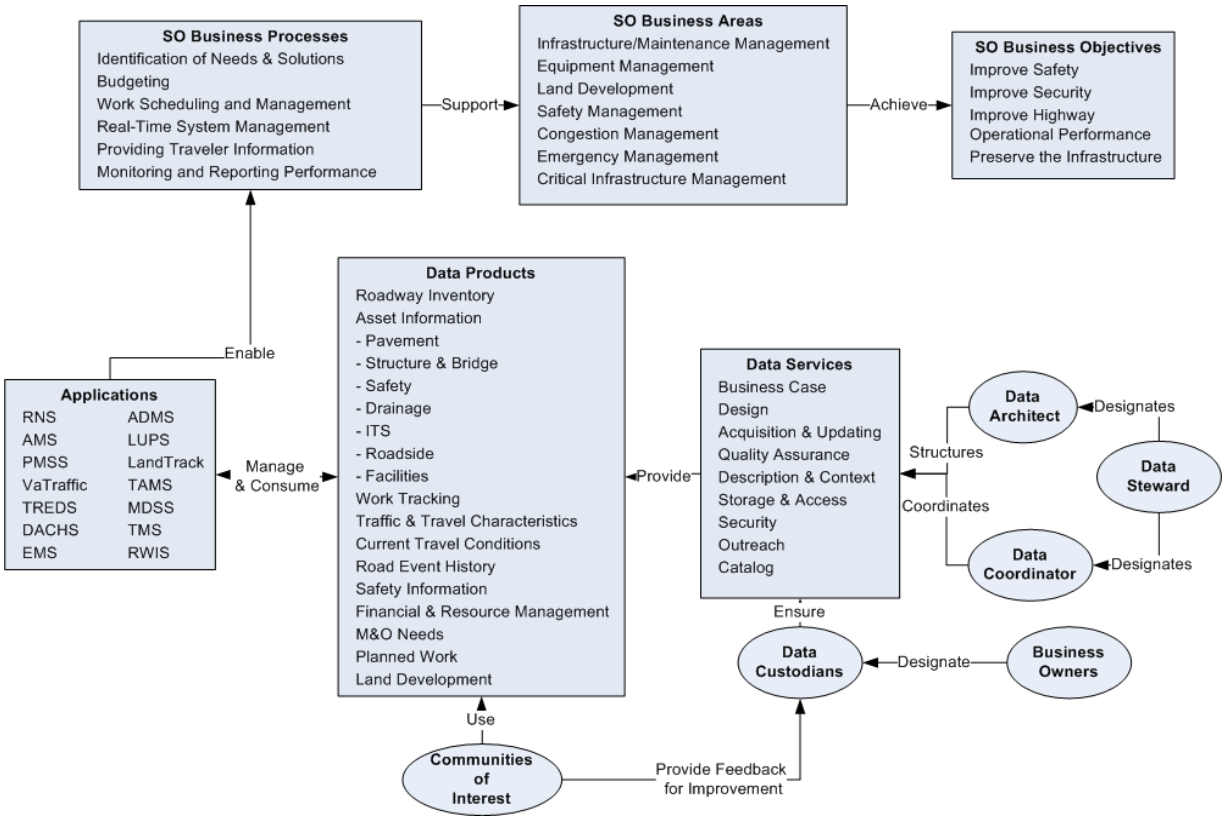
5. Implement a Data Business Plan to address the gaps and direct the management of data and information for the roadway travel mobility data programs at U.S. DOT.

Each of these steps is addressed in the proposed Data Management Plan for the U.S. DOT.

***Virginia Department of Transportation (VDOT)***

The Virginia Department of Transportation (VDOT), like other DOT examples presented here, has also implemented a Data Business Plan (DBP) for the System Operations (SO) Directorate to "provide a framework for making decisions about what data to acquire, how to get it, and how to make sure it is providing value commensurate with its cost" (VDOT SO Data Business Plan, June 2008). The DBP for the SO Directorate includes information about the VDOT Data Management Policy, and the roles and responsibilities of data business owners, data stewards, data custodians, data architects, and data coordinators, as well as defines roles of the Information Technology Division (ITD) in support of managing data and information. The data business plan framework for the SO Directorate is illustrated in Figure B-4.

**Figure B-4. VDOT Data Business Plan Framework** (Source: Virginia Department of Transportation System Operations Directorate Data Business Plan, Virginia Department of Transportation, June 26, 2008)




This framework identifies the data systems/products that are used to support core business functions of the division. In addition to this framework, VDOT also established a Business Governance Board in 2011, to support management of information and data at the department. This Board sets policies and

procedures for use of the intranet and makes all decisions regarding training, standard operating procedures, communications, requested functionality, and upgrades.

VDOT also created a Knowledge Management Office in 2003, which is dedicated to sharing and retaining critical organizational knowledge and supporting the management of transportation information throughout the department. The Knowledge Management Officer also serves as the chair of the Business Governance Board.

VDOT has also developed a seven-step Data Action Plan for implementing Data Governance at the department. The steps include the following:

1. Assign Roles – data business owners, data stewards, data custodians, etc.;
2. Produce Initial Data Catalog – includes information on data models, data dictionaries, data quality assurance processes, etc.;
3. Develop Data and Business Process Models – includes business use cases and diagrams;
4. Develop Business Requirements for Data Access – defines the needs for reporting and analysis;
5. Estimate Data Acquisition and Maintenance Costs – develops a standard methodology for the collection and maintenance of data;
6. Initiate Annual Business Data Review Process – includes two steps: data product review and data acquisition review; and
7. Establish Communication Protocols – establish clear lines of communication between Communities of Interest and data business owners, data stewards, data custodians, etc.

In support of creating networks to share and exchange knowledge and information, VDOT also uses what are referred to as Communities of Practice (COPs). Recommendations from the COPs often result in changes and improvements in business and workflow processes. As an example, the Construction Quality Managers COP helps to address the critical need for well-trained experienced inspectors and for quality control in the field.

The management of transportation information at VDOT relies on the collaborative efforts across several offices including the Office of Knowledge Management, the VDOT Research Library, the Administrative Services Office, the Communications Department (previously Public Affairs), and all other business offices at VDOT including the divisions and districts. Each office plays a vital role in providing data and information for assimilation and distribution of transportation information to support VDOT business operations and requests from external customers such as the governor's office, legislature, local governments, and the public.

***Applicability to U.S. DOT***

The experiences at VDOT illustrate how a knowledge management (KM) system helps to organize and categorize transportation data and information. The use of the KM system is very beneficial in capturing lessons learned from previous project experiences and also is used as a repository to store corporate knowledge about business process workflows and data and systems used to support such programs as the Highway Performance Monitoring System (HPMS). U.S. DOT should consider implementing a similar KM system to support roadway travel mobility data programs.

VDOT also maintains outreach to their Communities of Practice (which are similar to the recommended Communities of Interest for the roadway travel mobility data programs) in order to

ensure that stakeholders' needs continue to be met. It is recommended that meetings be conducted on a quarterly, semi-annual, or annual basis with the COIs for the roadway travel mobility data programs at U.S. DOT to ensure continued quality of data and information is provided to the stakeholders for those programs.

# B.3 Internal Best Practices

Additional internal best practices were documented in the *Needs and Gaps* white paper and are included here for their relevance to the efforts that will be undertaken in implementing the Data Business Plan at U.S. DOT.

## Data Quality

The following best practices for data quality were documented in the *Needs and Gaps* white paper:

- The Urban Congestion Reporting project found that data quality is best assured locally and immediately. Having data quality staff that are familiar with the network and minimizing the lag between data collection and data quality control were found to be critical elements of an effective data quality control process.
- HPMS has statistical specifications for data quality, as well as AASHTO's Guidelines for Traffic Data Programs. There are 20 data quality checks, and data that do not meet standards are rejected. Data quality performance measures need to be consistent across all U.S. DOT data programs. Development of data standards should be based on current ITS standards, gap analysis, and standards needed to support data capture/data sharing. The Federal role in standards should be similar to that of the U.S. Environmental Protection Agency (EPA) and vehicle emission standards.
- Federal Motor Carrier Safety Administration (FMCSA) requires robust data quality practices because their data is used to rate carriers based on crashes and inspections. As with any data, errors could occur during the data entry process (e.g., carrier ID entered incorrectly). DataQs is an online system that allows carriers to request a review of their data and challenge any errors they find. Carriers must submit evidence of why there is an error, and any challenges are automatically forwarded to the state agency that conducted the initial inspection. Carriers can track their request and final response through DataQs. The system helps those who are most impacted by the data to have input as to its accuracy.
- FMCSA also rates the completeness, timeliness, accuracy, and consistency of state-reported crash and roadside inspection data on a monthly basis. Results are displayed on the FMCSA Data Quality website. http://ai.fmcsa.dot.gov/DataQuality/dataquality.asp. FMCSA provides grant money for states to improve the quality of their data record systems, and they offer on-site support and training for state personnel responsible for reporting data.
- The FHWA Freight Performance Measures Initiative requires third party independent verification of data to ensure continuous data quality improvement. However, it is important to understand the legal arrangements for procurement of data from private parties, and to understand the distinctions between use versus ownership of the data. For example, there could be legal ramifications if purchased data is used for

purposes other than that contractually agreed. From a U.S. DOT perspective, it is better to take possession of the data product rather than the raw data. The contract language should broadly describe the main purpose of the program and the products desired from the data (e.g., data is to be used for congestion management and to better understand the impacts of freight operations).

***Outreach to Stakeholders and Customers***

Some of the best practice examples in outreach to stakeholders and customers (as identified in the Data Gaps white paper) include the following:

- FHWA has existing relationships with state partners through their individual field offices located in each state. These individual relationships make implementation of Federal programs actually work.
- FTA conducts stakeholder outreach through workshops, listening sessions, and group sessions to discuss various aspects of the connected vehicle program. They also actively attend conferences and give presentations to outreach to stakeholders.
- In terms of staying apprised of what data is available, the Bureau of Transportation Statistics' (BTS') primary success factor is establishing strong working relationships with the data owners. There are internal experts within BTS who are familiar with all the data sources available and have established contacts and knowledge of the various industries for which they collect data.
- BTS conducts outreach through workshops, listening sessions, meetings, conferences, professional working groups, Twitter, Facebook, and an online survey for customer satisfaction. They are also members of TRB committees and the North American Transportation Statistics Interchange. There is a regular cycle for responding to questions or correcting data. At conferences, BTS distributes their products and gathers feedback on how users are actually using the information BTS produces. BTS cited their stakeholder feedback process as another major success factor.
- The Integrated Corridor Management (ICM) Initiative is an example of joint management of an initiative by a core team of representatives from multiple U.S. DOT offices. The management team for the initiative is small and well structured, consisting of four people (one each from ITS JPO, FHWA Research, FHWA Operations, and FTA Research). Cited success factors include strong leadership within the program, mutual respect among team members (there are no "turf" issues), joint commitment to the goals and expected outcomes of the initiative, strong communication, and positive working relationships.
- Another example of a collaborative environment is FHWA's Office of Freight Management and Operations, which is working with the International Mobility and Trade Corridor Coalition, U.S. Customs, and the Niagara International Transportation Technology Coalition (NITTEC) to capture real-time data on border wait times at the Peace Bridge and Pacific Highway in the Pacific Northwest.

### *Applicability to U.S. DOT*

Each of the best practices presented in this section can serve as a menu of options for consideration by U.S. DOT in the implementation of Data Management Plan. The specific practices to be replicated will be determined after more in-depth discussions with internal U.S. DOT staff that support roadway travel mobility data programs and the stakeholders for each of these programs. Forums will be held in the spring of 2013 to address proposed steps for implementing data management at U.S. DOT.

# APPENDIX C. Best Practices in Data Governance

This section provides an overview of data governance and describes the terms used in a data governance program. Best practice examples are also presented from national and state experiences in developing and implementing data governance programs.

## C.1 Overview of Data Governance

To fully appreciate the value and importance of a data governance program, it is important to begin by defining several terms.

**Data governance** is defined as the execution and enforcement of authority over the management of data assets and the performance of data functions. The management of data assets is usually accomplished through a data governance board or council. This role is critical in successfully managing data programs that meet business needs and in supporting a comprehensive data business plan for the organization.

Support for a data governance program is performed under a data stewardship role, by data stewards.

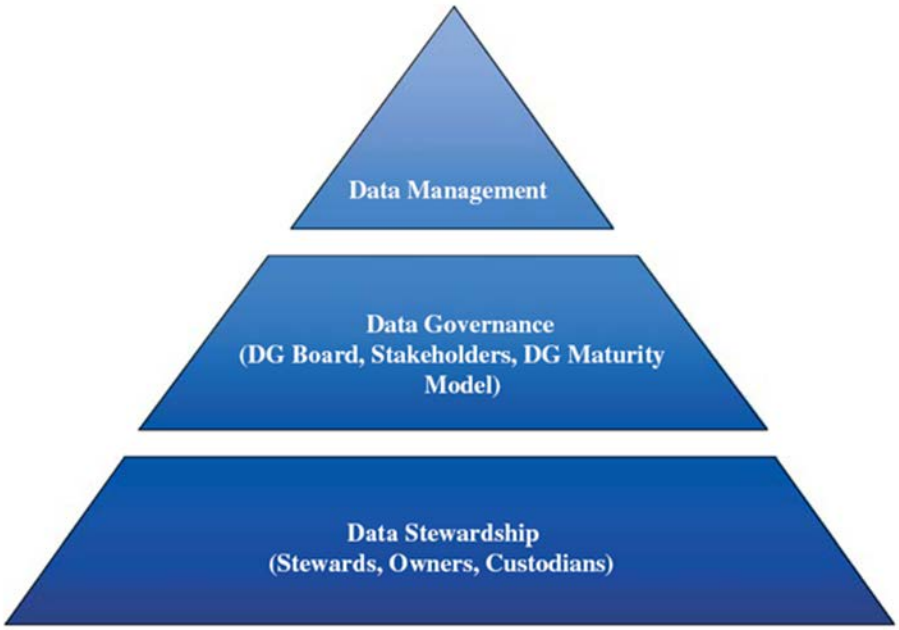
**Data stewardship** is defined as the formalization of accountability for the management of data resources. Data stewardship is a role performed by individuals within an organization known as data stewards.

The functions of data governance and data stewardship typically are part of an overall data management program within an organization.

**Data management** is defined as the development, execution, and oversight of architectures, policies, practices, and procedures to manage the information lifecycle needs of an enterprise in an effective manner as it pertains to data collection, storage, security, data inventory, analysis, quality control, reporting, and visualization.

There is a hierarchical relationship between data management, data governance, and data stewardship as illustrated in Figure C-1.

**Figure C-1. Data Management, Governance, Stewardship** (Source: Cambridge Systematics, Inc., January 2013)




***The U.S. DOT Roadway Transportation Data Business Plan includes all three of these elements.***

Data governance models are typically used to illustrate the data governance framework specific to an organization. There are a variety of data governance models which can be used to develop a data governance framework that is unique to an organization. Each model has its own advantages and disadvantages. There is not a one-size-fits-all approach when developing a data governance model and each agency should develop one that best suits the needs of the organization. Figure C-2 represents a generic data governance model illustrating how data governance fits within an organization.

**Figure C-2. Generic Data Governance Model** (Source: NCHRP Report 666: Target Setting Methods and Data Management to Support Performance-Based Resource Allocation by Transportation Agencies, Transportation Research Board, 2010)


The data users and stakeholders as well as the data stewards and custodians each serve a critical role in developing and maintaining data systems to meet user needs. To have a well functioning data governance structure, there must be open channels of communication between the providers of data and information and the users of the data. Outreach programs from the data providers to the data user community (stakeholders for the roadway travel mobility data programs) are an effective way to maintain this relationship. The benefits of implementing a data governance framework include improved quality of data collected and reported, which in turn has a positive impact on decision-making.

An example data governance model has been developed for the FHWA Office of Operations as illustrated in Appendix D.

## Data Governance Benefits

Having a well-defined and structured data governance program, with clearly defined roles and responsibilities for stewardship, offers significant benefits to an agency or office responsible for the management and oversight of data and information. The benefits usually include improvements in the processes and technology and tools used to manage data and information services. Several benefits from the use of data governance have been well documented by many organizations. Data governance:

- Provides a structured framework for utilizing technology, tools, processes, and persons/offices who manage data and information systems to support business operations;
- Ensures data programs support strategic vision, mission, goals;
- Promotes understanding of data as a valuable asset;
- Limits risks associated with loss of data and information;

- Encourages understanding and management of data from both a technical and business perspective;
- Provides enterprise-wide access to data standards and metadata;
- Provides central focus for identifying and controlling collection, storage, and disposition of data;
- Facilitates sharing of data with internal/external stakeholders, thereby reducing costs associated with data collection;
- Reduces redundancy in maintenance of duplicate data systems/increases efficiency in system maintenance;
- Ensures quality of data closest to source of data collection through the use of data standards;
- Provides quick resolution to issues regarding development, investment in data programs; and
- Provides opportunities to implement new/improved technologies including Business Intelligence tools for data programs.

## Data Governance Challenges

There are always challenges associated with implementation of any new initiative and this includes data governance. Some of the well-documented challenges based on experiences of other organizations are listed below.

- Implementing data governance requires a culture change to adapt the organization to the DG framework.
- Data governance requires a change in understanding of data and information systems as critical assets to the organization.
- External influences and/or policies/directives may prioritize other initiatives ahead of implementation of data governance.
- Staff does not understand their role in supporting data governance.
- There may be a lack of funding or available resources to support DG implementation.
- Lack of IT infrastructure of Business Intelligence (BI) tools such as knowledge management systems and GIS systems to support data governance and management of core business programs.

In addition, the National Association of State Chief Information Officers (NASCIO) has identified the following barriers that agencies need to overcome to implement data governance policies and procedures:[11]

- Determining rules and requirements;
- Gaining agreement regarding policies;
- Developing tools/software to implement data governance;

[11] Data Governance – Managing Information as an Enterprise Asset, April 2008, NASCIO.

- Cost of implementing policies;
- Incompatible systems; and
- Competing priorities within the organization.

Each of the challenges described above are handled in different ways in various agencies and organizations. The next sections describe several examples at national and state levels in successfully implementing components of a data governance framework.

# C.2 Best Practices in Data Governance

The next section describes some of the national and state agency best practices in the use of data governance to support data integration and sharing and exchange of data and information between internal and external stakeholders. Examples of success stories in data governances are provided to facilitate the next steps for U.S. DOT stakeholder offices interested in implementing data governance as part of the **Data Business Plan.**

## National Best Practices

Each of the examples provided in this section describe a particular aspect of data governance being used to improve the management of data and information within an organization. This includes sharing and exchange of data and information with internal and external stakeholders who have a business need for data and information from a particular office.

***U.S. Department of Health and Human Services (HHS) – Establishing IT Standards Committee***

HHS is currently in the process of developing a Nationwide Health Information Network. Two Federal Advisory Committees have been initiated to advise the HHS. The committees are the Health Information Technology Standards Committee to advise on federal health IT standards issues, and the Health Information Technology Policy Committee to establish a framework for governance of the Nationwide Health Information.

The nationwide health information network is a set of standards, services, and policies that enable secure health information exchange over the Internet. The network provides a foundation for the exchange of health information across diverse entities, within communities and across the country, helping to achieve the goals of the Health Information Technology for Economic and Clinical Health (HITECH) Act. This critical part of the national health IT agenda enables health information to follow the consumer, be available for clinical decision making, and support appropriate use of healthcare information beyond direct patient care to improve population health.

Capabilities currently supported by the Nationwide Health Information Network Exchange:

- Exchange of summary patient records for care coordination, including the Virtual Lifetime Electronic Record (VLER);
- Exchange of summary patient records for Social Security Administration (SSA) disability determination purposes; and
- Bio-surveillance and case reporting to CDC.

**Applicability to U.S. DOT**: This example demonstrates how the use of advisory committees can help to recommend or establish standards and policies for sharing of data and information across multiple stakeholder groups, thus providing necessary information for decision making.

***National Information Exchange Model (NIEM)***

The National Information Exchange Model (NIEM) is an "approach to driving standardized connections among and between governmental entities as well as with private sector and international partners which enable disparate systems to share, exchange, accept, and translate information. […] In Fiscal Year (FY) 2010, the Office of Management and Budget (OMB) provided guidance to all Federal Agencies to evaluate the adoption and use of NIEM as the basis for developing reference information exchanges to support specification and implementation of reusable cross-boundary services."[12]

The NIEM governance framework includes several entities that are similar to the recommended participants in the data governance framework for FHWA Office of Operations. These include a NIEM Executive Steering Council, NIEM Program Management Office, NIEM Communications and Outreach Committee (this would similar to the Communities of Interest), NIEM Technical Architecture Committee, and the NIEM Business Architecture Committee.

The following paragraphs describe how some federal agencies are addressing the challenge of implementing centralized governance, developing and implementing information exchange guidelines, creating collaborative sharing agreements and developing enterprise data management maturity, all of which are identified as challenges in the report of June 11, 2010. These agencies are committed to using the NIEM framework to facilitate the sharing and exchange of information across stakeholder groups (communities of interest). The following examples are excerpts from the report of June 11, 2010.

U.S. Department of Transportation (U.S. DOT)

The U.S. Department of Transportation is committed to using NIEM to support a department-wide capability to manage and share Suspicious Activity Reporting (SAR) information. The value expected is DOT's full participation in the Nationwide SAR Initiative (NSI), and ultimately to contribute in preventing another terrorist-type surprise attack on the nation. The information exchange at DOT is considered to be of high value. Currently, DOT creates SAR information, and stores this information in five different databases. Participation in the NSI is a priority of the National Security Staff and as such is seen as a high impact exchange.

**Applicability to U.S. DOT**: The use of NIEM or a similar framework supports the exchange of information across non-integrated databases. Integration of data at an enterprise level typically provides benefits to all offices within an organization by allowing access to data and information through a centralized portal.

Department of Agriculture (USDA)

USDA is committed to using NIEM and has initiated a project called the Acreage Crop Reporting Streamlining Initiative (ACRSI). The cross domain information exchange **allows for consistent reporting of common commodity data** for use by the Risk Management Agency (RMA), Farm

[12] Agency Information Exchange Functional Standards Evaluation, Federal CIO Council, June 11, 2010.

Services Agency (FSA), Natural Resources Conservation Service (NRCS), National Agricultural Statistics Service (NASS) and other USDA agencies. This cross domain exchange has significant potential for reuse across many federal agencies and private sector agriculture service providers.

**Applicability to U.S. DOT:** The use of a governance framework such as NIEM at the USDA provides for consistent reporting of data to internal and external stakeholder agencies.

Department of Labor (DOL)

The DOL reported that there is a potential opportunity for the use of NIEM for the Farm Labor Contractor Certification. DOL's Wage Hour Division (WHD) exchanges information with the Federal Bureau of Investigation (FBI) for the processing of FD-258 from data (Applicant Fingerprint Card) according to the detailed specification of an exchange (Agency Information Sharing Functional Standards – Template 3). This process is currently a heavily involved manual process and DOL believes that a highly automated exchange mechanism through an exchange standard like NIEM will result in a huge benefit for both DOL and FBI.

**Applicability to U.S. DOT:** The use of the NIEM or similar data and information sharing and exchange framework incorporates the use of business intelligence (BI) tools, which can be used to automate some current manual processes, thereby saving time and resources of an agency.

## State Best Practices

Several state agencies, including some state Departments of Transportation (DOTs), have also made significant progress in implementing standards, practices, tools, and technology infrastructure to strengthen data management and data governance programs at their respective agencies.

***Oregon Department of Transportation***

The Oregon Department of Transportation (ODOT) developed a charter to establish the Transportation Community of Interest Data Council in 2006. This Council "develops and provides policy, standards, processes, and procedures to ensure that data assets are well maintained and utilized. They provide guidance on data management activities, provide decisions on program improvements and resolve key issues for transportation assets."

ODOT has recognized the need for strong data governance to create and enforce data management standards. They have established a data management policy, which states:

The Oregon Data Governance model also includes well-defined roles and responsibilities for Data Stewards, Data Custodians, and the various Transportation Communities of Interest. The ODOT Data Governance structure also includes work groups, which were formed to provide work products used by the governance program including enterprise data management and reporting tools.

**Applicability to U.S. DOT**: The Oregon DOT data governance model provides an excellent template for defining the roles and responsibilities of all participants in a data governance framework, including the oversight council, communities of interest, data stewards, data business owners, etc. Many of these roles will also need to be defined for U.S. DOT stakeholder offices during implementation of their data governance program.

### *Washington State Department of Transportation (WSDOT)*

The Washington State Department of Transportation (WSDOT) has implemented a Data Council and a Data Stewardship Council to help support data governance at WSDOT. WSDOT further defines two categories of Data Stewardship: Business Stewardship and Technical Stewardship. The Business Stewards are executive, managerial, and operational stewards, while the Technical Stewards include the more traditional roles of system architects and database administrators. While these roles may be defined for larger application systems, WSDOT does acknowledge that for smaller applications, there may be just one or two people responsible for maintaining a specific data system, and there is no formal stewardship role defined for them.

The WSDOT model started with a smaller goal of gaining support at the executive level for a data policy. This eventually led to the support for the development of a data catalog, which was followed with the development of a Data Stewardship program. This proved to be an effective method for beginning implementation of a data governance framework.

The role of the Data Stewardship Council is to address data issues across the Department. An additional Data Council was also established to set standards for such components as data architecture, data modeling, and data stewardship.

The Department successfully encourages the use of data standards, by linking the funding of various programs, when possible, to the use of those standards within the organization. This is an effective means for gaining compliance with the data management policies and procedures for the Department. Defining distinct roles for Business and Technical Stewards also serves to strengthen this critical partnership and the overall data governance framework for the Department.

**Applicability to U.S. DOT:** The implementation of data governance at U.S. DOT should follow a similar approach of targeting a smaller group of participants to participate in the larger data governance initiative. Data governance participants should include those responsible for management of roadway travel mobility data programs.

WSDOT has been successful in linking funding for various programs to compliance with standards established within the organization. Although this may not be feasible for U.S. DOT, there may be an opportunity to encourage such compliance with standards for roadway travel mobility data programs, once the data governance framework is in place for individual stakeholder offices (or through OHPI's efforts to establish a Data Governance Advisory Council).

The WSDOT model also defines roles of business and technical stewards, each with a specific purpose in supporting data governance at WSDOT. A similar definition of roles is recommended for U.S. DOT data governance initiatives to ensure that business needs continue to be met for the roadway travel mobility data programs.

### *Minnesota Department of Transportation*

The Minnesota Department of Transportation (MnDOT) completed implementation of a Data Business Plan (DBP) in January 2011. This was the result of efforts that began in 2008 to develop a business plan for data. A Business Information Council (BIC) was formed to support and provide leadership for the development of the DBP. Council membership included senior managers and representatives from districts and specialty offices throughout the department. With adoption of the DBP in 2011, the BIC

was replaced with a seven-member Data Governance Board, which included the CIO and Data Management Coordinator.

The data business planning effort was identified as a flagship initiative in the department's Strategic Plan, and therefore, maintained support from executive leadership of the department.

The development and implementation of data governance at MnDOT included an assessment of existing data programs to determine where business needs were sufficiently being met and where improvements were needed (in data programs) to support other business needs. The tasks involved in developing the DBP and data management plan for MnDOT also helped the department to prioritize future investments in data programs to support strategic department business functions. The role of data governance at MnDOT is also clearly defined in the Data Business Plan.

**Applicability to U.S. DOT**: The development of a data business plan helps the organization to implement a data governance framework that is scaled to the size of the organization (identifying target offices as participants in the data governance initiative). A similar approach should be used to implement data governance to specifically target roadway travel mobility data programs at U.S. DOT. Similar experiences at other organizations have indicated that the successful implementation of data governance relies as much on scaling it appropriately to the organization, or particular offices, as it does on the actual governance framework designed for implementation.

***National Association of State Chief Information Officers (NASCIO)***

The National Association of State Chief Information Officers (NASCIO) recognized outstanding efforts of several state agencies in many categories related to management of information systems in its "2010 Best Practices." A few of these award winners are featured here for their experiences that are applicable to U.S. DOT stakeholder offices in implementing data governance programs in support of its Data Business Plan. The following information is excerpted from the *2010 Best Practices, NASCIO, Representing Chief Information Officers of the States, Showcasing Award Winning State Information Technology Programs (NASCIO, 2010)*

***Arkansas: Universal Financial Aid Management System***

The Arkansas Department of Higher Education was recognized in the category of Cross-Boundary Collaboration and Partnerships, which addresses identifying, planning, coordinating, sharing, integrating or joining up formerly non-integrated IT-related organizational goals and strategies. These could include governance and management, policies, business processes, data and information, systems and applications, services, technologies and infrastructure.

The Arkansas Department of Higher Education worked with Governor Mike Beebe and the General Assembly to improve the financial aid process for students looking to attend college for the first time or continue with their education. Millions of dollars in financial aid and scholarship funds went unused in Arkansas due to a cumbersome paper process and more than 21 different programs requiring individual processes and documentation. Students, parents, and counselors were not always aware of various programs or how best to gain access and apply. In a matter of minutes, students answer a few basic questions and the "YOUniversal" Financial Aid System determines aid for which they may be eligible across all 21 programs. Students may then submit one application, simplifying the process into one easy step.

The financial aid, scholarship, and loan awarding process is managed through the system that electronically notifies the student. Students manage their financial aid profile throughout their higher education experience within a secure student module.

Finally, the Department of Higher Education personnel manage the eligibility and award process for all aid programs within the administrative components of the system that replaces outdated disparate databases and systems previously used. Integration with third party sources is a key success factor and includes federal level FAFSA (family income) data, state level TRIAND high school transcript data, higher education institution enrollment verification data, and state financial integration for automated disbursement of funds. Data integration eliminates manual processing and errors, and expedites the review and award process.

**Applicability to U.S. DOT:** The integration of disparate databases and information systems helps to facilitate the exchange and sharing of data and information across various offices. The data governance framework should take into consideration the important role of U.S. DOT stakeholder offices in providing data and information from the roadway travel mobility data programs to other offices within and external to U.S. DOT and FHWA.

***Pennsylvania: Child Support Portal***

Pennsylvania's Department of Public Welfare's Child Support Enforcement Program was recognized in the category of Data Information and Knowledge Management**.** This category covers strategies, processes, applications, solutions, initiatives or programs that create, use, process, leverage, archive, or manage data, information, content, knowledge and intellectual value, property or capital.

Outcomes and benefits may include the provision of information-related content and services, as well as support for the development of policies, performance assessments, funding, decision-making and making government more transparent and/or inter-connected.

In Pennsylvania, roughly 1 in 10 people are involved in the Department of Public Welfare's Child Support Enforcement Program as a custodial parent, non-custodial parent, beneficiary, or employer. Customer service being a prime component of program administration, it was evident that customer "self- service" needed to be the primary approach. However, how can accurate, timely, and meaningful information be made available to child support customers cost-effectively while providing personalized service to each customer?

The solution was to provide citizens with a comprehensive "one stop" portal for 24x7 child support customer service and information delivery. This information must be consistent across communication channels such as integrated voice recognition, customer help desk, and web self service. By reducing the time spent answering common questions, case workers now focus their time on locating absent parents, establishing orders, and monitoring compliance. By improving data accuracy, they also spend less time updating case information.

Further, the burden on other social service programs has been reduced by having the non-custodial parent contribute to the child's financial well being and medical insurance. Employers are able to provide accurate information about non-custodial parents, enabling faster income attachment, which leads to increased collections. These improvements in service have helped Pennsylvania become one of the leading child support programs in the country by making it easy, cost effective, fast, and more accurate for citizens to interact with the program.

The realized return on investment includes:

- $1,462,268.40 saved due to customer demographic updates;
- Greater confidence in electronic service delivery mechanisms such as e-disbursement and e-collections which saved $50,000,000 over five years;
- Approximately $7,810,000 saved through reduced numbers of "no shows" to Appointments; and
- Approximately $430,000 saved as money on-hold is released to payees.

**Applicability to U.S. DOT**: Improving data accuracy and availability of data and information to all stakeholders reduces overall cost of maintaining data programs. In the case of Pennsylvania, this was the Child Support Portal; in the case of U.S. DOT, this applies to the roadway travel mobility data programs. The next section describes some of the current initiatives at U.S. DOT that support implementation of data governance at the agency.

# C.3 State of Data Governance at U.S. DOT

There are several ongoing data governance activities within U.S. DOT that are relevant to the Data Business Plan.

## FHWA Data Governance Advisory Council

The Office of Highway Policy Information has formed the FHWA Data Governance Advisory Council (DGAC) as part of the Data Integration Platform project. The DGAC has authority and responsibility to corporately advise on the utilization of FHWA's data resources and recommend major changes in FHWA data collection efforts that will result in increased consistency and coordination between existing and new data programs; the elimination of redundant data collection; the consolidation of data sources and resources; and compliance with external mandates. The DGAC will provide recommendations to the Investment Review Board (IRB) for approval and prioritization.

The DGAC will function as an Agency-level, senior leadership governance committee and will meet at least once each quarter.[13] Members of the council will include the following senior managers or their designees:

- Director of Highway Policy Information, who serves as the Chair;
- Representative from the Chief Financial Officer;
- Representative from the Chief Counsel Office;
- Representative from the Policy and Program Review, Office of Federal Lands Highway;
- Representative from Bridge Technology;
- Representative from the Transportation Performance Management;
- Representative from the Safety Programs;

[13] FHWA Data Governance Council Charter, December 1, 2012.

- Representative from Research, Development and Technology;
- Representative from the DA Council;
- Representative from the IT Advisory Group; and
- Representative from a Program Office that will rotate on an annual basis.

The next steps for the Data Governance Advisory Council will be to:

- Develop a proposed plan for corporately managing FHWA data;
- Review existing data collection efforts for need, consistency, and efficiency annually;
- Review and approve all new data collection efforts;
- Monitor and evaluate performance of data programs;
- Creation of a functional data dictionary;
- Recommend to the FHWA Investment Review Board ways to improve and streamline new and existing data collection efforts; and
- Coordinate FHWA data collection efforts with other modes within the Department.

This effort to develop the DBP will provide relevant examples for the Data Governance Advisory Council to use.

## FHWA Visualization Working Group

The FHWA Visualization Working Group provides leadership and advocates for the use of visualization and innovative data analysis techniques to accomplish the mission of FHWA.[14] The group is responsible for defining and initiating projects to promote the use of visualization techniques, such as applying visualization to current activities, development of resources, training, outreach, and representation at conferences. Future activities include developing a catalog of major datasets available and an internal catalog or data library that provides access to data. They have discussed the need for an open data policy and a place for data to reside so it can be accessible to both internal and external users. To achieve this, the Working Group meets every other month in order to coordinate efforts across individual FHWA offices.

Representatives are designated by respective offices of FHWA, including Headquarters offices, Research and Development, Federal Lands Highway, the Resource Center, and at least one Field Representative. The Associate Administrator for Research, Development, and Technology (RS&T) chairs the Working Group. RS&T also serves as Program Administrator and Secretariat and administers resources for the group.

## Connected Vehicle Senior Policy Working Group

The Connected Vehicle Senior Policy Working Group is researching and analyzing critical policy and institutional issues associated with Data Capture and Management (DCM) and Dynamic Mobility Application (DMA) Programs, including issues related to security, ownership, and liability of data. This

[14] FHWA Visualization Working Group Charter.

section provides a brief summary of program policy documents and relevant considerations for the Data Business Plan.

*Identification of Critical Policy Issues for the Data Capture and Management (DCM) and Dynamic Mobility Application (DMA) Programs.*[15] The document identifies the following critical policy issues related to governance of open data and open data environments that need to be addressed throughout the DCM Technical Program Phases:

- **Structure and Authority:** What form of governance structure(s) supports DCM data environments? Who will have what roles and responsibilities in decision making and dispute resolution? Decisions for upgrading and maintaining the data environments? Enforcement? What are the options regarding the level of ongoing federal involvement and, for each option, what are the roles and responsibilities of federal participants, and what are the associated costs? Can governance be implemented by the private sector or a hybrid of public- and private-sector stakeholders? Who currently has authority or is new authority needed?
- **Research Data Exchange (RDE) Data Manager:** What is the role of the RDE Data Manager, and are the appropriate policies defined to guide the Data Manager in operations, maintenance, and enforcement?
- **Federation of data environments:** What criteria should determine the appropriateness and eligibility of connecting external data environments with either the research or operational data environments? What are the associated costs and responsibilities of establishing and maintaining a relationship – for both the RDE/ODE and the external environment? What policies/rules are needed for adding or removing these external environments? For removing data sets? How does federation support data ownership? Revenue generation from the data ownership? How does federation impact liability or raise security risks? What are mitigation strategies?
- **Data Sharing Agreements:** What are standard components of data sharing agreement documents? What are important considerations and lessons learned from other agencies in implementing data sharing agreements?
- **Policy for Maintenance:** Who makes decisions about technology upgrades and flexibility of adding new technologies? Who manages the DCM system configuration?

*Policy Analysis and Recommendations for Development of the Dynamic Mobility Applications.*[16] This document presents a policy framework, analysis, and set of preliminary recommendations with regard to procurement and development choices associated with the application/application bundles. The document identifies the need for governance decisions regarding permissions and the implementation of user access control and other security technologies to manage access to the Open Source Applications Development Portal (OSADP). It also calls out Program and Portal governance groups

[15] Bartinique, I., D. Herring, A. Sheridan, S. Sloan, D. Waldron, and J. Hassol. Identification of Critical Policy Issues for the Data Capture and Management (DCM) and Dynamic Mobility Application (DMA) Programs. Draft Report, November 2011.

[16] Cuddy, M., A. Chachich, A. Brecher, M. Razo, and S. Sloan. Policy Analysis and Recommendations for Development of the Dynamic Mobility Applications (Draft Report). FHWA-JPO-12-033, US Department of Transportation, January 2012.

that would be responsible for selecting new DMA projects based on selection criteria that were used in selecting the first six DMA bundles.

*Metadata Guidelines for the Research Data Exchange.*[17] This document provides requirements for metadata that must be followed by registered users who submit data for the Research Data Exchange. In terms of standards, metadata must follow ASTM E2468-05, *Standard Practice for Metadata to Support Archived Data Management Systems*. Metadata content includes seven main sections: 1) Identification information; 2) Data quality information; 3) Spatial data organization information; 4) Spatial reference information; 5) Entity and attribute information; 6) Distribution; and 7) Metadata reference information. The document also defines the following roles and responsibilities in metadata management:

- **RDE Metadata Creator.** These individuals include RDE users (both registered users and registered users with additional access), the portal content manager, data management system content manager, and data sources. For projects under U.S. DOT procurement, the project team is required to create/update metadata along with the data set used for or generated from the project. For projects no longer under U.S. DOT procurement or for data from an external data management system, the data provider may volunteer to create/update metadata, or the RDE Portal Administrator may designate a Portal Content Manager to be the Metadata Creator.
- **RDE Metadata Manager.** The RDE Portal Administrator acts as the RDE Metadata Manager. When a new data set arrives, the Metadata Manager designates a Metadata Analyst from the RDE portal content managers to check the quality of the metadata. If the metadata conforms to the standard and passes the quality assurance test, the Metadata Manager releases the data set and metadata to the public through the RDE.
- **RDE Metadata Analyst.** The RDE portal content manager acts as the Metadata Analyst and is responsible for checking the quality of the metadata to see if it meets the metadata standard. If the metadata meets the standard, the Metadata Analyst performs the metadata quality assessment. The metadata is then passed back to the Metadata Manager.
- **RDE Metadata User.** These individuals include RDE registered users and registered users with additional access that use the metadata as the search criteria to find and use data sets. Metadata Users report any inappropriate or irrelevant metadata issues to the Metadata Manager.

## RDE/OSADP Policy Group

The recently created Research Data Exchange (RDE)/Open Source Applications Development Portal (OSADP) Policy Group is tasked to make specific recommendations regarding the RDE and OSADP efforts to the ITS Legal group, which has shared members with the Senior Policy Working Group.

[17] Jung, S., R. Glassco, K. Wunderlich. Metadata Guidelines for the Research Data Exchange. Final Report, December 30, 2011.

# APPENDIX D. Example Data Governance Framework

The following steps are required for any U.S. DOT stakeholder office interested in implementing data governance:

1. Map data programs to business objectives;
2. Define stakeholder roles and responsibilities;
3. Develop data governance model;
4. Develop data governance charter;
5. Develop data catalog; and
6. Assess data governance maturity.

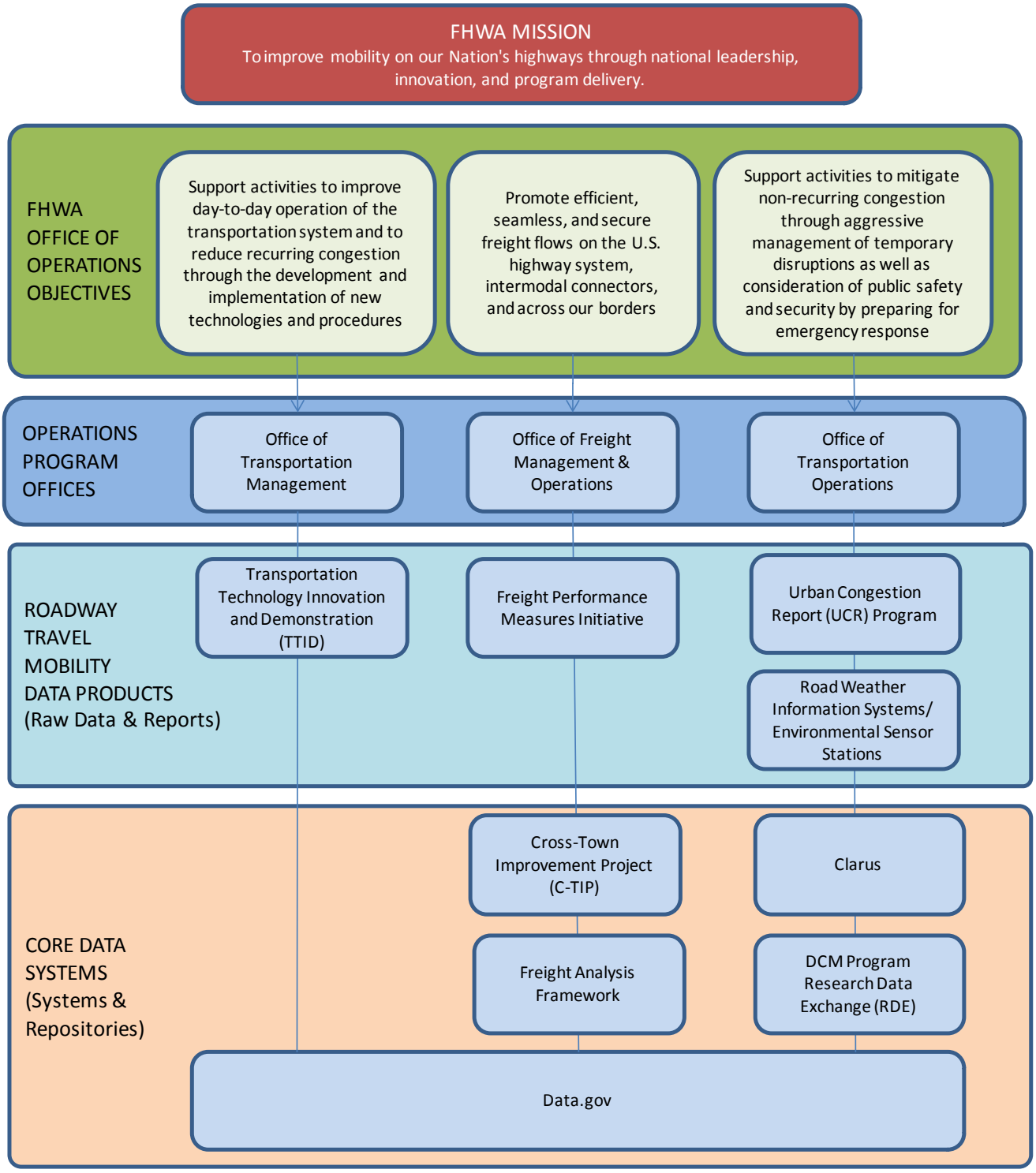
This appendix describes these steps in the context of an example Data Governance Framework for travel data within the FHWA Office of Operations.

## D.1 Map Data Programs to Business Objectives

The initial component of a Data Governance Framework is to define the relationship between the mission and business objectives of a U.S. DOT stakeholder office and how they map to the data programs managed by that office.

As an example, Figure D-1 illustrates the relationship between FHWA's mission, FHWA Office of Operations' business objectives, program offices, and core data systems that support roadway travel mobility data. Establishing and documenting this link between data and information as they relate to supporting business needs is a critical first step in implementing a Data Governance Framework.

**Figure D-1. Example Mapping of Mission, Business Processes, and Relationships to Core Data Systems** (Source: Cambridge Systematics, Inc., January 2013)



## D.2 Define Stakeholder Roles and Responsibilities

Using the hierarchical relationship between data management, data governance, and data stewardship illustrated in Appendix C, U.S. DOT stakeholder offices should define roles and responsibilities for data governance.

The following are example data governance roles for the FHWA Office of Operations:

- Data Governance Team – The designated individuals within FHWA Office of Operations responsible for the oversight of data programs to support the business functions of the office.
- FHWA Data Governance Advisory Council, Office of Operations Team Leader – Representative from the FHWA Office of Operations who will participate on the FHWA Data Governance Advisory Council.
- Data Business Owners – Individuals who manage the data and metadata for information systems within their area of responsibility. Data business owners are responsible for maintaining the data dictionaries for the data systems and for establishing business requirements for the use of roadway travel mobility data.
- Data Stewards – Individuals who ensure data is managed according to policies established by FHWA Office of Operations Data Governance Team.
- Community of Interest, Internal – Any persons or offices internal to U.S. DOT that collects, owns, maintains, uses or interfaces with, accesses, benefits from, or is otherwise affected by roadway travel mobility data.
- Community of Interest, External – Any persons or offices external to U.S. DOT that collects, owns, maintains, uses or interfaces with, accesses, benefits from, or is otherwise affected by roadway travel mobility data.

Table D-1 provides an example of stakeholder membership and responsibilities within the FHWA Office of Operations for each data governance role.

**Table D-1. Example Stakeholder Roles and Responsibilities**

| Data Governance Role | Membership | Responsibility |
|---|---|---|
| FHWA Office of Operations Data Governance Team | One representative from each of the FHWA Office of Operations offices – HOFM, HOTM, HOTO | Finalize data Governance Framework and plan with input from Communities of interest<br>Identify Data Stewards, Data Business Owners for roadway travel mobility data programs |
| Data Business Owners for Roadway Travel Mobility Data | TBD | Establish business requirements for use of roadway travel mobility data. |
| Data Stewards for Roadway Travel Mobility Data | TBD | Ensure data for each program is managed according to policies established by FHWA Office of Operations Data Governance Team |

| Data Governance Role | Membership | Responsibility |
|---|---|---|
| Internal Community of Interest | ITS Joint Program Office (JPO), RITA Bureau of Transportation Statistics (BTS) , FHWA Office of Human Environment – Bicycle & Pedestrian Program (HEPH), FHWA Office of Operations, FHWA Office of Transportation Management (HOTM), Turner-Fairbank Highway Research Center (TFHRC), FHWA Office of Freight Management & Operations (HOFM), FHWA Office of Performance Management, FHWA Office of Highway Policy Information (HPPI), FHWA Office of Performance Management, Volpe (Federal Transit Administration (FTA ) representative), Federal Motor Carrier Safety Association (FMCSA), FHWA Research and Development, Noblis (Research Data Exchange) | |
| Roadway Travel Mobility Data Community of Interest – External | External stakeholders | TBD |

## D.3 Develop Data Governance Model

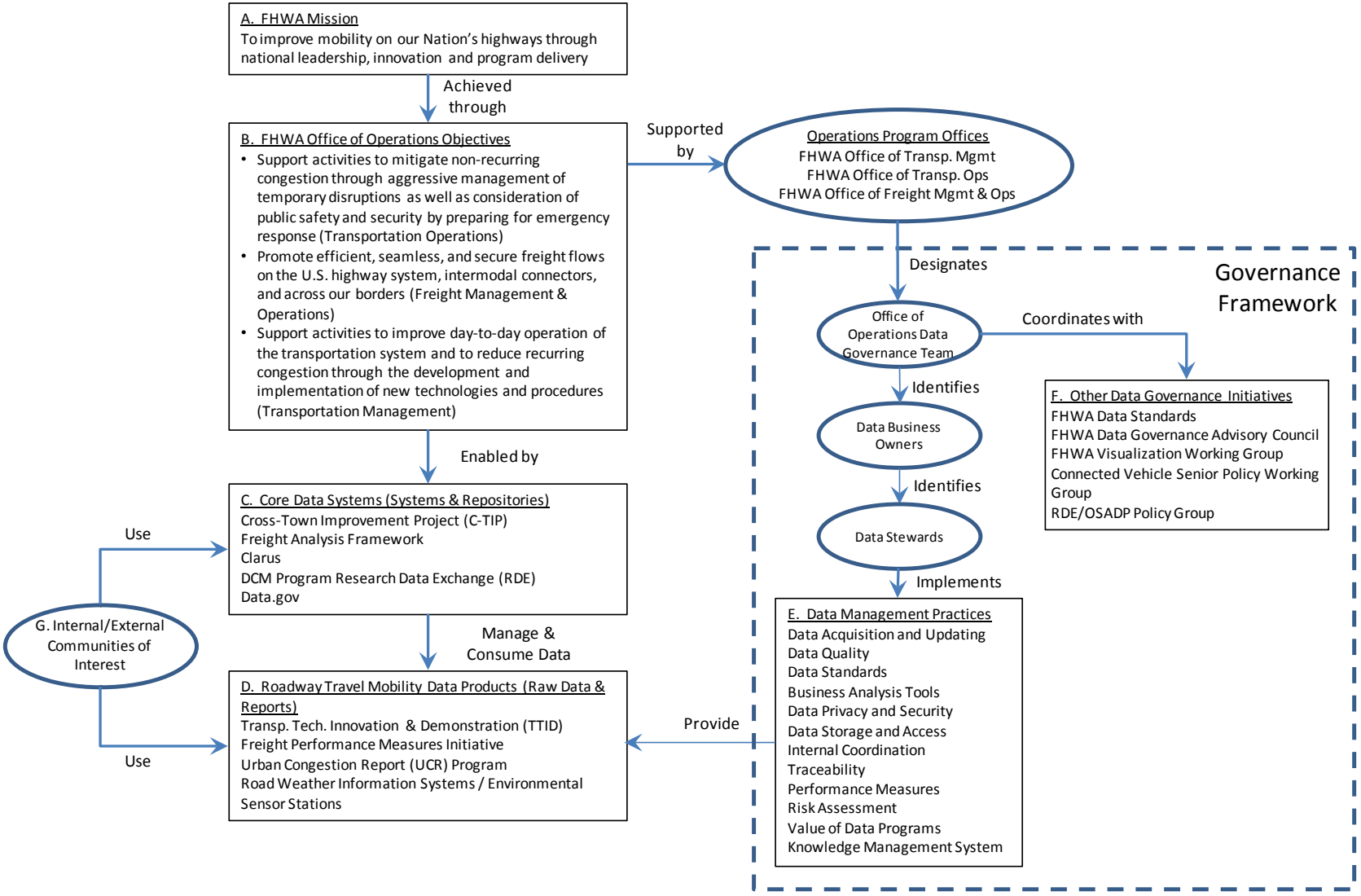
Once the mapping of data programs to agency and office mission and goals is accomplished, a data governance model diagram should be established to formalize the structure for managing the data programs.

Figure D-2 shows an example data governance model diagram for the FHWA Office of Operations. This model includes Data Business Owners and Data Stewards who will help implement data services to maintain the core data products and systems and meet the business objectives of the FHWA Office of Operations. The stakeholders are shown in oval shapes and the objectives, systems, products and initiatives are shown in rectangles.

The following describes the components of the example data governance model diagram:

- **A.** FHWA Mission;
- **B.** FHWA Office of Operations Objectives;
- **C.** Core Data Systems (Systems and Repositories);
- **D.** Roadway Travel Mobility Data Products (Raw Data and Reports);
- **E.** Data Services; and
- **F.** Communities of Interest.

**Figure D-2. Example Data Governance Model Diagram** (Source: Cambridge Systematics, Inc., January 2013)




U.S. Department of Transportation, Research and Innovative Technology Administration
Intelligent Transportation System Joint Program Office

# D.4 Develop Data Governance Charter

A data governance charter sets forth the purpose, mission, vision, goals and objectives, and data management policies for data governance within a U.S. DOT stakeholder office.

Below is an example data governance charter for the FHWA Office of Operations, which includes the Office of Transportation Management, Office of Freight Management & Operations, and the Office of Transportation Operations.

## Example Data Governance Charter

***Purpose***

This data governance charter is intended to provide guidance to the FHWA Office of Operations in the coordination of data management activities for the mobility data and connected vehicle data capture programs.

***Vision***

Business decisions of the FHWA Office of Operations (and the offices for which it provides guidance) are supported by well managed data programs, which comply with established DOT and Office of Operations policies, procedures, and standards for the collection (or capture), maintenance, and use of data, in particular, roadway travel mobility data.

***Mission***

To provide timely, accurate, high quality data and information pertaining to roadway travel mobility data programs managed by the FHWA Office of Operations that is easily shared with other stakeholder offices, both internal and external to the DOT.

***Goals and Objectives***

The goals and objectives established for the Data Business Plan can be used to guide data governance within the Office of Operations:

- GOAL 1. Improve the coordination and communication mechanism across U.S. DOT and FHWA offices involved with roadway travel mobility data to avoid investing resources in the same or similar types of data related programs.
  - Objective 1.1. Identify gaps and redundancies in roadway travel mobility data programs.
  - Objective 1.2. Devise ‘‘rules of engagement’’ regarding collaboration of the data functions for roadway travel mobility data.
- GOAL 2. Improve the coordination of the data capture activities associated with ITS JPO-sponsored research in wirelessly connected vehicle technologies with roadway travel mobility data programs within U.S. DOT and FHWA.
  - Objective 2.1. Identify how current and planned data from the connected vehicle initiative can support existing roadway travel mobility data programs.

- Objective 2.2. Identify how data from roadway travel mobility data programs within U.S. DOT and FHWA can support the connected vehicle initiative.
- Objective 2.3. Identify existing/future data inventory and data structures/policies/governance practices that could be applicable to the Research Data Exchange.

***Data Management Policy***

The following data management policy is stated in the draft FHWA Data Governance Advisory Council memorandum:

> *FHWA information is an enterprise resource that must be managed from an enterprise perspective. Put another way, the FHWA will treat its data as agency-wide assets and not as items defined by, and embedded in, individual programs. The Resource Center, Divisions, and Program Offices, as stewards of agency data, are responsible for the timeliness, accuracy, understandability, availability, and security of the data under their stewardship.*
>
> *The FHWA will maintain an Enterprise Data Management Program. The program invokes the principles of information stewardship as well as data architecture, data registration, data standardization, data certification, and data life-cycle management.*
>
> *Data management, especially data standardization, enables many information technology cost control activities, such as consolidating applications and servers.*
>
> *The Enterprise Data Management Program and its products will provide a framework to guide, support and promote integrating data management principles systematically into the FHWA's new and evolving information systems and processes.*

In support of the FHWA Data Management Policy, the following data management policy shall be implemented in January 2013 to coincide with the implementation of the data governance framework for the FHWA Office of Operations:

> *The FHWA Office of Operations shall manage data from roadway travel mobility data programs as valuable assets. Data stewards shall be identified for each of these programs, and these stewards will have responsibility for ensuring the quality and accuracy of the data for these programs. All roles and responsibilities pertaining to the management of these data programs shall be documented in a Data Governance Manual for the Office of Operations.*
>
> *Data management of roadway travel mobility data programs shall be done in accordance with standards established by the Office of Operations Data Governance Team. Standards will include Metadata Standards to be followed in the management of these data programs.*

## D.5 Develop Data Catalog

Another integral component of a Data Governance Framework is a data catalog. The purpose of a data catalog is to provide a centralized location for information about the data used by stakeholders involved with roadway travel mobility data programs. Involvement means the office is performing one or several of the following functions related to data: collection, analysis, reporting, dissemination, or providing guidance to other stakeholders related to those functions.

The catalog should be reviewed and revised by data business owners within each U.S. DOT stakeholder office to ensure that all data systems, data standards, roles, and responsibilities, etc., are correctly identified. It should also be revised at least on an annual basis, or monthly if changes occur that require updating the information listed in the catalog.

An example data catalog for the FHWA Office of Operations is provided in Table D-2. It builds on the *Data Inventory* and includes the following components:

- List of data programs within U.S. DOT.
- List of data business owners within that program, with their contact information. Data business owners are defined as the U.S. DOT office who manage the data and metadata for information systems within their area of responsibility for a business unit (roadway travel mobility data), maintain the data dictionaries for the data systems within their office, and establish business requirements for the use of roadway travel mobility data.
- List of data stewards responsible for the data program, with their contact information. Data stewards ensure data is managed according to U.S. DOT policies.
- Instructions for accessing data standards and definitions associated with the collection and use of the data.

**Table D-2. Example Data Catalog**

| Agency | Data Program | Business Owner | Data Steward | Access Instructions |
|---|---|---|---|---|
| Federal Highway Administration (FHWA) | Highway Performance Monitoring System (HPMS) | Office of Highway Policy Information http://www.fhwa.dot.gov/ policyinformation/ | Ralph Gillman, HPMS Program Team Leader, 202-366-5042 | Access to HPMS Software Version 7.0 restricted to authorized users. A UPACS account must be established and permission granted through the FHWA Field Division in each State. Some HPMS data for States and MPOs are publicly accessible through the HPMS website. |
| | National Transportation Atlas Databases | RITA Bureau of Transportation Statistics http://www.its.dot.gov/ | Dave Smallen, 202-366-5568; Mark Bradford, 202-366-6810 | NTAD databases are accessible to the public and are available for download through the BTS website.[18] Users can also request an NTAD DVD through the BTS bookstore. |
| | Archived Data User Service (ADUS) | Office of Highway Policy Information http://www.fhwa.dot.gov/ policyinformation/ | Tianjia Tang, 202-366-0180; James Pol, ITS Joint Program Office, 202-366-4374, james.pol@fhwa.dot.gov | Documents for ADUS program activities, standards and resources are available to the public and can be accessed through the FHWA ADUS website.[19] |
| | Transportation Technology Innovation and Demonstration (TTID) | Office of Operations http://ops.fhwa.dot.gov/ | Bob Rupert, 202-366-2194, Robert.Rupert@dot.gov; Jimmy Chu, 202-366-3379, Jimmy.Chu@dot.gov | Access to the data is provided through the ITIP Stakeholder Applications Data Warehouse,[20] which is available to FHWA and the participating State agencies (currently 27). Each FHWA Program and participating agency is provided with a User ID and password for the data warehouse. The only restriction on use is that users are not allowed to sell the data. |
| | Urban Congestion Report (UCR) program | Office of Operations http://ops.fhwa.dot.gov/ | Rich Taylor, FHWA Office of Transportation Management, 202-366, 1327, rich.taylor@dot.gov | Urban Congestion Reports are accessible by the public and are available for download through the Urban Congestion Report website.[21] |

[18] RITA BTS National Transportation Atlas Database website, http://www.bts.gov/ publications/national_transportation_atlas_database/.
[19] FHWA Office of Highway Policy Information Archived Data User Service website, http://www.fhwa.dot.gov/policy/ohpi/travel/adus.cfm.
[20] ITIP Stakeholder Applications Data Warehouse website, http://stakeholder.traffic.com.
[21] FHWA Urban Congestion Report website, http://www.ops.fhwa.dot.gov/ perf_measurement/ucr/.

| Agency | Data Program | Business Owner | Data Steward | Access Instructions |
|---|---|---|---|---|
| Federal Highway Administration (FHWA) (continued) | National Speed Data Collection and Reporting Program | Office of Highway Policy Information http://www.fhwa.dot.gov/policyinformation/ | Tianjia Tang, 202-366-2236, tianjia.tang@dot.gov | TBD |
| | Traffic Volume Trend (TVT) Reports | Office of Highway Policy Information http://www.fhwa.dot.gov/policyinformation/ | Daniel Jenkins, Office of Highway Policy Information, 202-366-1067, daniel.jenkins@dot.gov | TVT reports are available online[22] to all users. |
| | Travel Monitoring and Analysis System (TMAS) | Office of Highway Policy Information http://www.fhwa.dot.gov/policyinformation/ | Daniel Jenkins, Office of Highway Policy Information, 202-366-1067, daniel.jenkins@dot.gov | TMAS is accessible via the FHWA's User Profile and Access Control System (UPACS) website.[23] Access is limited to State DOT, FHWA Division office personnel, and others with a legitimate reason for accessing the system. User authorization must be obtained through the UPACS Administrator within the FHWA Division office in the individual State. Contact information for all the Division offices is available on the FHWA website.[24] |
| | Vehicle Travel Information System (VTRIS) | Office of Highway Policy Information http://www.fhwa.dot.gov/policyinformation/ | Ralph Gillman, (202) 366-5042; David Jones, (202) 366-5053 | Report data that are based on VTRIS summary data are accessible to the public and are available through the VTRIS W-Table website.[25] VTRIS Reports are available to the public between the following hours: Monday-Friday: 7 a.m. To 10 p.m. (EST), Sunday: 12 Noon To 10 p.m. (EST). A User's Guide consisting of the manual, appendix, and database structure supplementary tables is provided on the FHWA Office of Highway Policy Information website.[26] |

[22] FHWA Office of Highway Policy Information, Traffic Volume Trends website, http://www.fhwa.dot.gov/policyinformation/travel_monitoring/tvt.cfm.
[23] FHWA Information Systems – UPACS Login website, https://fhwaapps.fhwa.dot.gov.
[24] FHWA Field Offices website, http://www.fhwa.dot.gov/field.html#fieldsites.
[25] VTRIS W Table Generation website, https://fhwaapps.fhwa.dot.gov/vtris-wp/default.aspx.
[26] FHWA Office of Highway Policy Information website, VTRIS Users Guide page, http://www.fhwa.dot.gov/ohim/ohimvtis.cfm.

| Agency | Data Program | Business Owner | Data Steward | Access Instructions |
|---|---|---|---|---|
| Federal Highway Administration (FHWA) (continued) | National Transportation Statistics | RITA Bureau of Transportation Statistics http://www.its.dot.gov/ | BTS Information Services: 800-853-1351 or RITAInfo@dot.gov | NTS is accessible to the public and the datasets are available for download on the BTS website.[27] |
| | Road Weather Information Systems (Environmental Sensor Stations) | Office of Operations – Road Weather Management http://ops.fhwa.dot.gov/weather/index.asp | Paul Pisano, Road Weather Management Program, FHWA, (202) 366-1301, paul.pisano@dot.gov | Access to raw sensor data is generally limited to internal State DOT (or other owner agency) staff. States participating in the *Clarus* initiative make sensor data publicly available through the *Clarus* System website.[28] |
| | *Clarus* | Office of Operations – Road Weather Management http://ops.fhwa.dot.gov/weather/index.asp | Paul Pisano, Road Weather Management Program, FHWA, (202) 366-1301, paul.pisano@dot.gov; Ben McKeever, (202) 366-4876, ben.mckeever@dot.gov | *Clarus* data is accessible to the public and is available on the *Clarus* System website.[29] Subscriptions reports that are fulfilled periodically (such as every 5, 10, 15, 20, 25, or 30 minutes) may require a password. |
| | Border Crossing Data | RITA Bureau of Transportation Statistics http://www.its.dot.gov/ | BTS Information Services: 800-853-1351 or RITAInfo@dot.gov | BTS provides access to the data through an interactive searchable interface called North American Border Crossing/Entry Data Online Data Tool.[30] The interface is accessible to the public. |
| | North American Transborder Freight Data | RITA Bureau of Transportation Statistics http://www.its.dot.gov/ | BTS Information Services: 800-853-1351 or RITAInfo@dot.gov | BTS provides access to the data through an interactive searchable interface called North American Transborder Freight Data.[31] This allows users to create multi-variable cross-tabulations on port, geography and commodity for all modes of transportation. Search results can be viewed online and then downloaded. This data is available online to all users on the website. |

[27] BTS National Transportation Statistics website, http://www.bts.gov/publications/ national_transportation_statistics/.

[28] At the time of publication, Clarus is being transitioned to the Meteorological Assimilation Data Ingest System (MADIS). Users can request access to Clarus data through the MADIS Data Application website, http://madis.noaa.gov/data_application.html.

[29] *Ibid.*

[30] North American Border Crossing/Entry Data Online Data Tool, http://transborder.bts.gov/programs/international/transborder/TBDR_BC/TBDR_BC_Index.htmlhttp://www.bts.gov/programs/international/transborder/TBDR_BC/ TBDR_BC_Index.html.

[31] North American Transborder Freight Data Interface website, http://transborder.bts.gov/programs/international/transborder/http://www.bts.gov/programs/international/transborder/TBDR_QA.html.

| Agency | Data Program | Business Owner | Data Steward | Access Instructions |
|---|---|---|---|---|
| Federal Highway Administration (FHWA) (continued) | Commodity Flow Survey | RITA Bureau of Transportation Statistics http://www.its.dot.gov/ | Commodity Flow Survey team: cfs@dot.gov | CFS reports and tables are accessible to the public and are available for download on the BTS website.[32] |
| | Intermodal Passenger Connectivity Database | RITA Bureau of Transportation Statistics http://www.its.dot.gov/ | TranStats Customer Support, Phone: (800)853-1351, Email: ritainfo@dot.gov | The database is accessible to the public and is available for download on the TranStats website.[33] |
| | Cross-Town Improvement Project (C-TIP) | Office of Freight Management and Operations http://www.ops.fhwa.dot.gov/freight/ | Randy Butler, Office of Freight Management – FHWA, 202-366-9215, Randy.butler@dot.gov; Ron Schaefer, SAIC, 618-257-8435, schaeferrl@saic.com | TBD |
| | FHWA Freight Performance Measures Initiative | Office of Freight Management & Operations http://www.ops.fhwa.dot.gov/freight/ | Ed Strocko: 202-366-2997, Ed.Strocko@dot.gov | Data from the Freight Performance Measures initiative is available on the FPM web application.[34] Access is limited, and a user must have an account to log into FPMweb. The website currently operates under a ‘‘limited user’’ phase. |
| | Freight Analysis Framework (FAF) | Office of Freight Management & Operations http://www.ops.fhwa.dot.gov/freight/ | Michael Sprung: 202-366-9047, Michael.Sprung@dot.gov | FAF data products can be downloaded at the FAF3 website including: Data Extraction Tool,[35] National Freight Transportation Maps[36], and Highway Network Data Files.[37] Users can also download the entire FAF3 2007 and 2040 flows database in either Microsoft Access 2003 or CSV format.[38] |

[32] Commodity Flow Survey Website, http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/commodity_flow_survey/index.html; Commodity Flow Survey Website, http://www.bts.gov/publications/ commodity_flow_survey/.

[33] BTS TranStats Website – Intermodal Passenger Connectivity Database, http://www.transtats.bts.gov/DL_SelectFields.asp?Table_ID=1180&DB_Short_Name=Transnet.

[34] Freight Performance Measures Web Application (FPMweb) website, https://www.freightperformance.org/fpmweb/user_login.aspx.

[35] FAF3 Data Extraction Tool website, http://faf.ornl.gov/fafweb/Extraction0.aspx.

[36] FAF3 National Freight Transportation Maps, http://faf.ornl.gov/fafweb/dataMap.aspx.

[37] FAF3 Highway Network Data Files, http://faf.ornl.gov/fafweb/networkdata.aspx.

[38] FAF3 2007 and 2040 flows database website, http://ops.fhwa.dot.gov/freight/freight_analysis/faf/index.htm.

| Agency | Data Program | Business Owner | Data Steward | Access Instructions |
|---|---|---|---|---|
| Federal Highway Administration (FHWA) (continued) | International Freight Data System (IFDS) | RITA Bureau of Transportation Statistics http://www.its.dot.gov/ | Long Nguyen, RITA BTS, long.nguyen@dot.gov | Access (review and use the information) to the IFDS will be limited to the statutory authority of each DOT OA. Information will be automatically and securely downloaded from ACE/ITDS to the IFDS database managed by RITA. The IFDS database will then make the information available to participating DOT OAs. Only those DOT OAs that have statutory authority to view the data in order to execute their international missions will be given access to the IFDS. |
| | National Household Travel Survey (NHTS) | RITA Bureau of Transportation Statistics http://www.its.dot.gov/ | Adella Santos, NHTS Program Manager, 202-366-5021, adella.santos@dot.gov | All of the data for 1969, 1977, 1983, 1990, 1995, 2001 and 2008 are accessible to the public and are available for download on the NHTS website.[39] The website serves as a central facility to obtain information on the NHTS, status updates on the dataset or publications, publications using NHTS data, and the disaggregate NHTS data which is available for download and analysis by the user community. |
| | Integrated Corridor Management (ICM) Program | RITA Bureau of Transportation Statistics http://www.its.dot.gov/ | Brian Cronin, (202) 366-8841, brian.cronin@dot.gov; Steven Mortensen, (202) 493-0459, steven.mortensen@dot.gov; Robert Sheehan, (202) 366-6817, robert.sheehan@dot.gov; Dale Thompson, 202-493-3420, dale.thompson@dot.gov | Currently, accessibility to corridor data is limited to ICM team members. |
| | TFHRC Data Resources Testbed | Turner-Fairbank Highway Research Center http://www.fhwa.dot.gov/research/ | Carol Tran: 202-493-3315, carol.tan@dot.gov | Open to both internal and external researchers. |

[39] NHTS website, http://nhts.ornl.gov/.

| Agency | Data Program | Business Owner | Data Steward | Access Instructions |
|---|---|---|---|---|
| Federal Highway Administration (FHWA) (continued) | DCM Program Research Data Exchange (RDE) | RITA Bureau of Transportation Statistics http://www.its.dot.gov/ | Dale Thompson, dale.thompson@dot.gov; Gene McHale, gene.mchale@dot.gov | The Research Data Exchange is accessible through a common web-based Data Portal.[40] Users must register with the system to obtain access to data and participate in the community resources provided by the Research Data Exchange. |
| | DCM Program Test Data Sets | RITA Bureau of Transportation Statistics http://www.its.dot.gov/ | Dale Thompson: (202) 493-0259, dale.thompson@dot.gov | Test data sets and resulting documentation will be made available through the Research Data Exchange.[41] |
| | Dynamic Mobility Application Datasets | RITA Bureau of Transportation Statistics http://www.its.dot.gov/ | Katherine K. Hartman, (202) 366-2742, kate.hartman@dot.gov; Robert Rupert, (202) 366-2194, robert.rupert@dot.gov; Steven Mortensen, (202) 493-0459, steven.mortensen@dot.gov; Randy Butler, (202) 366-9215, randy.butler@dot.gov | Dynamic Mobility Application data sets and resulting documentation will be made available through the Research Data Exchange.[42] |
| | Omnibus Household Survey | RITA Bureau of Transportation Statistics http://www.its.dot.gov/ | Joy Sharp, RITA BTS, joy.sharp@dot.gov | All OHS data for 2000, 2001, 2002, 2003, 2005, and 2009 are accessible to the public and are available for download on the BTS website[43] http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/subject_areas/omnibus_surveys/household_survey/index.html. The website serves as a central facility to obtain information on the OHS, sample design and data collection procedures, OmniStats publications using OHS data, and the raw OHS data which is available for download and analysis by the user community. |

[40] U.S. DOT Research Data Exchange (RDE) Website, https://www.its-rde.net/.
[41] U.S. DOT Research Data Exchange (RDE) Website, https://www.its-rde.net/.
[42] U.S. DOT Research Data Exchange (RDE) Website, https://www.its-rde.net/.
[43] Omnibus Household Survey website, http://www.bts.gov/programs/ omnibus_surveys/household_survey/.

| Agency | Data Program | Business Owner | Data Steward | Access Instructions |
|---|---|---|---|---|
| Federal Transit Administration (FTA) | National Transit Database (NTD) | Federal Transit Administration http://www.fta.dot.gov/ | John D. Giorgis, NTD Program Manager, Federal Transit Administration, 202-366-5430, john.giorgis@dot.gov | NTD data is accessible to the public and is available on the NTD Program website.[44] |
| Federal Motor Carrier Safety Administration (FMCSA) | Motor Carrier Management Information System (MCMIS) | Federal Motor Carrier Safety Administration http://www.fmcsa.dot.gov/ | Bill Bannister, William.Bannister@dot.gov | Most information is available to the public through the MCMIS Data Dissemination Program.[45] Personally identifiable information (PII) can only be accessed by designated individuals, through a written request that the FMCSA Information System Division Office reviews and approves. Responsibility for disseminating the data, and the authority to charge prescribed fees for the data, have been delegated to a government contractor, Computing Technologies, Inc. |
| Federal Motor Carrier Safety Administration (FMCSA) | Licensing and Insurance (L&I) Database | Federal Motor Carrier Safety Administration http://www.fmcsa.dot.gov/ | 1-800-832-5660 | Some L&I information is accessible to the public and is available on the public page of the L&I public facing website.[46] Representatives of insurance companies, surety companies, and financial institutions can register with the FMCSA as an electronic insurance filer to view file certifications and cancellations, last transmissions, and print confirmation, acceptance, and reject reports through the L&I Internal Information System.[47] Insurance filers can request a user name and password by calling (202) 385-2423 M-F 8:00 a.m. – 4:30 p.m. EST or (202) 385-2400 after hours or holidays to leave a message. |

[44] National Transit Data Program Website, http://www.ntdprogram.gov/ntdprogram/ data.htm.
[45] MCMIS Catalog and Documentation website, http://mcmiscatalog.fmcsa.dot.gov.
[46] FMCSA L&I Public website, http://li-public.fmcsa.dot.gov/LIVIEW/pkg_html.prc_limain.
[47] Licensing and Information Internal Information System website, http://li-public.fmcsa.dot.gov/LIVIEW/pkg_menu.prc_menu.

| Agency | Data Program | Business Owner | Data Steward | Access Instructions |
|---|---|---|---|---|
| All Stakeholders | Data.gov | Hosted by the U.S. General Services Administration. Departments/Agencies submitting datasets to Data.gov retain ownership of the data | Contact information varies. For each dataset, a link is provided to contact the dataset owner. These messages are then sent to the relevant agency for review | Data.gov datasets are accessible to the public and are available for download through the Data.gov website.[48] The Dataset Management System (DMS) website provides an automated process for departments/agencies to submit datasets to the Data.gov catalog. |
| I-95 Corridor Coalition | I-95 Corridor Coalition Data | I-95 Corridor Coalition http://www.i95coalition.org/i95/Default.aspx | Bill Stoeckert, I-95 Corridor Coalition, 774-207-0367, wstoeckert@yahoo.com; Phil Tarnoff, University of Maryland, 301-403-4619, tarnoff@eng.umd.edu; Stan Young, University of Maryland, 301-403-4593, seyoung@umd.edu | Access is limited to I-95 Coalition member agencies and their contractors. Access to data is granted once a member agency has signed a Data Use Agreement through the University of Maryland. The I-95 Traffic Monitoring website[49] provides a visualization of traffic conditions on a map and access to archived data. |
| State DOTs | State Roadway Inventory Programs | State DOTs | Varies with State, but several State DOTs have a data management office responsible for maintaining the road inventory data. This may be the responsibility of the Planning Division, the Asset Management Office or other offices, depending upon the organizational structure of the DOT | The State DOT determines how and what data is provided internally and externally. This is usually through a web interface, or FTP server if large quantities of data are requested. |
| | State Traffic Monitoring Stations (for Planning Data Programs) | State DOTs | Varies at the State and local government levels | The State DOT determines how and what traffic data is provided internally and externally. This is usually through a web interface, or FTP server if large quantities of data are requested. |
| Michigan DOT (MDOT) | VII Data Use Analysis and Processing (DUAP) Project | Michigan DOT http://www.michigan.gov/mdot/ | Michigan DOT, 517-373-2090 | Access to data is limited to MDOT and its associated stakeholders. |

[48] Data.gov website, http://www.data.gov.
[49] I-95 Traffic Monitoring website, http://i95.inrix.com.

## D.6 Assess Data Governance Maturity

An important component of establishing the need for data governance and for starting the discussion regarding data governance with stakeholders is to assess the state of data governance within an organization using a maturity model.

A data management maturity model is a framework describing aspects of the development of an organization with respect to a certain process. It is a helpful tool to assess where an organization stands with respect to implementing certain processes. A maturity model also can be used to benchmark for comparison or assist an agency in understanding common concepts related to an issue or process. A typical maturity model identifies levels and characteristics of those levels. The model can be used to assess an agency's status and assist in identifying next steps to achieve success toward an ultimate goal state.

An example maturity model (Table D-3) was developed to document levels of maturity related to the development and implementation of data management and data governance principles with respect to employees/processes, technology/tools and institutional/governance. It was developed based on *NCHRP 666*. The criteria (employees/processes, technology/tools and institutional/governance) are described in more detail below:

- **Employees/processes** – This refers to the willingness, understanding, and commitment of employees within the agency (e.g., FHWA Office of Operations) to embrace data management. It also refers to processes that may be in place to assure employees understand and appreciate the value of data management.
- **Technology/tools** – This refers to the use of tools and techniques designed to assist the agency (e.g., FHWA Office of Operations) managing, integrating, analyzing, and reporting data.
- **Institutional/governance** – Refers to the institutional structure within an agency (e.g., FHWA Office of Operations) to ensure consistent management of data programs.

The levels of maturity are somewhat generic in nature and are described as: 0 – Ad Hoc; 1 – Aware; 2 – Planning; 3 – Defined; 4 – Managed; 5 – Integrated; and 6 – Continuously Improving. Table 3-3 documents the levels of maturity within the categories.

This or a similar model could be used to assess the overall status of data management within the FHWA Office of Operations as it pertains to roadway travel mobility data programs.

**Table D-3. Example Data Management Maturity Model**

| Level | 0 – Ad Hoc | 1 – Aware | 2 – Planning | 3 – Defined | 4 – Managed | 5 – Integrated |
|---|---|---|---|---|---|---|
| **A. Technology/ Tools** | No tools in place | Planning for tools to support data management in some offices | Planning for tools to support data management across agency or for a specific office | Implemented some tools to support data management but not widespread across agency | Widespread implementation of tools to support data management but not integrated | Integrated, widespread implementation of tools to support data management and performance measurement |
| **B. Employees/ Awareness** | Not aware of need for improved data management | Aware of need for improved data management<br>No action has been taken | Aware of need for improved data<br>Some steps have been made within agency to improve technology or institutional setting to support data management in at least one office | Aware of need for improved data management<br>Some steps have been made within agency to improve both technology and institutional setting to support data management in more than one office | Aware of need for improved data<br>Improvements are under way to improve both technology and institutional setting to support data management across the agency | Aware of need for improved data management<br>Technology and institutional processes are in place to support data management for performance measures |
| **C. Institutional/ Governance** | No data governance in place | Agency is discussing needs/plans for data governance | Some level of data program assessment and formulation of roles for data managers is underway in one or more offices of agency | Data Business Planning underway – including development of governance model for multiple offices in agency | Data Business Plan developed with data assessment complete and data governance structure defined | Fully operational data governance structure in place |

During the assessment, stakeholders should discuss steps to achieve the desired levels of maturity for data governance. For example, the following action steps could be used by the FHWA Office of Operations to advance data governance and ensure consistent management of data programs:

1. Formally adopt the data governance principles identified in the Data Business Plan and incorporate them into FHWA Office of Operations policies, standards, and processes and publish them in a Data Governance Manual.
2. Develop and formally adopt a data governance policy within the FHWA Office of Operations and incorporate this policy into a published Data Governance Manual.
3. Identify a champion or team leader for data governance within the FHWA Office of Operations to lead the FHWA Office of Operations Data Governance Team and participate in the agency-wide FHWA Data Governance Advisory Council.
4. Create an FHWA Office of Operations Data Governance Team to include representatives from each of the FHWA Office of Operations offices (HOFM, HOTM, and HOTO).
5. Assign data governance roles within the FHWA Office of Operations, including data stewards, data business owners, data custodians, data collectors, and data architects for roadway travel mobility data.
6. Define detailed responsibilities for the previously mentioned data governance roles within the FHWA Office of Operations, including data stewards, data business owners, data custodians, data collectors, and data architects for roadway travel mobility data.
7. Identify and establish the internal/external Roadway Travel Mobility Data Working Groups and Communities of Interest.
8. Conduct group meetings with the Data Working Group and Internal Community of Interest to: 1) begin ongoing forums of communication and coordination; 2) assess the overall status of data management within the FHWA Office of Operations as it pertains to roadway travel mobility data; 3) discuss steps to achieve the desired levels of maturity for data governance using the Data Management Maturity Model; and 3) implement the Data Business Plan.
9. Finalize the Data Governance Framework and Charter based on stakeholder input.
10. Update the Data Catalog to identify Business Owners and Data Stewards responsible for the data programs, once those roles have been assigned. Business Owners and Data Stewards should be identified for roadway travel mobility data.
11. Establish communication protocols with the Data Working Group and Internal/External Communities of Interest members for change management, issue escalation, and information dissemination.

# APPENDIX E. FHWA Data Governance Advisory Council Charter

# CHARTER

# FEDERAL HIGHWAY ADMINISTRATION (FHWA) DATA GOVERNANCE COUNCIL

## Article I. Purpose, Authority, and Duration

**Purpose:** The Federal Highway Administration (FHWA) Data Governance Advisory Council, hereinafter referred to as the "DGAC," is formally chartered and empowered to provide strategic review and oversight of all FHWA data collection efforts. The DGAC will consider guidance and information provided by the DOT Secretary, the FHWA Administrator and the FHWA Chief Information Officer (CIO) as part of its processes and functions. The DGAC has authority and responsibility to corporately advise on the utilization FHWA's data resources and recommend major changes in FHWA data collection efforts that will result in increased consistency and coordination between existing and new data programs; the elimination of redundant data collection; the consolidation of data sources and resources; and compliance with external mandates. The DGAC will present recommendations to the Investment Review Board ORB) for approval and prioritization.

**Authority:** The Council is formed under delegated authority from the Secretary of the DOT and FHWA Administrator, and in support of the Department's implementation of the following laws:

- The Clinger-Cohen Act of 1996;
- Federal Acquisition Streamlining Act of 1994 (FASA);
- Federal Information Security Management Act of 2002 (FISMA);
- E-Government Act of 2002 (E-Gov Act);
- Paperwork Reduction Act of 1995 (PRA); and
- Government Performance and Results Act of 1993 (GPRA).

The DGAC is also formed to assist in the Agency's compliance with various regulatory, policy, or procedural requirements of the OMB, and the DOT.

**Duration:** The DGAC is considered a permanent FHWA governance body.

## Article II. Scope and Responsibilities

Scope: The DGAC is an Agency-level, senior leadership governance committee whose scope includes:

- Corporately provide advice on the management of FHWA data assets;
- Providing recommendations on FHWA strategic data decisions and resource allocations to the FHWA leadership to obtain initial approval of data policies and standards; and
- Annually reviewing the FHWA data programs and making change recommendations to the FHWA-IRB for approval.

**Responsibilities:** The DGAC shall be responsible for orchestrating FHWA's major data collection efforts, including the pre-selection, selection, control and evaluation of individual data and entire data programs.

In addressing these stages, the DGAC shall perform the following functions:

- Develop a proposed plan for corporately managing FHWA data;
- Review existing data collection efforts for need, consistency and efficiency annually;
- Review and approve all new data collection efforts;
- Monitor and evaluate performance of data programs;
- Creation of a functional data dictionary;
- Recommend to the FHWA Investment Review Board ways to improve and streamline;
- existing and new data collection efforts; and
- Coordinate FHWA data collection efforts with other modes within the Department.

## Article III. Membership

**Membership:** The FHWA DGAC membership includes the following senior managers or their designees, who are all voting members:

- Director of Highway Policy Information, who serves as the Chair;
- Representative from the Chief Financial Office;
- Representative from the Chief Counsel Office;
- Representative from the Policy and Program Review, Federal Lands Highway;
- Representative from Bridge Technology;
- Representative from the Transportation Performance Management;
- Representative from the Safety Programs;
- Representative from Research, Development and Technology;
- Representative from the DA Council;
- Representative from the IT Advisory Group; and
- Representative from a Program Office that will rotate on an annual basis.

Note: Only one representative from each office is permitted.

## Article IV. Schedule

The DGAC will meet regularly at a time and place set by the Chair. The DGAC will meet at least once each quarter.

## Article V. Effective Date and Review

This charter is effective as of December 1, 2012. There are no cancellations associated with the implementation of this DGAC Charter.

David R. Winter, P.E.
Director, Office of Highway Policy Information

Sarah J. Shores
Associate Administrator for Administration

# APPENDIX F. Data Mapping

Figures F-1 through F-12 illustrate the relationship between data sets, core data systems, and internal/external stakeholders for the data. The figures illustrate how data sets currently (and in the future will) support the core data systems for roadway travel mobility data and connected vehicle data capture. The data mapping represents an example of a data management tool that can be applied.

**Figure F-1. Highway Performance Monitoring System (HPMS) Data Mapping** (Source: Cambridge Systematics, Inc., January 2013)

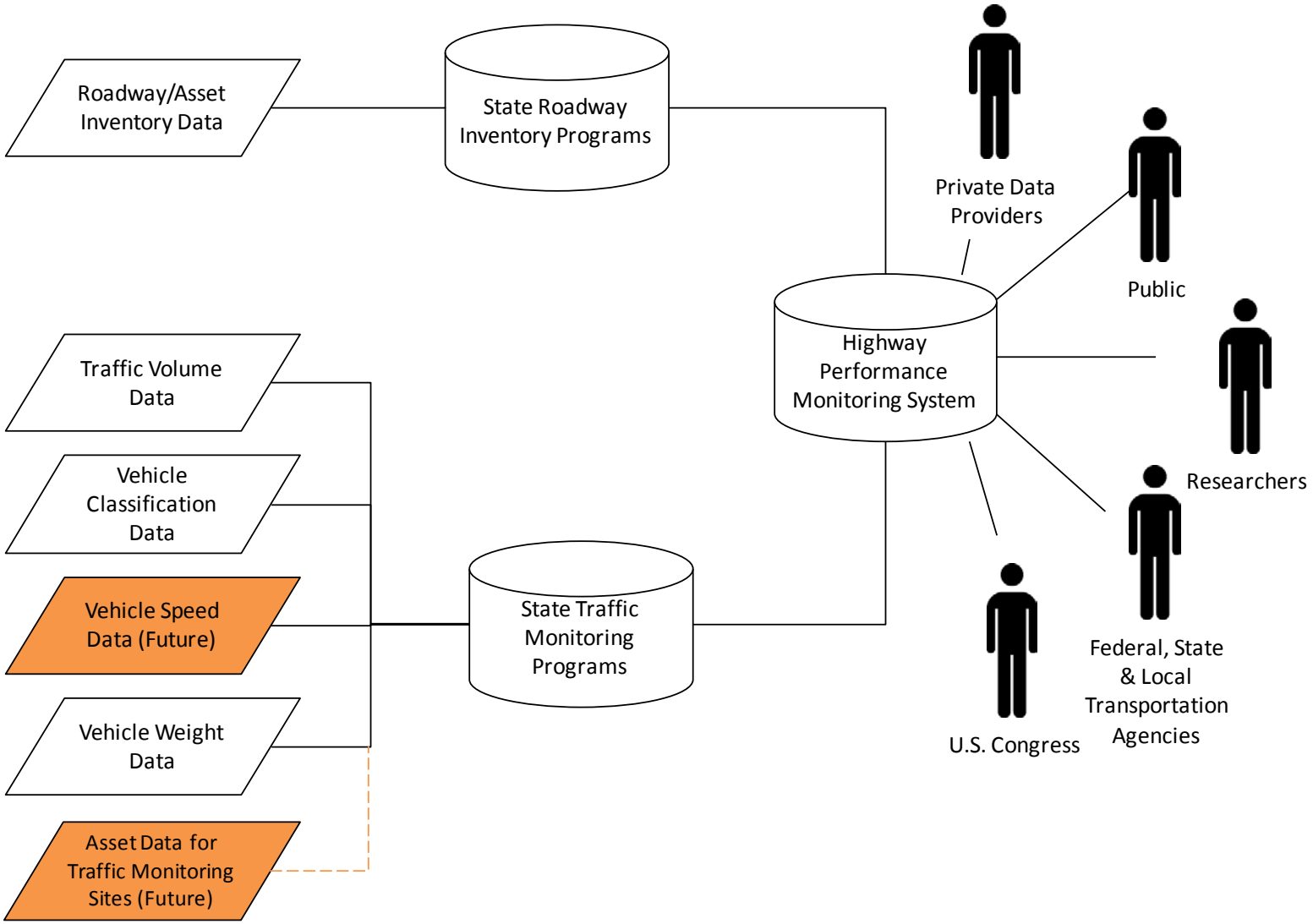
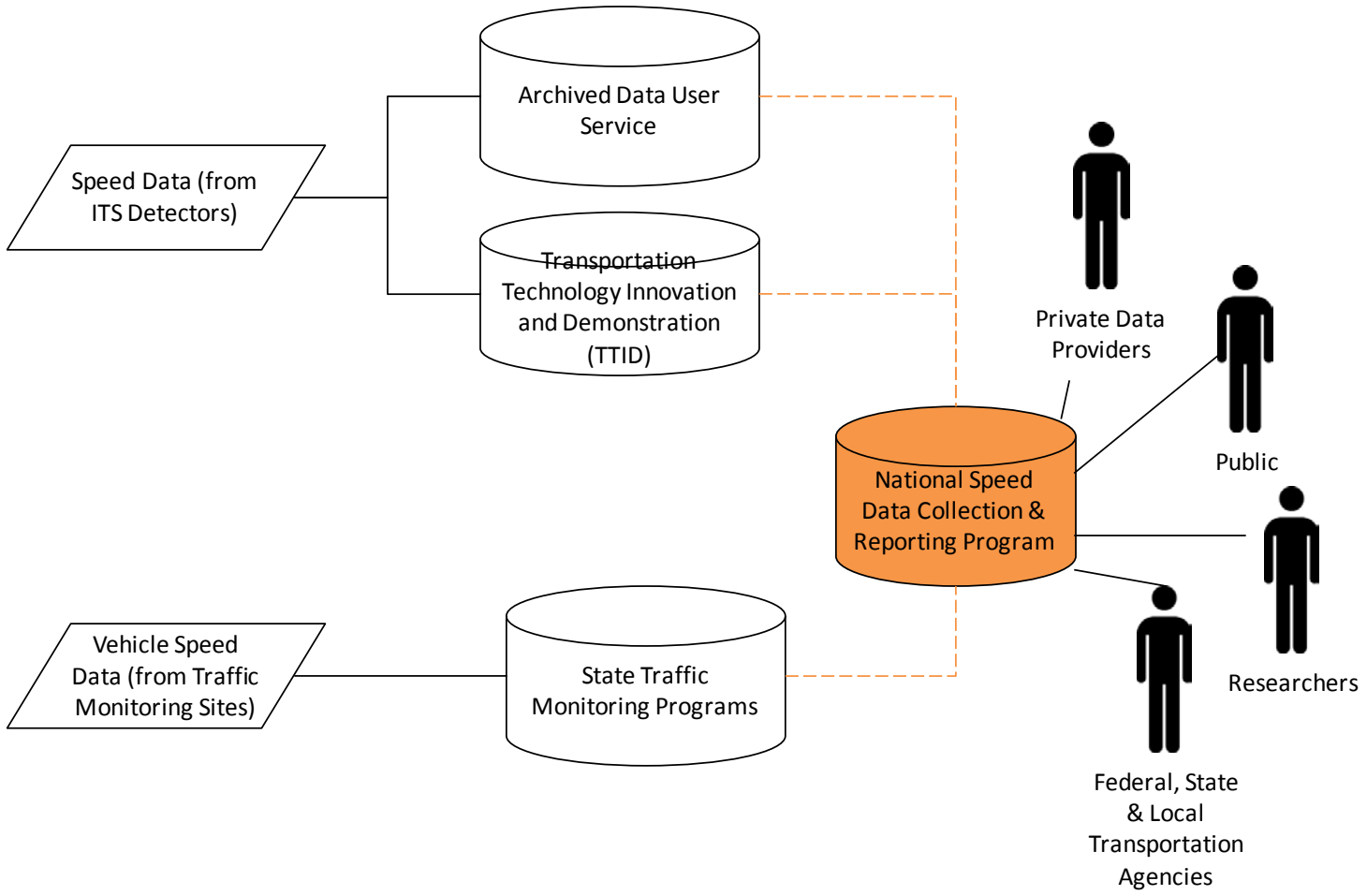


**Figure F-2. National Speed Data Collection and Reporting Program (Future) Data Mapping** (Source: Cambridge Systematics, Inc., January 2013)




NOTE: Orange symbols represent future data systems/programs.
Orange dashed lines represent future connections.

**Figure F-3. Travel Monitoring and Analysis System (TMAS) Data Mapping** (Source: Cambridge Systematics, Inc., January 2013)

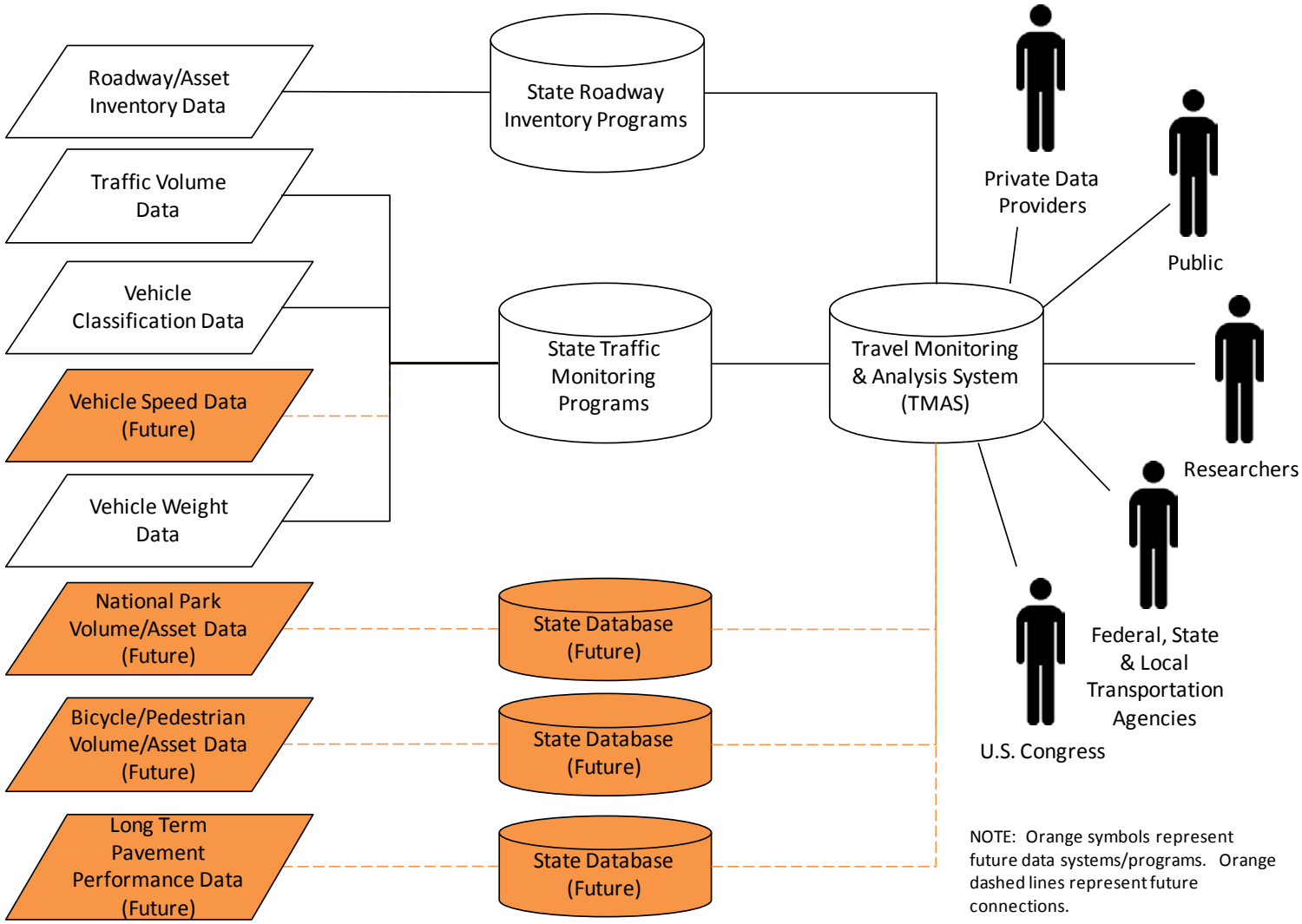


**Figure F-4. Vehicle Travel Information System (VTRIS) Data Mapping** (Source: Cambridge Systematics, Inc., January 2013)

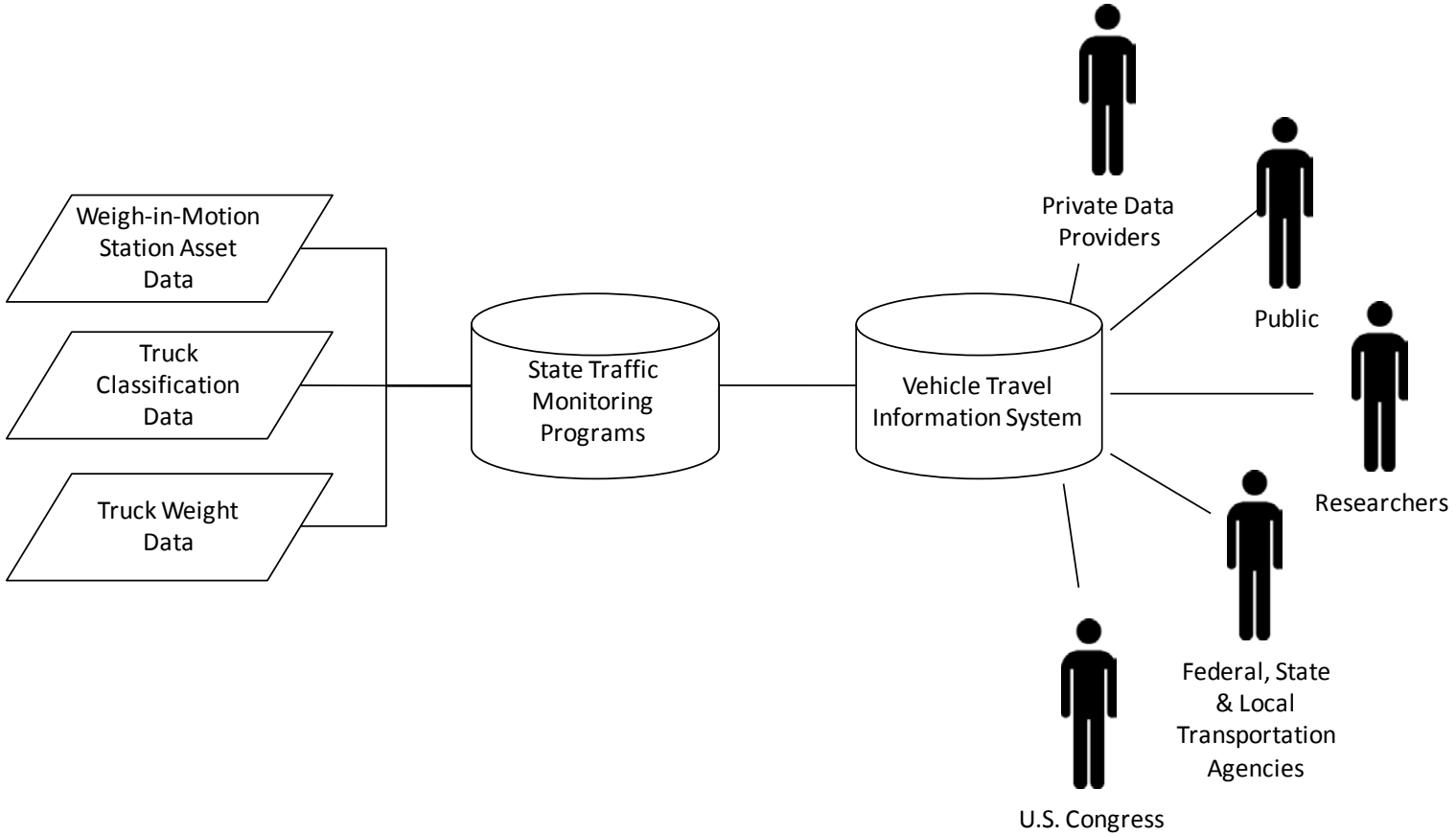


**Figure F-5. *Clarus* Data Mapping** (Source: Cambridge Systematics, Inc., January 2013)

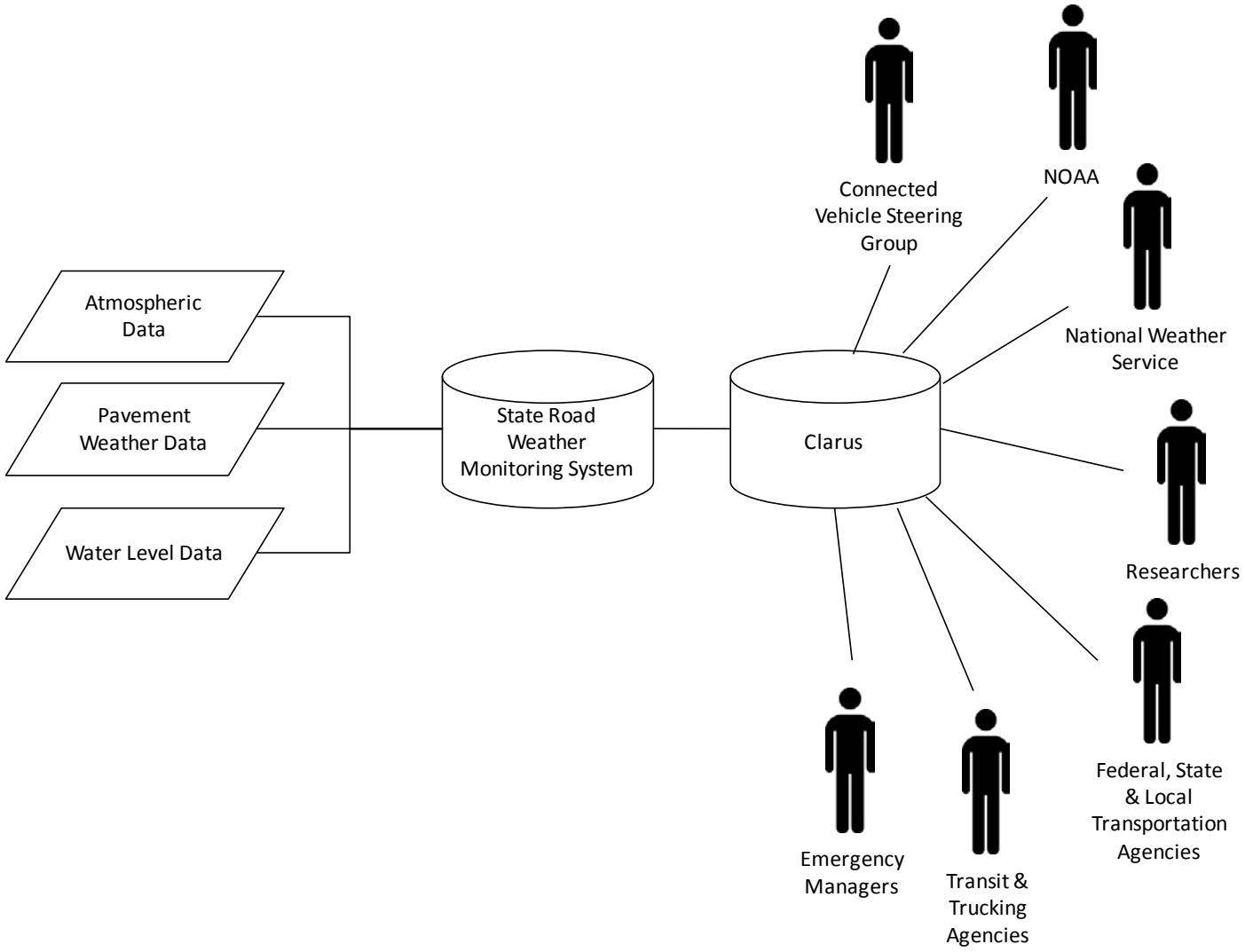


**Figure F-6. North American Transborder Freight Data Mapping** (Source: Cambridge Systematics, Inc., January 2013)

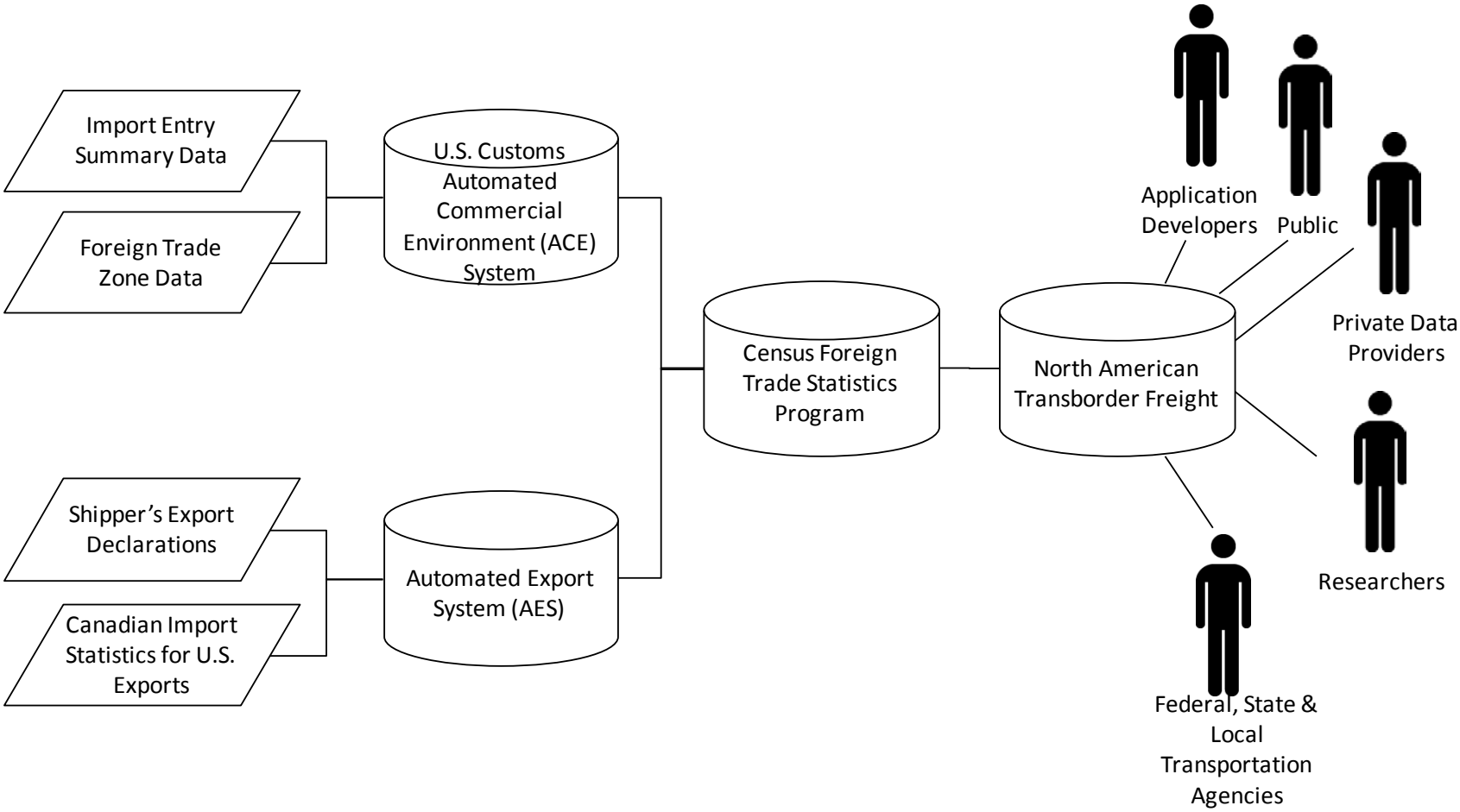
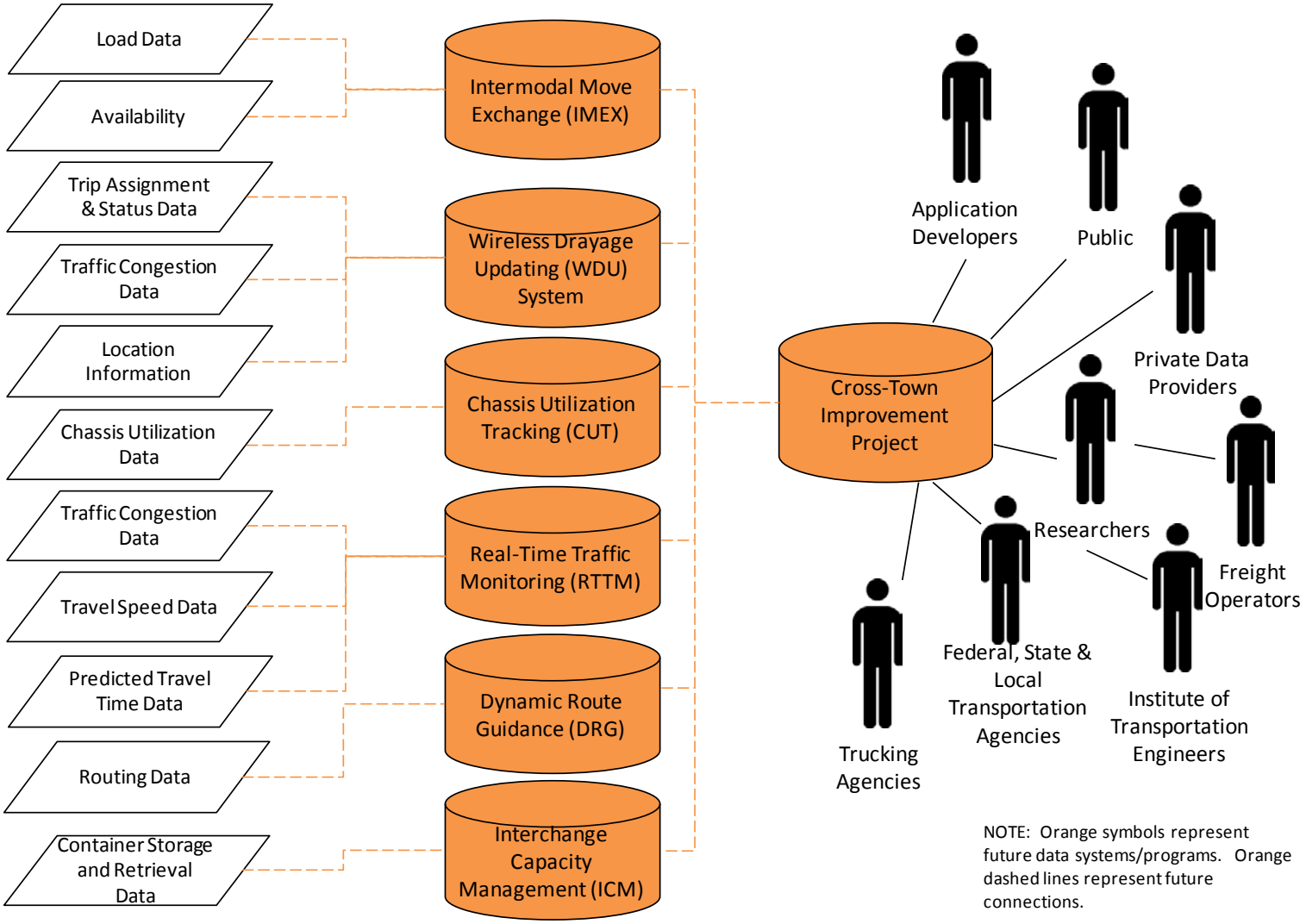


**Figure F-7. Cross-Town Improvement Project (C-TIP) (Future) Data Mapping** (Source: Cambridge Systematics, Inc., January 2013)




U.S. Department of Transportation, Research and Innovative Technology Administration
Intelligent Transportation System Joint Program Office

**Figure F-8. Freight Analysis Framework Data Mapping** (Source: Cambridge Systematics, Inc., January 2013)

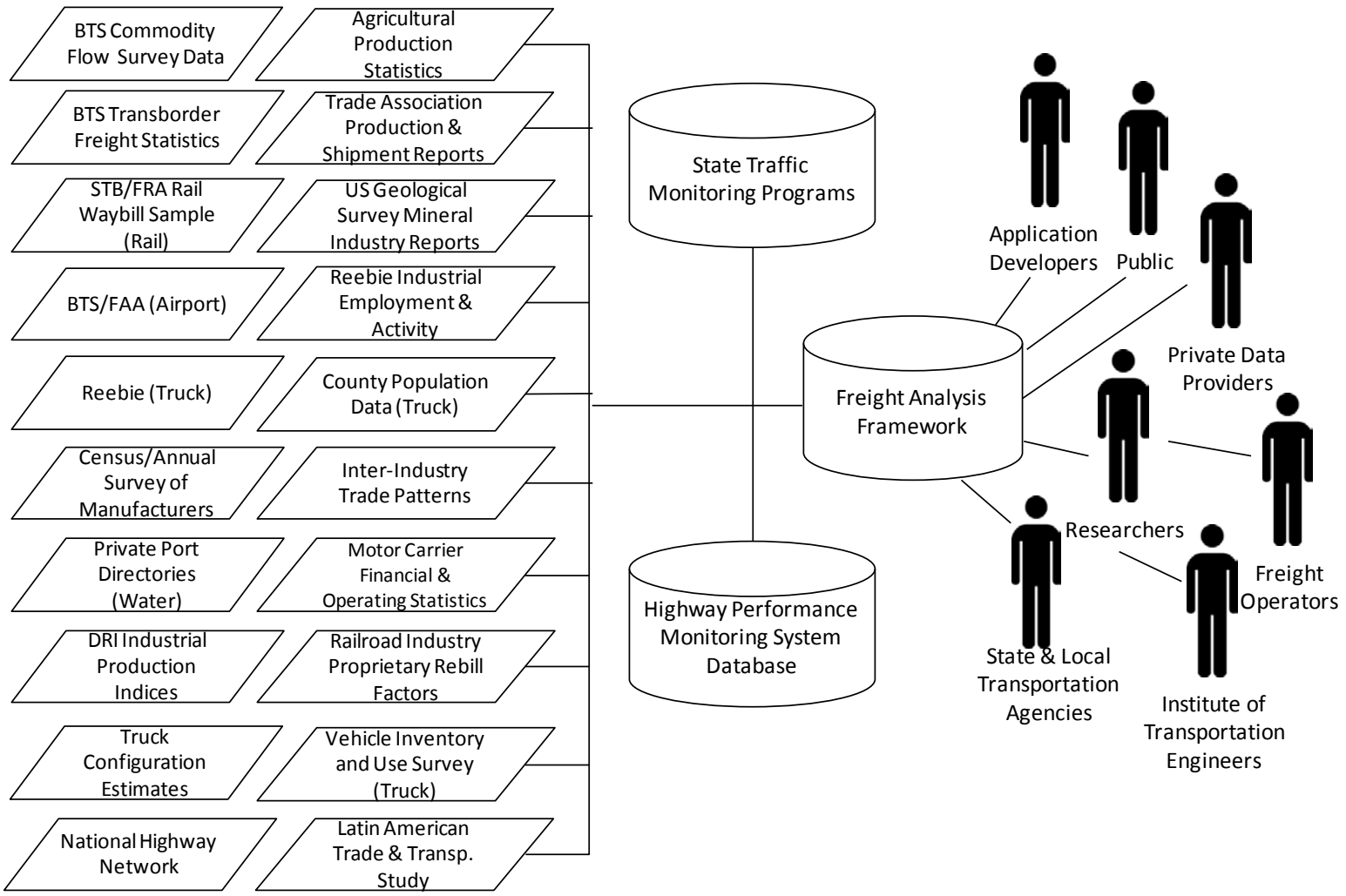


**Figure F-9. Freight Performance Measures Initiative Data Mapping** (Source: Cambridge Systematics, Inc., January 2013)

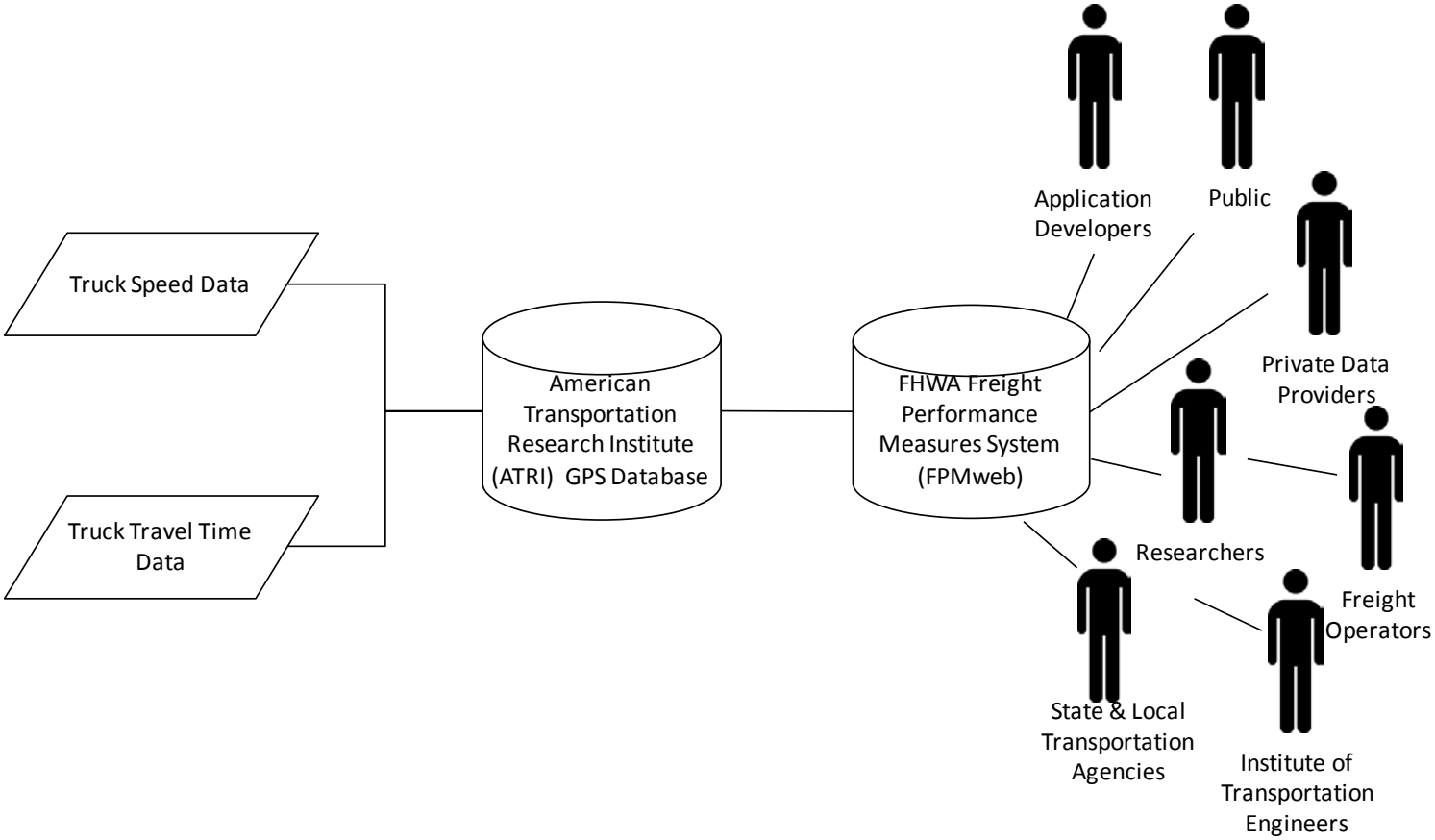


**Figure F-10. International Freight Data System (IFDS) (Future) Data Mapping** (Source: Cambridge Systematics, Inc., January 2013)

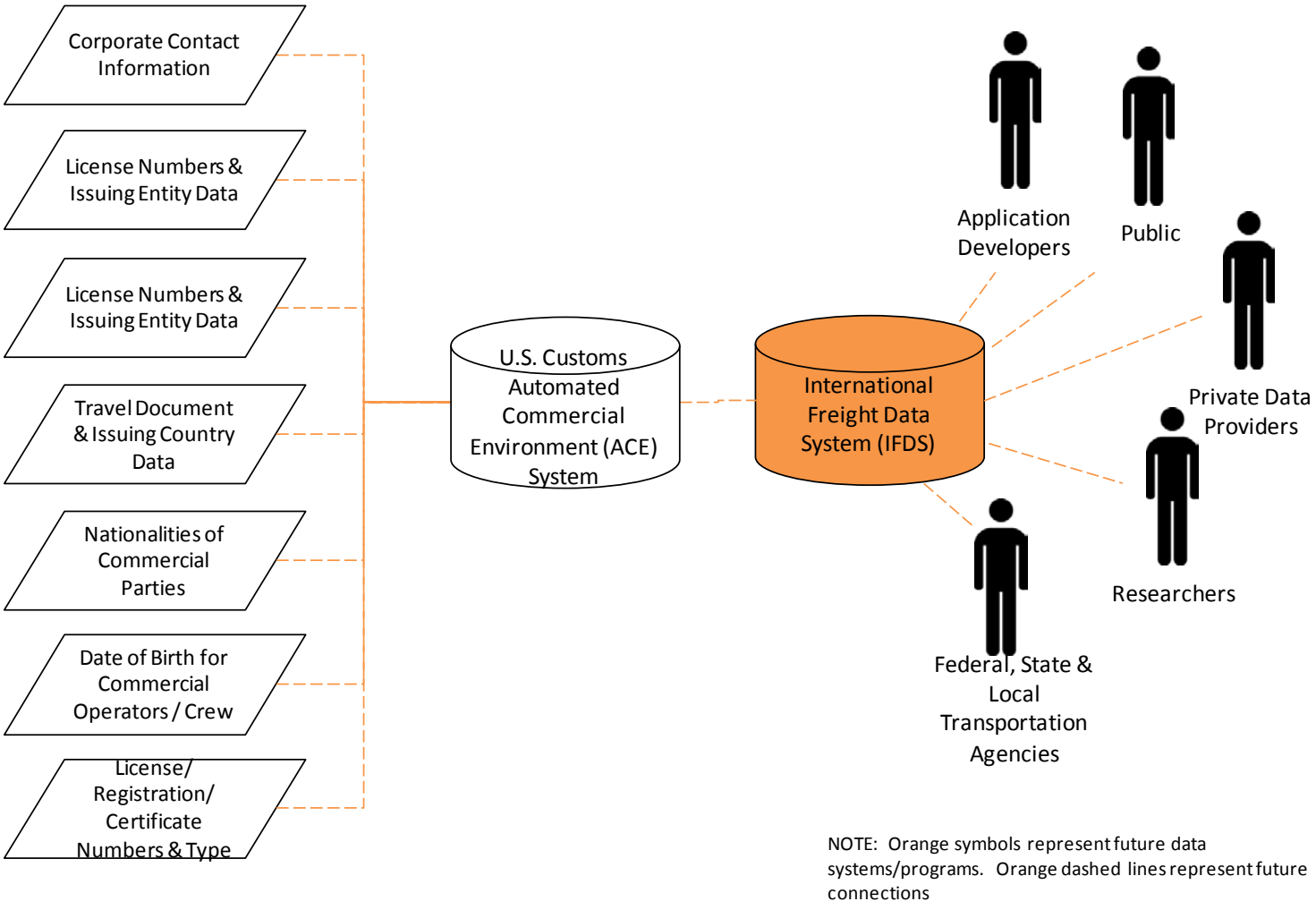


**Figure F-11. Data.gov Data Mapping** (Source: Cambridge Systematics, Inc., January 2013)

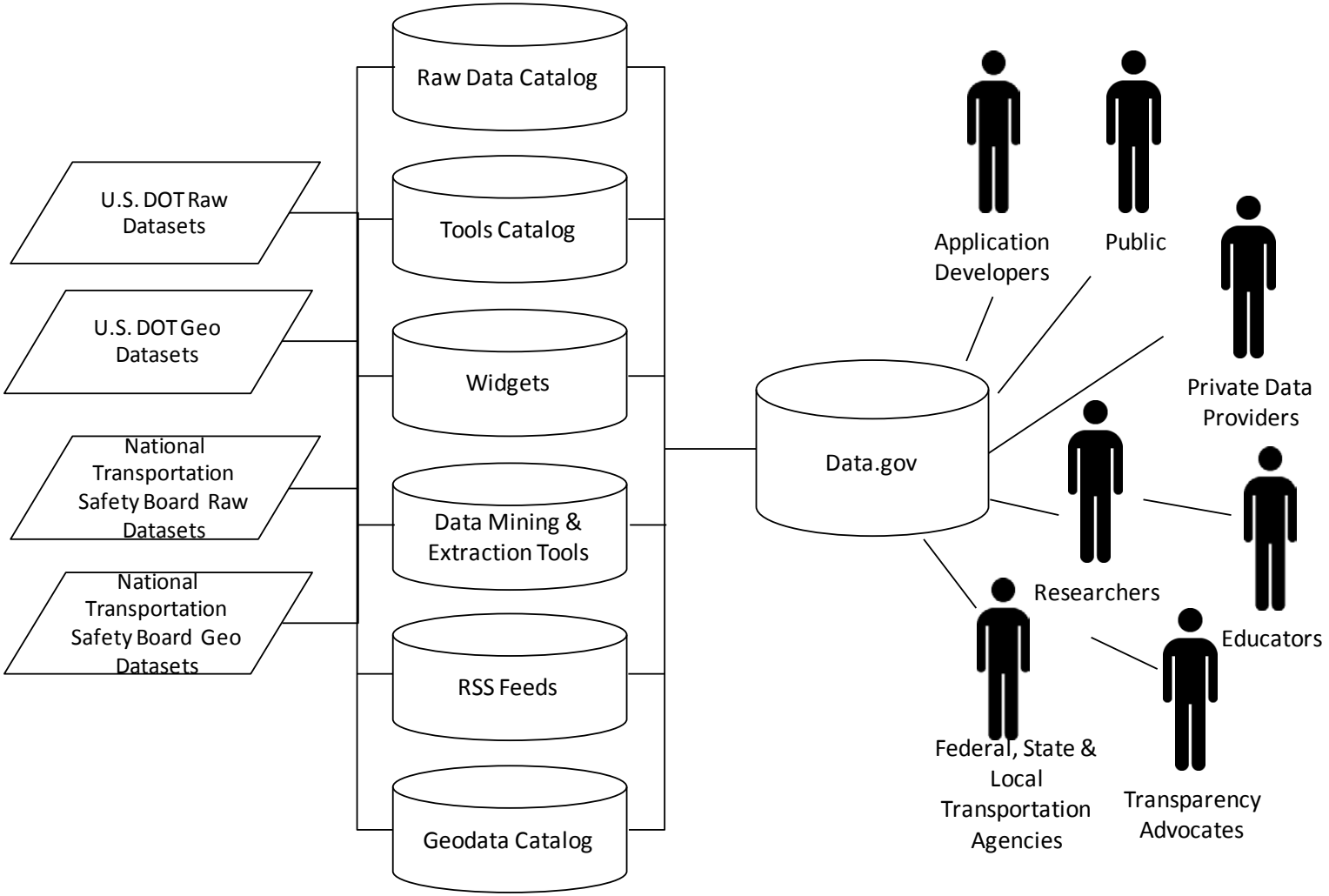
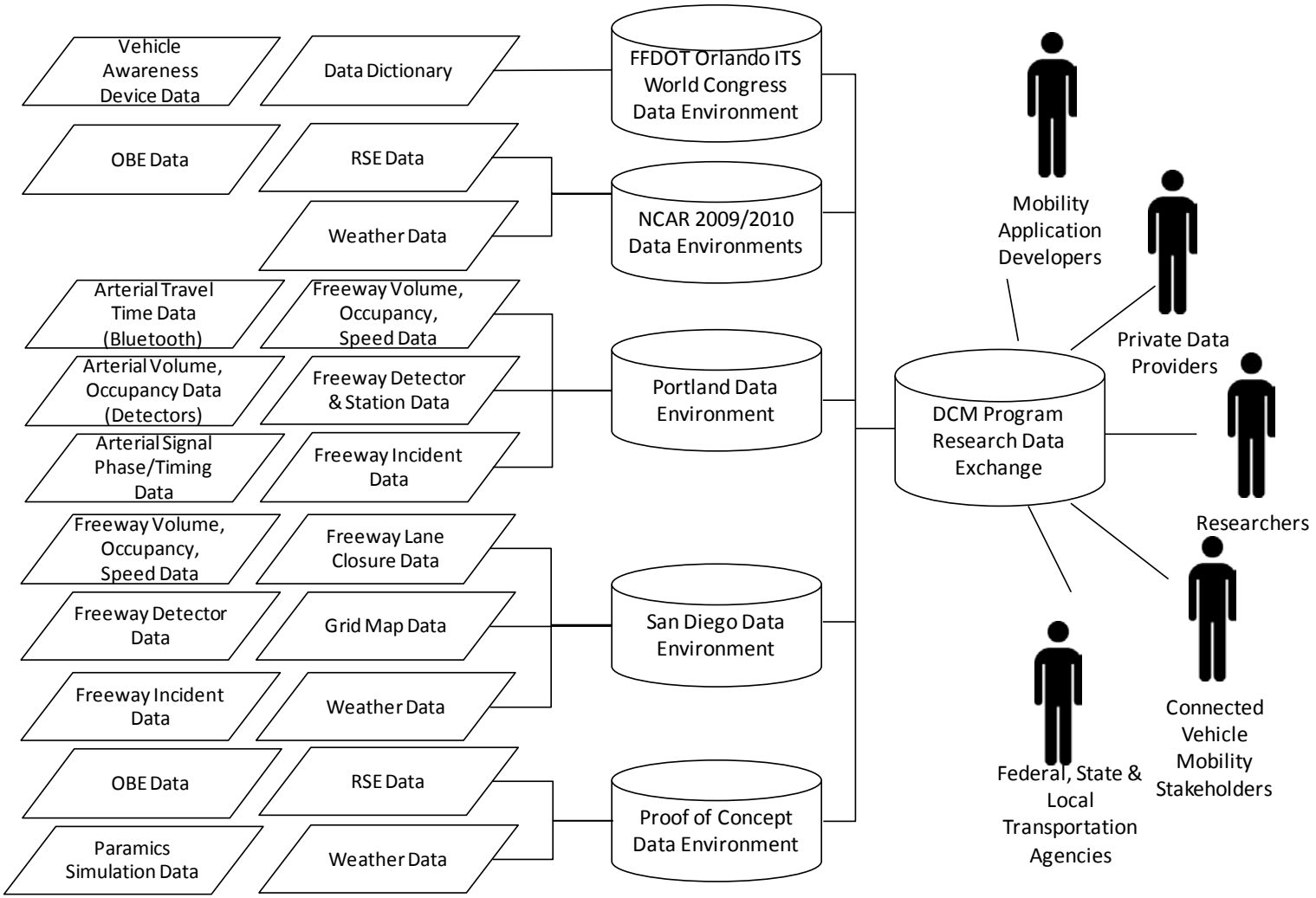


**Figure F-12. DCM Program Research Data Exchange Data Mapping** (Source: Cambridge Systematics, Inc., January 2013)




U.S. Department of Transportation, Research and Innovative Technology Administration
Intelligent Transportation System Joint Program Office

# APPENDIX G. U.S. DOT Mobility Data Coordination Group – Memorandum of Understanding

This Memorandum of Understanding (MOU) outlines the participants, the rules of engagement and operating guidelines for the U.S. DOT Mobility Data Coordination Group. By signing this MOU, each office agrees to participate in the Coordinating Group's activities and to share data and information with other participants.

## G.1 Participants

The members of the U.S. DOT Mobility Data Coordination Group are listed below. Each office will appoint a designated representative and alternate to attend the Group meetings and teleconferences.

U.S. DOT Mobility Data Coordination Group participants

- ITS Joint Program Office (JPO)
- RITA Bureau of Transportation Statistics (BTS)
- FHWA Office of Human Environment – Bicycle & Pedestrian Program (HEPH)
- FHWA Office of Highway Policy Information (HPPI)
- FHWA Office of Operations (HOTO)
- FHWA Office of Operations – Road Weather Management
- FHWA Office of Transportation Management (HOTM)
- Turner-Fairbank Highway Research Center (TFHRC)
- FHWA Office of Freight Management and Operations (HOFM)
- FHWA Office of Performance Management (TPM)
- Volpe (Federal Transit Administration (FTA ) representative)
- Federal Motor Carrier Safety Association (FMCSA)
- Noblis (DCM Program Technical Support)

Data Working Groups

The U.S. Roadway Transportation Data Business Plan recommended that several sub-groups be formed to coordinate and address subject specific issues in greater detail. The members of these working groups are listed below. Each office will appoint a designated representative and alternate to attend the Group meetings and teleconferences.

*Travel Data Working Group*

- ITS Joint Program Office (JPO)
- FHWA Office of Highway Policy Information (HPPI)

- FHWA Office of Performance Management (TPM)
- RITA Bureau of Transportation Statistics (BTS)
- FHWA Office of Operations (HOTO)
- FHWA Office of Transportation Management (HOTM)
- Turner-Fairbank Highway Research Center (TFHRC)
- FHWA Office of Freight Management & Operations (HOFM)

*Connected Vehicle Data Capture Working Group*

- ITS Joint Program Office (JPO)
- Turner-Fairbank Highway Research Center (TFHRC)
- FHWA Research and Development
- Noblis (DCM Program Technical Support)

*Other Working Groups to be determined- including Climate, Modal, and Infrastructure Working Groups*

## G.2 Rules of Engagement

The U.S. Roadway Transportation Data Business Plan recommended a list of activities or rules of engagement that the Mobility Data Coordination Group and the Working Groups should pursue. These are suggested activities that each group may use to form meeting agendas and action plans. The activities are listed in general order of priority, with sharing and coordinating activities to be conducted first and more labor intensive developmental activities occurring later.

1. Share RFP's for current and upcoming initiatives.
2. Share current initiatives, activities, and best practices related to specific types of roadway travel mobility data.
3. Share current activities and best practices related to data strategies, policies, standards, metadata, architecture, procedures and metrics.
4. Identify opportunities to coordinate resources to reduce data redundancy and implement cost sharing strategies for the collection, management, and maintenance of roadway travel mobility data. Redundancy could be addressed through data standardization and an annual review of data programs to identify where duplicate data collection and storage can/should be eliminated in and replaced with a single source of data for specific data programs. This will help to ensure that data is collected once and used many times.
5. Identify opportunities to reduce redundancy in the development and maintenance of duplicate data systems, promote efficiency in system maintenance, and promote open source initiatives.
6. Identify needs and opportunities to integrate national data sets to support performance measurement and asset management purposes.
7. Identify needs and opportunities to create links between existing data sets and connected vehicle data sets in the future.

8. Develop data standards and stewardship recommendations for consideration by the FHWA Data Governance Advisory Council.. Data standards are very important to achieve harmonization across stakeholders and should be a prominent data management practice in the rules of engagement.
9. Explore methods to enhance access to information and data for the roadway travel mobility data programs. This includes developing web portals easily accessible by internal and external stakeholders for each of these programs to obtain data and information as needed. This will facilitate sharing of data with internal/external stakeholders, thereby reducing costs associated with data collection.
10. Understand and promote the value of data as a U.S. DOT-wide asset.

The following table excerpted from the U.S. Roadway Transportation Data Business Plan (Table 3.1) outlines roles and responsibilities of each Group.

**Table G-1. Stakeholder Roles and Responsibilities**

| Role | Membership | Responsibility |
|---|---|---|
| U.S. DOT Mobility Data Coordination Group | • ITS Joint Program Office (JPO)<br>• RITA Bureau of Transportation Statistics (BTS)<br>• FHWA Office of Human Environment – Bicycle and Pedestrian Program (HEPH)<br>• FHWA Office of Highway Policy Information (HPPI)<br>• FHWA Office of Operations (HOTO)<br>• FHWA Office of Operations – Road Weather Management<br>• FHWA Office of Transportation Management (HOTM)<br>• Turner-Fairbank Highway Research Center (TFHRC)<br>• FHWA Office of Freight Management and Operations (HOFM)<br>• FHWA Office of Performance Management (TPM)<br>• Volpe (Federal Transit Administration (FTA ) representative)<br>• Federal Motor Carrier Safety Association (FMCSA)<br>• Noblis (DCM Program Technical Support) | Finalize Data Coordination Framework with input from Data Working Groups and Internal Community of interest<br>Develop and approve U.S. Mobility Data Coordination Group charter |

| Role | Membership | Responsibility |
|---|---|---|
| Travel Data Working Group | • FHWA Office of Highway Policy Information (HPPI)<br>• RITA Bureau of Transportation Statistics (BTS)<br>• FHWA Office of Operations (HOTO) | Address stakeholder needs related to travel data<br>Identify and address gaps and redundancies in travel data programs<br>Devise "rules of engagement" regarding collaboration and coordination<br>Develop data standards and stewardship recommendations for consideration by the U.S. DOT Mobility Data Coordination Group |
| Connected Vehicle Data Capture Working Group | • ITS Joint Program Office (JPO)<br>• Turner-Fairbank Highway Research Center (TFHRC)<br>• FHWA Research and Development<br>• Noblis (DCM Program Technical Support) | Address stakeholder needs related to connected vehicle data capture activities<br>Identify and address gaps and redundancies in connected vehicle data capture activities<br>Devise "rules of engagement" regarding collaboration and coordination<br>Develop data standards and stewardship recommendations for consideration by the U.S. DOT Mobility Data Coordination Group |

## G.3 Operational Guidelines

The following is a recommended list of operating guidelines for the U.S. DOT Mobility Data Coordination Group.

1. The Mobility Data Coordination Group will meet quarterly.
2. Each Working Group will meet quarterly prior to Mobility Data Coordination Group.
3. The Working Groups will report to Mobility Data Coordination Group.
4. The Mobility Coordination Data Group will be chaired by the FHWA Office of Operations champion.
5. Each Working Group's chair will rotate among the participants.
6. The Mobility Data Coordination Group chair will be the representative to the FHWA Data Governance Advisory Council.
7. The Mobility Data Coordination Group is responsible for disseminating relevant mobility data information to the internal and external communities of interest.

Agreed to by:

_____

U.S. DOT Office

_____

Date

_____

Name, Title

# APPENDIX H. Data Inventory

## H.1 Introduction

### Background

The white paper, *Data Capture and Management: Needs and Gaps in the Operation and Coordination of U.S. DOT Data Capture and Management Programs,*[50] examined current data capture, archiving, and provisioning activities, including the Intelligent Transportation Systems (ITS) Real-Time Data Capture and Management (DCM) program and other data-related Federal programs. The white paper identified data capture and management gaps through independent research and outreach with U.S. DOT program stakeholders. Gaps identified in the document include the operations, maintenance and management of data programs, stakeholder and program cooperation/coordination, and the role of U.S. DOT in data integration. The *Needs and Gaps* white paper provides the foundation for the development of a Data Business Plan (DBP) for managing the U.S. DOT Data Capture and Management programs.

Phase 1 of the Data Business Plan will address these gaps by developing an implementable process to accomplish the following goals:

- GOAL 1. Improve the coordination and communication mechanism across U.S. DOT and FHWA offices involved with roadway travel mobility data to avoid investing resources in the same or similar types of data related programs.
- GOAL 2. Improve the coordination of the connected vehicle data capture program with other related data programs within U.S. DOT and FHWA.

The outcome of Phase 1 will be a Data Business Plan (DBP) describing a systematic process for the U.S. DOT to follow while conducting roadway travel mobility data-related activities. The Plan will describe action steps for meeting the following objectives:

- Objective 1.1. Identify gaps and redundancies in roadway travel mobility data programs.
- Objective 1.2. Devise “rules of engagement” regarding collaboration of the data functions for roadway travel mobility data.
- Objective 2.1. Identify how current and planned data from the connected vehicle initiative can support existing data programs.
- Objective 2.2. Identify how data from other mobility related programs within U.S. DOT and FHWA can support the connected vehicle initiative.

[50] A. Vandervalk, *Data Capture and Management, Needs and Gaps in the Operation and Coordination of U.S. DOT Data Capture and Management Programs*, Report No. FHWA-HOP-11-004, Cambridge Systematics, November 2010

- Objective 2.3. Identify existing/future data inventory and data structures/ policies/governance practices that could be applicable to the Research Data Exchange.

## Purpose

This report documents a data inventory of relevant national, State, and regional data programs that support the goals and objectives above. The purpose of the data inventory is to:

- Examine all Data Business Plan stakeholders and document the data programs they are using, their objectives, and how connected vehicle data could benefit these programs; and
- Create a standardized or level playing field for how all stakeholders are using data.

The report is organized as follows:

- Section H.1 Background on the Data Business Plan project and objectives of the Data Inventory report.
- Section H.2 A high level and detailed data inventory of relevant national, State, and regional data programs.
- Section H.3 Recommendations and next steps for the Roadway Transportation Data Business Plan project.

# H.2 Data Inventory

This section provides a high level and detailed inventory of relevant national, State, and regional data programs that support the goals and objectives of the Roadway Transportation Data Business Plan (described in Section 1.1). This is a comprehensive inventory that was compiled based on the *Data Capture and Management Program: Inventory of Data Warehousing Efforts*[51] report, the *Needs and Gaps* white paper, and U.S. DOT stakeholder input.

## Methodology

The data programs compiled in the *Inventory of Data Warehousing Efforts* report were reviewed as a starting point for the data inventory. Additional data programs were identified from both traditional (e.g., traffic monitoring programs, traffic management centers, etc.) and nontraditional (e.g., probe data, connected vehicle data, etc.) sources. The list of data programs was categorized and revised based on stakeholder input, and a preliminary assessment of these programs was presented in the *Needs and Gaps* white paper.

At the second kickoff meeting for the Roadway Transportation Data Business Plan project, which took place in Washington, D.C. on October 25, 2011, stakeholders agreed that core data programs for the Data Business Plan should include those related to roadway mobility, freight (FMCSA), and connected

[51] Noblis, *IntelliDrive[SM] Data Capture and Management Program: Inventory of Data Warehousing Efforts (Draft)*. Prepared for the RITA ITS Joint Program Office, February 2010.

vehicles. Therefore, data programs related to pavement, bridges and safety were excluded from the data inventory. Additional refinement of the data program list was accomplished through the stakeholder interview process, which took place in December 2011 and January 2012. The purpose of the interviews was to gain consensus on specific data programs/databases to be included in the Data Business Plan (data programs must be within the purview of one or more of the stakeholders) and major deficiencies within these databases.

The high-level data inventory was compiled to provide an assessment regarding data type (infrastructure, travel, climate, connected vehicle), timeliness, availability (existing, future), geographic scope (national, regional, project specific), mode (vehicle, truck freight, transit, bicycle/pedestrian), and data source hierarchy (raw data source, report, system, repository). The high-level data inventory is presented in table format, includes a brief description of the data programs, and identifies their owners.

The detailed data inventory provides a more comprehensive inventory for each data program. It is presented in text format and includes an inventory of the data owner, description, purpose, data source hierarchy, geographic scope/coverage, data format, system requirements to access the data, data quality, metadata, timeliness, longevity/future plans for the data program, primary users or customers, accessibility, and contact information for each data program/initiative.

The research for conducting both the high level and detailed data inventories was accomplished through Internet research, review of relevant reports, and interviews with relevant stakeholders.

## High Level Data Inventory

The high-level data inventory summarizes the following information for each data program:

- **Description.** A brief description of the data program/initiative.
- **Data Owner.** Identification of the FHWA Program Office or other entity responsible for the provision, capture, and management of data to support the data program.
- **Data Type.** Identifies the primary type of data included in the data program/initiative – infrastructure, travel, and climate – described as follows:
    - **Infrastructure Data** – Infrastructure data includes roadway geometry, roadway inventory, intersection characteristics, and the state of system controls. This could include roadway alignment, number and width of lanes, location of shoulder and median, posted speed limit, intersection configuration, or currently active signal-timing plan.
    - **Travel Data** – Travel data includes vehicle location, presence and speed within the system, internal vehicle status such as fuel consumption rate, or externally measured data such as recorded external temperature. Transit vehicles may contribute location, speed and status data, as well as passenger counts and schedule adherence data. Freight carriers may supplement a standard location and position report with gross weight data or data regarding the type and time-critical nature of goods carried. Public sector fleet vehicles may be able to contribute other key data related to their primary functions, such as snowplows reporting blade position or estimates of roadway snow depth. Additional data could include a multimodal trace of individual travelers through the transportation system.

  - **Climate Data** – Climate data includes data on weather and pavement surface conditions collected from roadway weather information systems (RWIS).
- **Timeliness.** Identifies the collecting/reporting frequency of the data program.
- **Availability.** Indicates whether the data program currently exists or will be available in the future.
- **Geographic Scope.** Indicates whether the data program is national, regional, or project specific in geographic scope/coverage.
- **Mode.** Identifies the modes that are included in the data program. In the original inventory, this column included surface, freight, air, and marine modes. Based on stakeholder input, this column was revised to include roadway transportation modes only (e.g., vehicle, truck freight, transit, and bicycle/pedestrian).
- **Data Source Hierarchy.** Indicates whether the data program is a source of raw data, a report, a system, or a repository.
- **Resources**. Website addresses for items mentioned in the section.

The high-level data inventory is shown in Table H.1. Currently available data programs are indicated with a solid filled circle (●), while future data sources are indicated with a hollow circle (○).

**Table H-1. Data Programs**

| Data Capture Program | Description | Data Owner | Data Type | | | Timelines | Availability | | Geographic Scope | | | Mode | | | | Data Source Hierarchy | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Infrastructure | Travel | Climate | Collection/Reporting Frequency | Existing | Future | National Data Program | Regional Data Program | Project Specific Data | Vehicle | Truck Freight | Transit | Bicycle/Pedestrian | Raw Data | Report | System | Repository |
| Infrastructure (inventory) | | | | | | | | | | | | | | | | | | | |
| State Roadway Inventory Programs | State roadway inventory programs document the extent (length), point and linear feature data of public roadways, along with attributes such as classification, ownership, physical conditions, pavement conditions, highway performance monitoring information, and more. Data is collected and stored for both on-state system and off-state system roadways. | State DOTs | ● | | | Varies | ● | | ● | | | ● | | | | ● | | | |
| Highway Performance Monitoring System (HPMS) | HPMS is a national level highway information system that includes data on the extent, condition, performance, use, and operating characteristics of the nation's highways. | FHWA Office of Highway Policy Information | ● | ● | | Annually | ● | | ● | | | ● | | | | | | ● | |

| Data Capture Program | Description | Data Owner | Data Type | | | Timelines | Availa-bility | | Geographic Scope | | | Mode | | | | Data Source Hierarchy | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Infrastructure | Travel | Climate | Collection/Reporting Frequency | Existing | Future | National Data Program | Regional Data Program | Project Specific Data | Vehicle | Truck Freight | Transit | Bicycle/Pedestrian | Raw Data | Report | System | Repository |
| Infrastructure (inventory) (continued) | | | | | | | | | | | | | | | | | | | |
| National Transporta-tion Atlas Databases | NTAD is a set of nationwide geographic databases of transportation facilities, networks, and associated infrastructure. The datasets include spatial information for transportation modal networks and intermodal terminals, as well as related attribute information for these features. | RITA Bureau of Transportation Statistics | ● | | | Annually | ● | | ● | | | ● | ● | ● | | | | | ● |
| Travel (speed and volume) | | | | | | | | | | | | | | | | | | | |
| Archived Data User Service (ADUS) | ADUS enables transportation agencies to retain ITS-generated volume, speed, and lane occupancy data and make them available for analysis. | FHWA Office of Highway Policy Information | | ● | | Collection: Varies (typ. 20-60 sec)<br>Archived: Varies (typ. 5-15 min) | ● | | | ● | | ● | | | | ● | | | |

| Data Capture Program | Description | Data Owner | Data Type | | | Timelines | Availa-bility | | Geographic Scope | | | Mode | | | | Data Source Hierarchy | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Infrastructure | Travel | Climate | Collection/Reporting Frequency | Existing | Future | National Data Program | Regional Data Program | Project Specific Data | Vehicle | Truck Freight | Transit | Bicycle/Pedestrian | Raw Data | Report | System | Repository |
| Travel (speed and volume) (continued) | | | | | | | | | | | | | | | | | | | |
| Transporta-tion Technology Innovation and Demon-stration (TTID) | The TTID Program is the primary source of loop detector data on many major interstate highways, particularly in smaller metropolitan areas that do not have ITS infrastructure in place. Polled data is converted into travel information (e.g., congestion levels, speed, etc.), and the data is then shared with local traffic management centers. | FHWA Office of Operations | | ● | | Collection: 60 sec<br>Archived: 1 min, 5 min, 15 min, 1 hr, 24 hr | ● | | | ● | ● | ● | | | | ● | | | |
| Urban Congestion Report (UCR) program | Reports characterize emerging traffic congestion and reliability trends at the national and city level. The reports utilize archived traffic operations data gathered from state DOTs and a private traffic information company. Performance measures include congested hours, travel time index, planning time index, percent change in VMT, and percent of usable data. | FHWA Office of Operations | | ● | | Compiled from 5-min interval data<br>Reported quarterly | ● | | | ● | ● | ● | | | | | ● | | |

| Data Capture Program | Description | Data Owner | Data Type | | | Timelines | Availability | | Geographic Scope | | | Mode | | | | Data Source Hierarchy | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Infrastructure | Travel | Climate | Collection/Reporting Frequency | Existing | Future | National Data Program | Regional Data Program | Project Specific Data | Vehicle | Truck Freight | Transit | Bicycle/Pedestrian | Raw Data | Report | System | Repository |
| Travel (speed and volume) (continued) | | | | | | | | | | | | | | | | | | | |
| State Traffic Monitoring Stations (for Planning Data Programs) | State DOTs install and maintain traffic monitoring stations throughout the state to support their traffic monitoring programs. These stations are established at both permanent and portable sites to collect volume, vehicle classification, and truck weight data. | State DOTs | | ● | | Hourly intervals (typ) | ● | | ● | | | ● | ● | | ○ | ● | | | |
| National Speed Data Collection and Reporting Program | FHWA is evaluating the feasibility of establishing a national speed data collection and reporting program. The proposed program would standardize speed data collection procedures, and provide monthly speed trends at the national level to evaluate travel trend and travel conditions. | FHWA Office of Highway Policy Information | | ○ | | TBD | | ○ | | | | ○ | | | | | | ○ | |

| Data Capture Program | Description | Data Owner | Data Type | | | Timelines | Availability | | Geographic Scope | | | Mode | | | | Data Source Hierarchy | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Infrastructure | Travel | Climate | Collection/Reporting Frequency | Existing | Future | National Data Program | Regional Data Program | Project Specific Data | Vehicle | Truck Freight | Transit | Bicycle/Pedestrian | Raw Data | Report | System | Repository |
| Travel (speed and volume) (continued) | | | | | | | | | | | | | | | | | | | |
| Traffic Volume Trend (TVT) Reports | The Traffic Volume Trends (TVT) report is published monthly by the Federal Highway Administration (FHWA). The report estimates the vehicle miles traveled (VMT) by state and several functional classes of roads. The estimates are based on the Highway Performance Monitoring System (HPMS) and monthly traffic counts from automatic traffic recorders (ATRs). | FHWA Office of Highway Policy Information | | ● | | Compiled from hourly interval data<br>Reported monthly | ● | | ● | | | ● | | | | | ● | | |
| Travel Monitoring and Analysis System (TMAS) | TMAS provides online data submitting capabilities to State traffic offices to submit travel monitoring data to FHWA. These data currently include monthly volume data for the Traffic Volume Trends report, vehicle classification, and truck weight data. | FHWA Office of Highway Policy Information | | ● | | Varies | ● | | ● | | | ● | ● | | ○ | | | ● | |
| Vehicle Travel Information System (VTRIS) | VTRIS is database management system designed to assist states in processing and storing vehicle classification and weigh-in-motion data. | FHWA Office of Highway Policy Information | | ● | | Hourly, daily, monthly, annually | ● | | ● | | | ● | ● | | | | | ● | |

| Data Capture Program | Description | Data Owner | Data Type | | | Timelines | Availability | | Geographic Scope | | | Mode | | | | Data Source Hierarchy | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Infrastructure | Travel | Climate | Collection/Reporting Frequency | Existing | Future | National Data Program | Regional Data Program | Project Specific Data | Vehicle | Truck Freight | Transit | Bicycle/Pedestrian | Raw Data | Report | System | Repository |
| Travel (speed and volume) (continued) | | | | | | | | | | | | | | | | | | | |
| National Transportation Statistics | NTS presents statistics on the U.S. transportation system, including its physical components, safety record, economic performance, energy use, and environmental impacts. | RITA Bureau of Transportation Statistics | ● | ● | | Quarterly | ● | | ● | | | ● | ● | ● | ● | | ● | | |
| Climate (weather) | | | | | | | | | | | | | | | | | | | |
| Road Weather Information Systems (Environmental Sensor Stations) | RWIS is a combination of technologies that collects, transmits, models, and disseminates weather and road condition information. | FHWA Office of Operations – Road Weather Management | | | ● | Varies | ● | | | ● | | ● | | | | ● | ● | | |
| Clarus | Clarus is a robust data assimilation, quality checking, and dissemination system that provides near real-time atmospheric and pavement observations from road weather information systems, environmental sensor stations (ESS), and mobile observations from Automated Vehicle Location (AVL) equipped trucks. | FHWA Office of Operations – Road Weather Management | | | ● | 5-min intervals | ● | | ● | | ● | ● | | | | | | ● | |

| Data Capture Program | Description | Data Owner | Data Type | | | Timelines | Availability | | Geographic Scope | | | Mode | | | | Data Source Hierarchy | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Infrastructure | Travel | Climate | Collection/Reporting Frequency | Existing | Future | National Data Program | Regional Data Program | Project Specific Data | Vehicle | Truck Freight | Transit | Bicycle/Pedestrian | Raw Data | Report | System | Repository |
| Modal | | | | | | | | | | | | | | | | | | | |
| Border Crossing Data | Border Crossing/Entry Data provides summary statistics for incoming crossings at the U.S.-Canadian and the U.S.-Mexican border at the port level. Data are available for trucks, trains, containers, buses, personal vehicles, passengers, and pedestrians. | RITA Bureau of Transportation Statistics | | ● | | Collected monthly, but updates are sent to BTS quarterly | ● | | ● | | | ● | ● | ● | ● | | | | ● |
| North American Transborder Freight Data | The North American Transborder Freight Database contains freight flow data by commodity type and by mode of transportation (rail, truck, pipeline, air, vessel, and other) for U.S. exports to and imports from Canada and Mexico. | RITA Bureau of Transportation Statistics | | ● | | Monthly | ● | | ● | | | | ● | | | | | | ● |

| Data Capture Program | Description | Data Owner | Data Type | | | Timelines | Availability | | Geographic Scope | | | Mode | | | | Data Source Hierarchy | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Infrastructure | Travel | Climate | Collection/Reporting Frequency | Existing | Future | National Data Program | Regional Data Program | Project Specific Data | Vehicle | Truck Freight | Transit | Bicycle/Pedestrian | Raw Data | Report | System | Repository |
| Modal (continued) | | | | | | | | | | | | | | | | | | | |
| Commodity Flow Survey | The CFS is the primary source of national and state-level data on domestic freight shipments in mining, manufacturing, wholesale, auxiliaries, and selected retail industries. Data are provided on the types, origins and destinations, values, weights, modes of transport, distance shipped, and ton-miles of commodities shipped. The CFS covers shipments by truck, rail, water, air, pipeline and multiple modes. CFS outputs include tabular data and reports. | RITA Bureau of Transportation Statistics | | ● | | 5 years | ● | | ● | | | | ● | | | | ● | | |
| Intermodal Passenger Connectivity Database | A nationwide data table of passenger transportation terminals, with data on the availability of connections among the various scheduled public transportation modes at each facility. In addition to geographic data for each terminal, the data elements describe the availability of rail, air, bus, transit, and ferry services. | RITA Bureau of Transportation Statistics | | ● | | Varies | ● | | ● | | | | | ● | | ● | | | |

| Data Capture Program | Description | Data Owner | Data Type | | | Timelines | Availability | | Geographic Scope | | | Mode | | | | Data Source Hierarchy | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Infrastructure | Travel | Climate | Collection/Reporting Frequency | Existing | Future | National Data Program | Regional Data Program | Project Specific Data | Vehicle | Truck Freight | Transit | Bicycle/Pedestrian | Raw Data | Report | System | Repository |
| Modal (continued) | | | | | | | | | | | | | | | | | | | |
| National Transit Database (NTD) | The NTD is a national database of statistics for the transit industry. The NTD is comprised of data reported by more than 600 transit agencies across the U.S., which is then analyzed and compiled into reports published by FTA and made available to the public on the NTD Program website. | Federal Transit Admin-istration | | ● | | Monthly, Annually | ● | | ● | | | | | ● | | ● | ● | | |
| Cross-Town Improve-ment Project (C-TIP) | C-TIP is a freight information sharing/transfer system that will enable coordination to maximize loaded movements and minimize unproductive movements within the Intermodal Transportation Network. | FHWA Office of Freight Manage-ment and Operations | | ○ | | Varies | | ○ | | ○ | ○ | | ○ | | | | | ○ | |
| FHWA Freight Performance Measures Initiative | The FPM initiative includes a data processing tool that estimates average operating speeds for trucks traveling on interstate highways. Speeds are calculated using confidential, onboard data from several hundred thousand trucks including periodic time, location, speed, and anonymous unique identification information. | FHWA Office of Freight Manage-ment and Operations | | ● | | 2-3 min intervals, available monthly | ● | | ● | | | | ● | | | ● | | | |

| Data Capture Program | Description | Data Owner | Data Type | | | Timelines | Availability | | Geographic Scope | | | Mode | | | | Data Source Hierarchy | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Infrastructure | Travel | Climate | Collection/Reporting Frequency | Existing | Future | National Data Program | Regional Data Program | Project Specific Data | Vehicle | Truck Freight | Transit | Bicycle/Pedestrian | Raw Data | Report | System | Repository |
| Modal (continued) | | | | | | | | | | | | | | | | | | | |
| Freight Analysis Framework (FAF) | The FAF integrates data from a variety of sources to create a comprehensive picture of freight movement among states and major metropolitan areas by all modes of transportation. With data from the 2007 Commodity Flow Survey and additional sources, FAF version 3 provides estimates for tonnage and value, by commodity type, mode, origin, and destination for 2007, the most recent year, and forecasts through 2040. Also included are truck flows assigned to the highway network for 2007 and 2040. | FHWA Office of Freight Management and Operations | | ● | | Annually | ● | | ● | | | | ● | | | | | ● | |
| Motor Carrier Management Information System (MCMIS) | MCMIS contains information on the safety fitness of commercial motor carriers (truck and bus) and hazardous material (HM) shippers subject to the Federal Motor Carrier Safety Regulations (FMCSR) and the Hazardous Materials Regulations (HMR). Components include carrier registration (census file) information, inspection data, and crash data for crashes involving a large truck or bus. | Federal Motor Carrier Safety Administration | | ● | | Registration – Varies; Inspection – 10 days; Crash – 90 days | ● | | ● | | | | ● | | | | | ● | |

| Data Capture Program | Description | Data Owner | Data Type | | | Timelines | Availability | | Geographic Scope | | | Mode | | | | Data Source Hierarchy | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Infrastructure | Travel | Climate | Collection/Reporting Frequency | Existing | Future | National Data Program | Regional Data Program | Project Specific Data | Vehicle | Truck Freight | Transit | Bicycle/Pedestrian | Raw Data | Report | System | Repository |
| Modal (continued) | | | | | | | | | | | | | | | | | | | |
| Licensing and Insurance (L&I) Database | For-hire and passenger carriers are required to have minimum levels of liability insurance, and this information is collected through the Licensing and Insurance (L&I) Database. | Federal Motor Carrier Safety Administration | | ● | | Biennial Updates | ● | | ● | | | | ● | | | | | ● | |
| International Freight Data System (IFDS) | The IFDS will examine transportation trade data and how long it takes goods to move between O-D pairs. It will provide a one stop shop for international trade data for all DOT modal administrations. | RITA Bureau of Transportation Statistics | | ○ | | TBD | | ○ | | | | | ○ | | | | | ○ | |
| Travel behavior | | | | | | | | | | | | | | | | | | | |
| National Household Travel Survey (NHTS) | The NHTS is the nation's inventory of daily and long-distance travel. The dataset allows analysis of daily travel by all modes, including characteristics of the people traveling, their household, and their vehicles. | RITA Bureau of Transportation Statistics | | ● | | 5 to 7 years | ● | | ● | | | ● | | ● | ● | ● | ● | | |

| Data Capture Program | Description | Data Owner | Data Type | | | Timelines | Availability | | Geographic Scope | | | Mode | | | | Data Source Hierarchy | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Infrastructure | Travel | Climate | Collection/Reporting Frequency | Existing | Future | National Data Program | Regional Data Program | Project Specific Data | Vehicle | Truck Freight | Transit | Bicycle/Pedestrian | Raw Data | Report | System | Repository |
| Travel behavior (continued) | | | | | | | | | | | | | | | | | | | |
| Omnibus Household Survey | The OHS is conducted to collect information about specific transportation issues (e.g., cell phones, distracted driving) and to assess public satisfaction and attitudes about the transportation system and its interaction with DOT agencies (e.g., airline passenger opinions on security screening procedures). | Transportation Research Board | | ● | | Annually | ● | | ● | | | ● | | ● | ● | ● | ● | | |
| Other Data Programs | | | | | | | | | | | | | | | | | | | |
| Data.gov | The Data.gov is a data repository for Federal datasets and tools. It contains searchable catalogs that provide access to raw datasets and various tools from different Federal agencies. | All stakeholders | ● | ● | ● | Varies | ● | | ● | | | ● | ● | ● | ● | | | | ● |

| Data Capture Program | Description | Data Owner | Data Type | | | Timelines | Availability | | Geographic Scope | | | Mode | | | | Data Source Hierarchy | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Infrastructure | Travel | Climate | Collection/Reporting Frequency | Existing | Future | National Data Program | Regional Data Program | Project Specific Data | Vehicle | Truck Freight | Transit | Bicycle/Pedestrian | Raw Data | Report | System | Repository |
| Other Data Programs (continued) | | | | | | | | | | | | | | | | | | | |
| Integrated Corridor Manage-ment (ICM) Program | The U.S. DOT is partnering with eight "Pioneer Sites" in a 5-year initiative to develop, deploy and evaluate ICM concepts in eight of the nation's busiest corridors. The ICM Initiative aims to advance the state of the practice in transportation corridor operations to manage congestion. This initiative will provide the institutional guidance, operational capabilities, ITS technology and technical methods needed for effective ICM systems. | RITA ITS Joint Program Office | | ● | | Varies | ● | | | ● | ● | ● | | ● | | ● | | | |
| I-95 Corridor Coalition Data | The I-95 Corridor Coalition Vehicle probe project acquire travel times and speeds on freeway and arterials using probe technology to provide regional management of traffic and traveler information useful to long-distance travelers. | I-95 Corridor Coalition | | ● | | 5 min | ● | | | ● | ● | ● | | | | | | | ● |

| Data Capture Program | Description | Data Owner | Data Type | | | Timelines | Availability | | Geographic Scope | | | Mode | | | | Data Source Hierarchy | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Infrastructure | Travel | Climate | Collection/Reporting Frequency | Existing | Future | National Data Program | Regional Data Program | Project Specific Data | Vehicle | Truck Freight | Transit | Bicycle/Pedestrian | Raw Data | Report | System | Repository |
| Other Data Programs (continued) | | | | | | | | | | | | | | | | | | | |
| VII Data Use Analysis and Processing (DUAP) Project | The DUAP project is evaluating how new connected vehicle data sources may be combined with other data sources, and how use of these new data sources may impact MDOT's current transportation system operations through improved methods and processes. | Michigan DOT | | ● | | Varies based on data type. Typ. 4 to 20 seconds | ● | | | ● | ● | ● | | | | ● | | | |
| Real-Time Data Capture and Management Program | | | | | | | | | | | | | | | | | | | |
| TFHRC Data Resources Testbed | The testbed will host real-time and archived transportation data from a variety of sources to support performance measurement and transportation system management applications. | Turner-Fairbank Highway Research Center | | ○ | | TBD | | ○ | | ○ | | ○ | | | | ○ | | | |

| Data Capture Program | Description | Data Owner | Data Type | | | Timelines | Availability | | Geographic Scope | | | Mode | | | | Data Source Hierarchy | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Infrastructure | Travel | Climate | Collection/Reporting Frequency | Existing | Future | National Data Program | Regional Data Program | Project Specific Data | Vehicle | Truck Freight | Transit | Bicycle/Pedestrian | Raw Data | Report | System | Repository |
| Real-Time Data Capture and Management Program (continued) | | | | | | | | | | | | | | | | | | | |
| DCM Program Research Data Exchange | The RDE is a new decentralized, Internet-based, federated system of systems. The RDE serves as a research repository for the Real-Time Data Capture and Management Program, and it contains data environments (collections of related data sets and possibly real-time data feeds), individual data sets, and real-time data feeds. | RITA ITS Joint Program Office | ● | ● | ● | Varies based on data environment, data set, and data feeds | ● | | | ● | ● | ● | | | | | | | ● |
| DCM Program Test Data Sets | Well-documented, quality test data sets are available from recent or ongoing operations, field tests, or simulations of emerging technologies support mobility, environment, transit, freight, weather, and other surface transportation research. | RITA ITS Joint Program Office | ● | ● | ● | Varies for each test data set | ● | | | ● | ● | ● | ● | ● | | ● | | | |
| Dynamic Mobility Application Datasets | Well-organized data and associated descriptors (data environments) that include data from field tests and advanced simulation models in support of high priority arterial, freeway, regional, and corridor applications. | RITA ITS Joint Program Office | ○ | ○ | | TBD | | ○ | | ○ | | ○ | ○ | ○ | ○ | ○ | | | |

**Legend:** ● Existing Data Program ○ Future Data Program

## Data Source Hierarchies

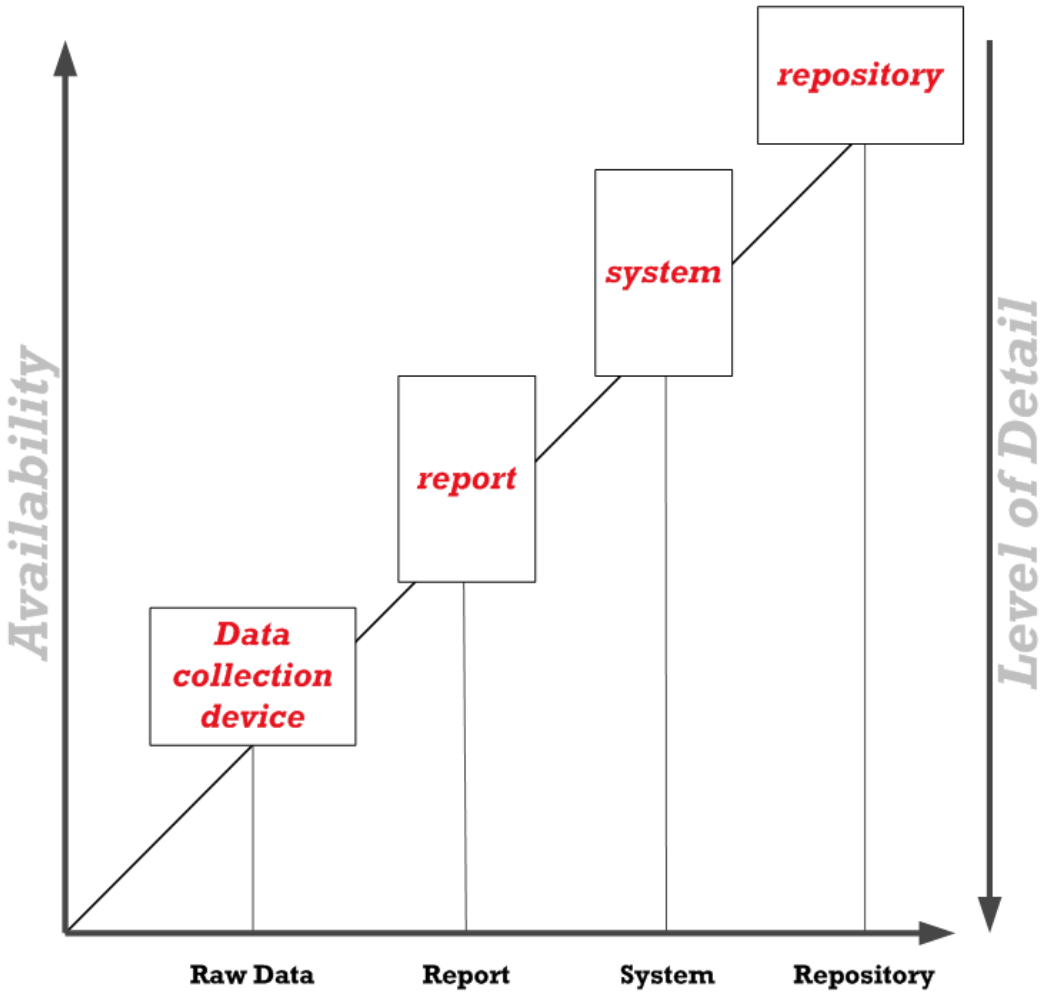
The white paper, *Data Capture and Management: Needs and Gaps in the Operation and Coordination of U.S. DOT Data Capture and Management Programs*, introduced the concept of a data source hierarchy. This concept recognizes that the data programs included in the data inventory are not all fundamental sources of data; rather, some are a compilation of data from other sources. It is important to realize that the sources of data are often the same, but the data are often aggregated and reported differently based on the end user needs. The data hierarchy identifies four levels of detail:

1. **Raw Data**. This is raw data that is acquired directly from the source through a data collection device such as a sensor, automatic traffic data recorder (ATR), or road tube. Examples include traffic volumes acquired directly from road tubes, weather data acquired directly from weather gauges, or roadway/asset inventory data acquired directly from a field inventory.
2. **Report.** This is raw data that has been quality checked, aggregated, and then generated into a report. Examples include Traffic Volume Trend (TVT) reports and Urban Congestion Reports.
3. **System.** A system that facilitates the processing and reporting of data to meet State or Federal reporting requirements as needed. Examples include the Travel Monitoring and Analysis System.
4. **Repository.** An information retrieval system that stores and provides access to data. Repositories also seek to preserve and promote the use of data. They may be in the form of a digital library, data warehouse, or website.

The four levels of the hierarchy differ in terms of availability and level of detail. In terms of availability, access to a raw data source is typically limited to selected users, while a repository is widely availability to a broad base of users. In terms of level of detail, the raw data source typically provides the highest level of detail in the data itself, while a repository may only provide aggregated or rolled up data. Figure  provides an overview of a basic data source hierarchy.

This section presents data source hierarchies for each of the data types – infrastructure, travel, and climate data. It is important to distinguish between these levels of data sources because certain coordination mechanisms will relate only to certain levels. In addition, the levels will help to identify true data sources and assist in prioritizing action steps to eliminate gaps.

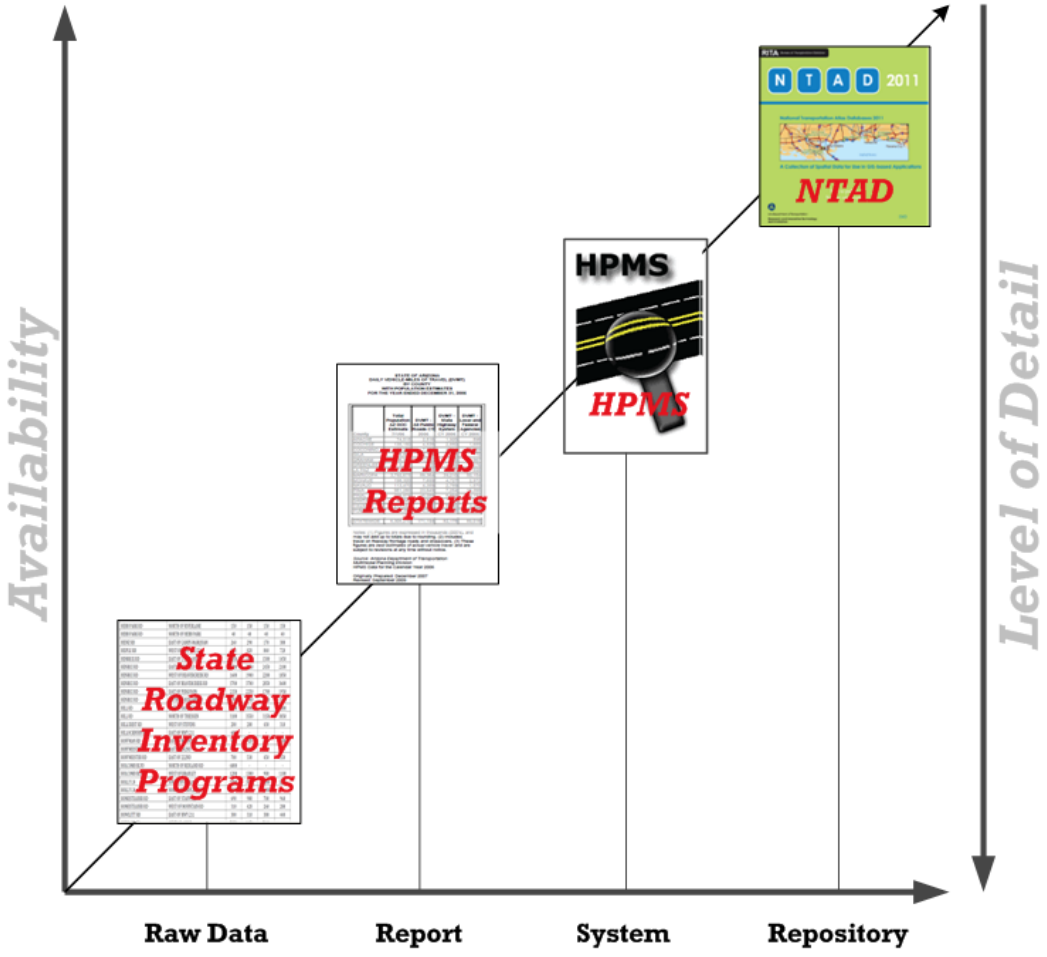
**Figure H.1 Basic Data Source Hierarchy** (Source: Cambridge Systematics, Inc., June 2012)



***Infrastructure Data***

Figure H.2 shows the data source hierarchy for infrastructure data. Raw roadway/asset inventory data are collected at the source through a field inventory conducted as part of a State Roadway Inventory Program. The data is compiled at the State level using the Highway Performance Monitoring System (HPMS). Infrastructure data are made available through the National Transportation Atlas Database (NTAD), which serves as a repository.

**Figure H.2 Data Hierarchy for Infrastructure Data** (Source: Cambridge Systematics, Inc., June 2012)



### *Travel Data*

Figure H.3 shows the data source hierarchy for volume data for planning purposes. Raw traffic volume data is collected at the source with a traffic data collection device such as an automatic traffic data recorder (ATR) or road tube at designated State traffic monitoring stations. It is compiled at both the State and national level into traffic volume reports. These are reported monthly to FHWA and published as Traffic Volume Trend (TVT) reports. Those reports are then used to develop the Highway Performance Monitoring System (HPMS). Results of HPMS are reported to FHWA through the Travel Monitoring and Analysis System (TMAS).

**Figure H.3 Data Hierarchy for Volume Data – Office of Highway Policy Information** (Source: Cambridge Systematics, Inc., June 2012)

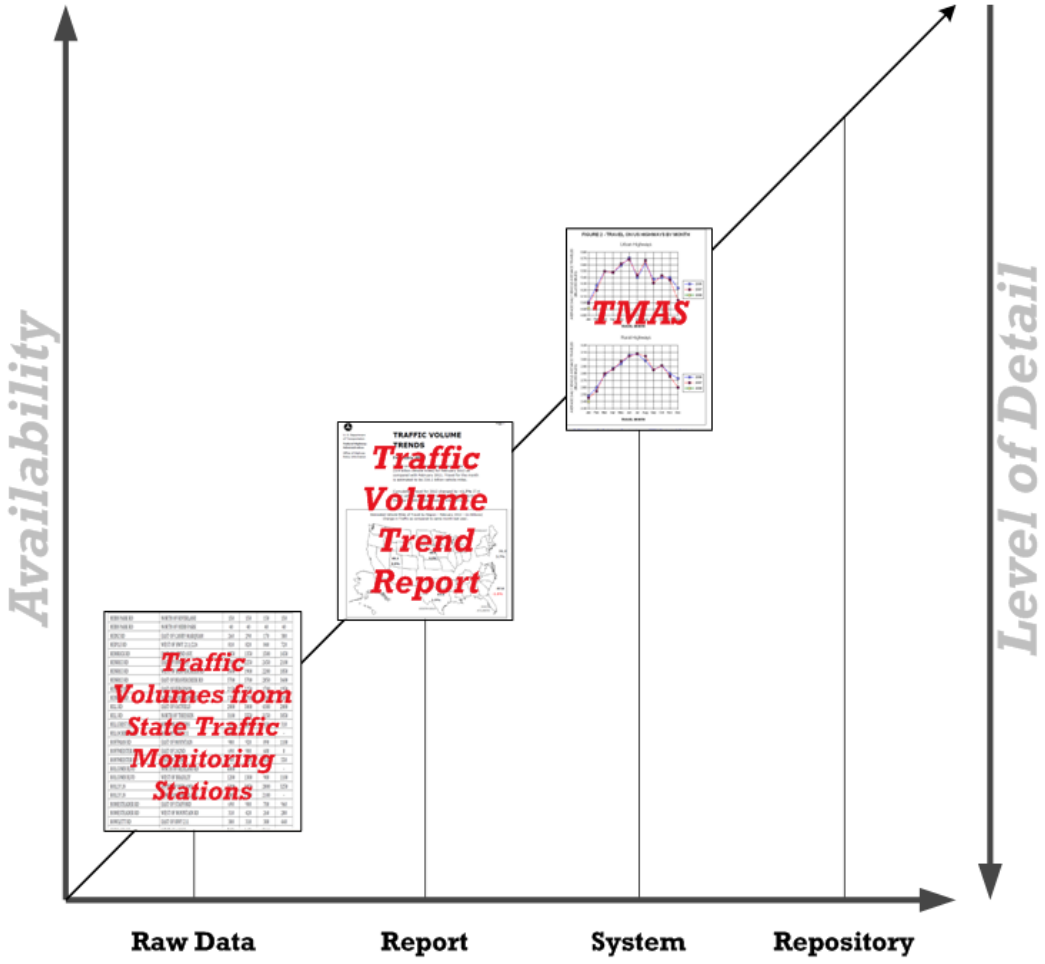


Figure H.4 shows the data source hierarchy for speed data for planning purposes. Raw speed data is collected at the source with a traffic data collection device such as an automatic traffic data recorder (ATR) at designated State traffic monitoring stations. Speed data is a requirement of the Traffic Monitoring Guide, and the data is reported to FHWA through the Travel Monitoring and Analysis System (TMAS).

**Figure H.4 Data Hierarchy for Speed Data – Office of Highway Policy Information** (Cambridge Systematics, Inc., June 2012)

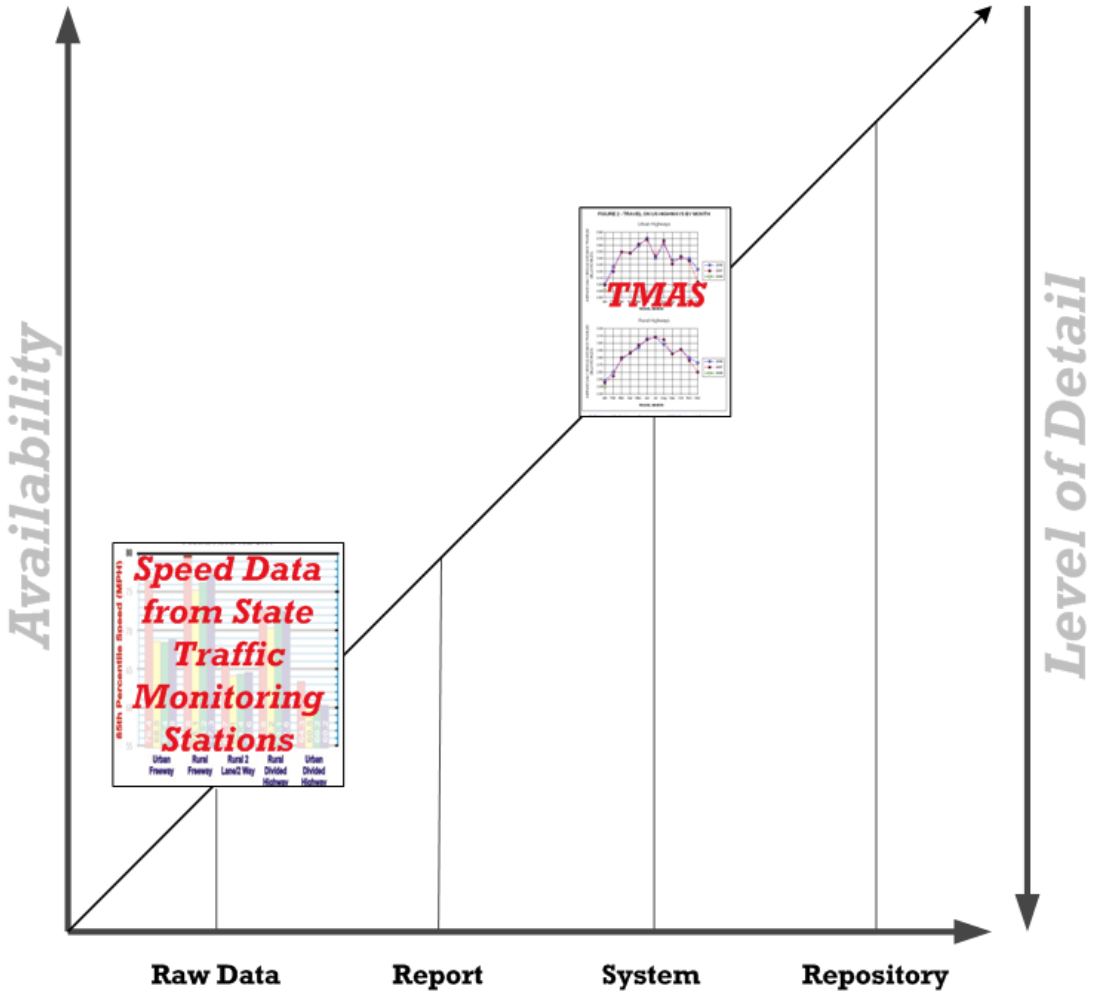


Figure H.5 shows the data source hierarchy for volume and speed data for operations purposes. Raw traffic volume, speed, and lane occupancy data is collected at the source through ITS data collection technologies (e.g., radar detection, in-pavement loops, magnetic detectors), probe vehicle based methods (e.g., toll tag readers, Bluetooth device matching, wireless location, crowd-sourcing), or private data sources (e.g., Navteq, Inrix). It is compiled at both the State and national level into Urban Congestion Reports. Some archived traffic data (such as Transportation Technology Innovation and Demonstration (TTID)/Traffic.com program data) is made available for FHWA research through the Intelligent Transportation Infrastructure Program (ITIP) Stakeholder Applications Data Warehouse.

**Figure H.5 Data Hierarchy for Volume and Speed Data – Office of Operations** (Source: Cambridge Systematics, Inc., June 2012)

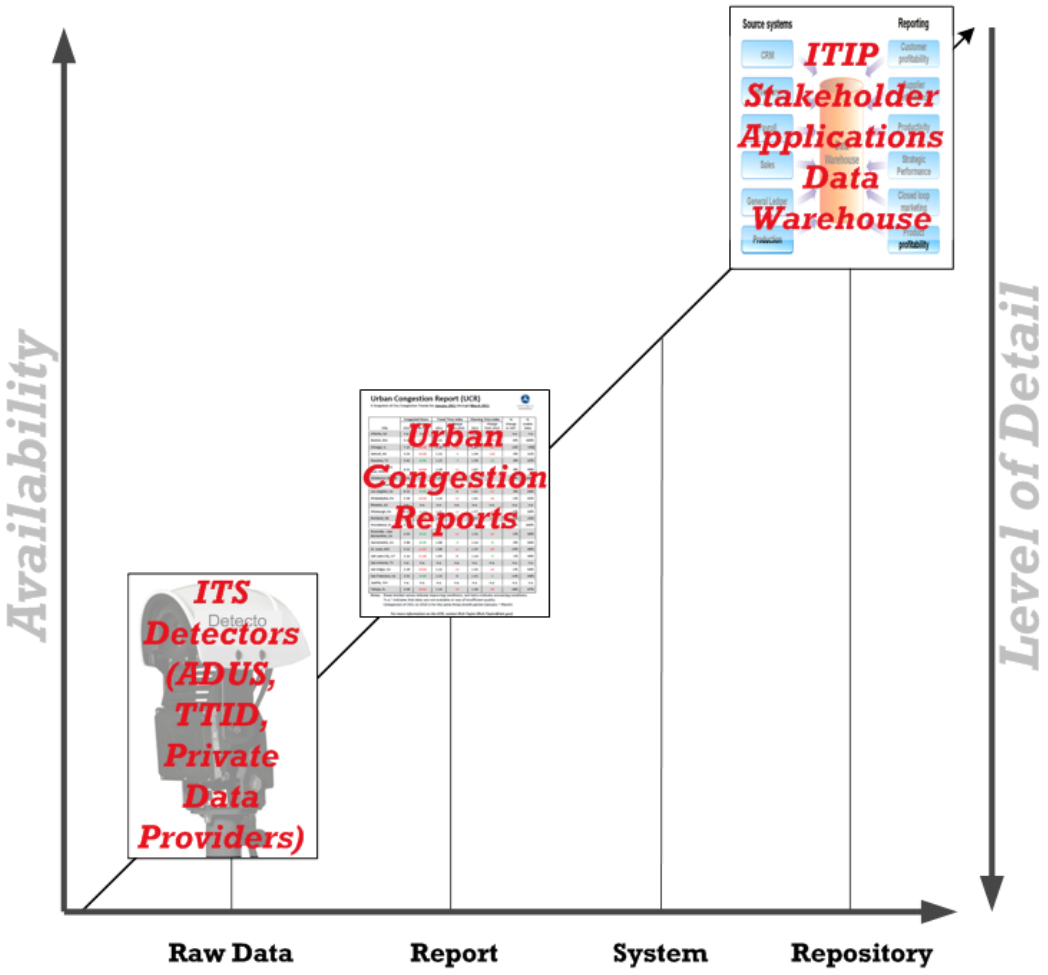
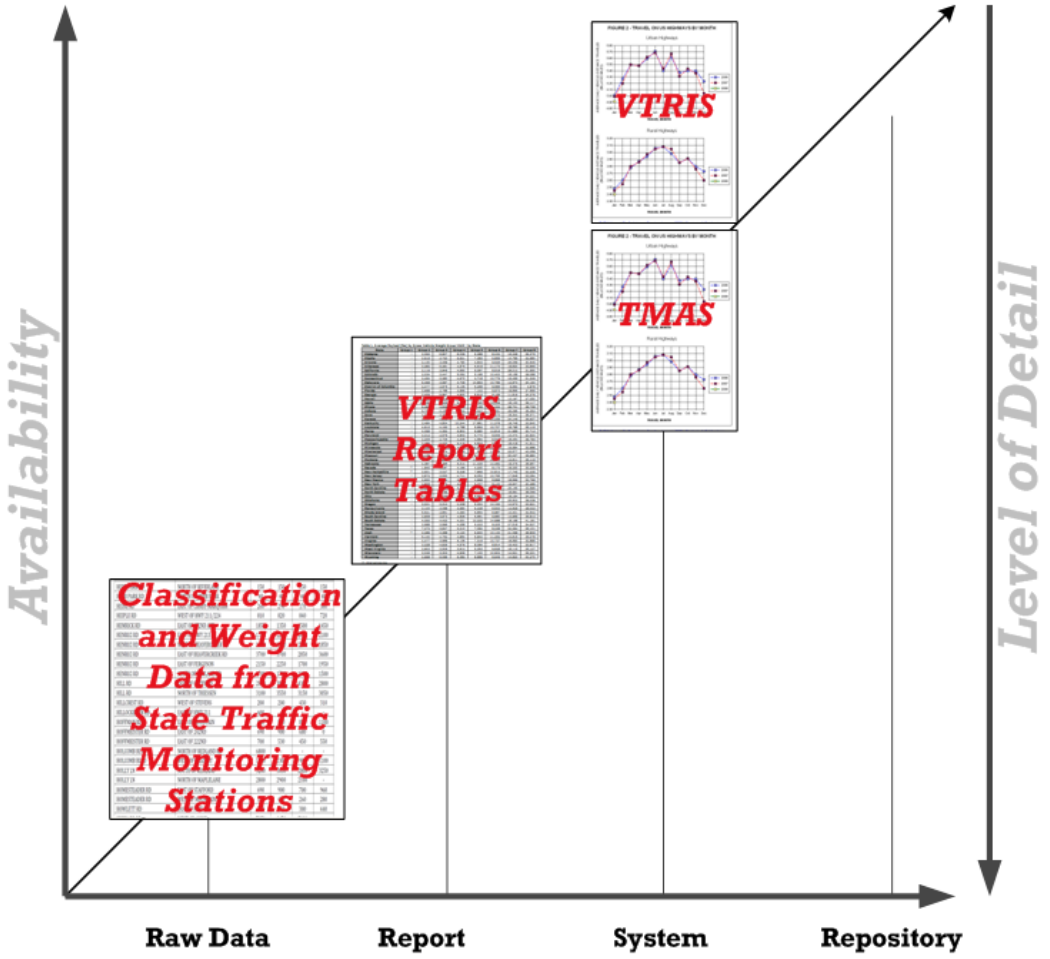


Figure H.6 shows the data source hierarchy for vehicle classification and weight data. Raw vehicle classification and weight data is collected at the source with a traffic data collection device such as an automatic traffic data recorder (ATR) or weigh-in-motion device at designated State traffic monitoring stations. It is compiled at the State level into the Vehicle Travel Information System (VTRIS), which produces summary report tables. Results of HPMS are reported to FHWA through either VTRIS or the Travel Monitoring and Analysis System (TMAS).

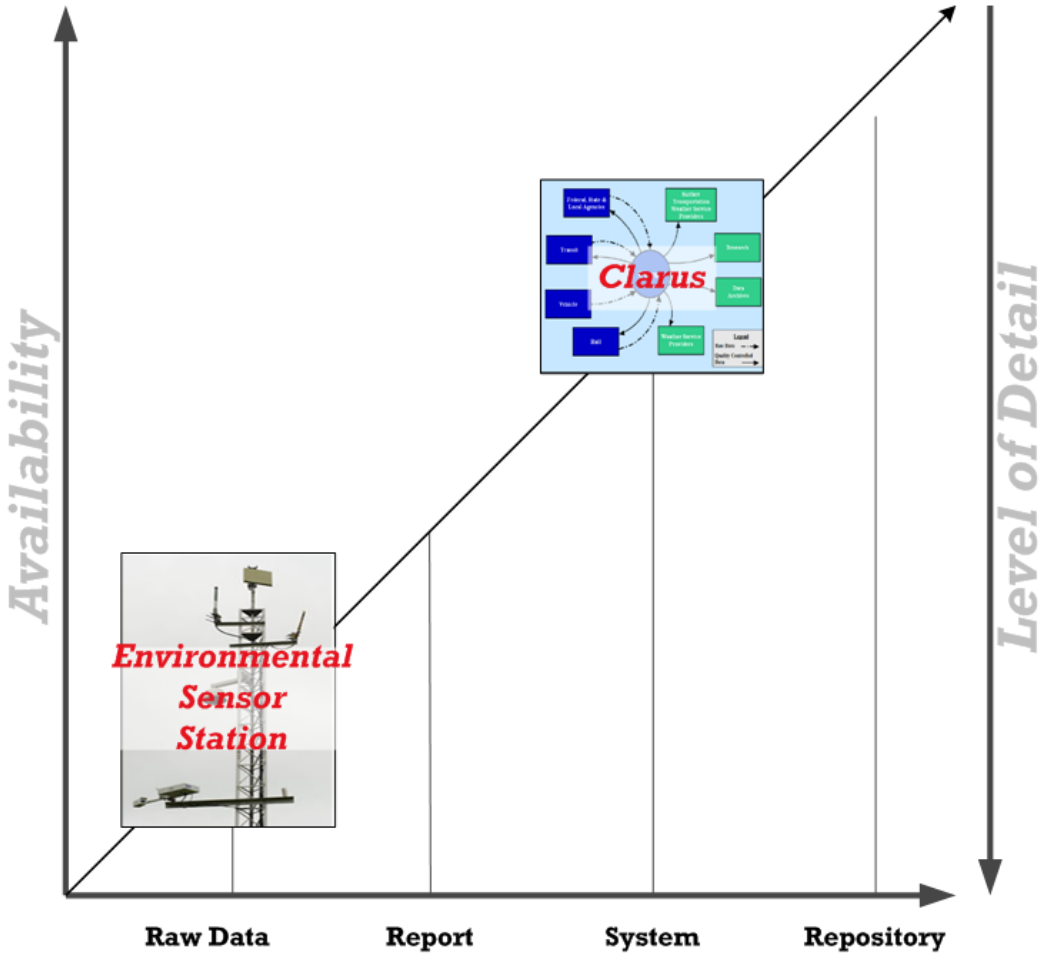
**Figure H.6 Data Hierarchy for Vehicle Classification and Weight Data** (Source: Cambridge Systematics, Inc., June 2012)



### *Climate Data*

Figure H.7 shows the data source hierarchy for climate data. Raw climate and road weather data is collected at the roadway through environmental sensor stations (ESS). The sensors connect to a remote processing unit adjacent to the road, which transmits the information to a central server. The road weather data is compiled into road weather reports, which are used to support maintenance activities in a State. The data is also submitted to Clarus, which facilitates assimilation, quality checking, and dissemination of the data.

**Figure H.7 Data Hierarchy for Climate Data** (Source: Cambridge Systematics, Inc., June 2012)



# Detailed Inventory – Infrastructure Data

***State Roadway Inventory Programs***

**Data Owner:** State DOTs

**Description:** State roadway inventory programs document the extent (length), point and linear feature data of public roadways in each State, along with roadway attributes covering roadway classification, ownership, physical conditions, traffic volumes, pavement conditions, highway performance monitoring information, and more. This data is collected and stored for both on-state system and off-state system roadways. Much more information is traditionally maintained for the on-state system, since the State DOTs are responsible for these roads. In some States, the State DOT responsibilities also include the collection of data on lower functionally classified roads that may be considered as off-state system.

**Purpose:** The purpose of roadway inventory programs is to support future planning, design, construction, and maintenance of a State's roadway network. This information is used to determine the need for future investment in the State's transportation infrastructure. The data is also used for transportation modeling (travel demand and pavement deterioration) at the State and national level and is used for Federal reporting purposes with the Highway Performance Monitoring System and the State's Highway Safety Program. It is also used to support the statewide strategic transportation plan and the transportation plans developed by the Metropolitan Planning Organizations (MPOs) and Councils of Government (COGs).

**Data Source Hierarchy:** The State roadway inventory programs consist of point and linear features along the public road network in each State. They are classified as a raw data source in the data hierarchy.

**Geographic Scope/Coverage:** The roadway inventory programs document the public road network in the States. Exceptions may include roads located within military facilities, Indian Nations, and other areas where roadway data collection is traditionally not allowed.

**Format:** Each State maintains their road inventory data in a variety of formats. Several States use geo-databases or relational databases as the preferred repository. This coincides with the recent implementation of geo-spatial databases at the Federal level for the Highway Performance Monitoring System (HPMS). The business intelligence (BI) tools available in a Geographic Information System (GIS) environment provide easier methods for maintenance of a State's road inventory network.

**System Requirements:** System requirements vary from state to state; however, the road inventory data is usually not 'real-time' data. In most States, it takes some time to QC the data and to make it ready for publication. In addition, the reporting requirements for HPMS include documenting the State's road network as of December 31 of the previous calendar year. As such, States maintain an end-of-year road inventory database, as well as a current database for use in reporting with HPMS and other State or Federal reports as needed.

**Data Quality:** Each State is responsible for the quality control and validation of the data in their road inventory programs. This QC process can be facilitated through the use of tools such as GIS.

**Metadata:** Most States will maintain metadata for their GIS data, which may include road inventory data, but they do not usually maintain metadata specifically for road inventory databases.

**Timeliness:** Again, due to the requirements for monthly and annual reports to FHWA, the road inventory data is usually available for use on a monthly and annual basis. The timeliness of the data depends upon how many offices are responsible for the collection, processing, and distribution of the data. Usually, the data is collected by the individual districts of a State DOT and then aggregated and distributed at a central office.

**Longevity/Future Plans:** States continue to enhance or in some cases revamp their existing road inventory databases to meet future needs and take advantage of the opportunities and efficiencies offered by new technology.

**Users:** Users may include public and private entities at the Federal, State, and local levels.

**Accessibility:** The State DOT determines how and what data is provided internally and externally. This is usually through a web interface, or FTP server if large quantities of data are requested.

**Contact:** Varies with State, but several State DOTs have a data management office responsible for maintaining the road inventory data. This may be the responsibility of the Planning Division, the Asset Management Office or other offices, depending upon the organizational structure of the DOT.

***Highway Performance Monitoring System (HPMS)***

**Data Owner:** FHWA Office of Highway Policy Information

**Description:** The Highway Performance Monitoring System (HPMS) is a national level highway information system that includes data on the extent, condition, performance, use, and operating characteristics of the nation's highways. The HPMS data are collected by the 50 States, the District of Columbia and U.S. territories. Some States also utilize HPMS data collected by the MPOs, and other jurisdictions. The data is reported to FHWA annually. The data collected include roadway inventory, traffic volumes and some limited classification data, geometric, pavement, route, and special network data items.

HPMS provides both infrastructure and traffic data. For infrastructure, HPMS provide limited data for all highway sections within its scope (e.g., number of lanes) and more detailed data for a sample of highway sections (e.g., measures of pavement condition).

**Purpose:** The Federal Highway Administration (FHWA) has the responsibility to assure that adequate highway transportation information is available to support its functions and responsibilities, including those of the Administration and the Congress. The primary purpose of the Highway Performance Monitoring System (HPMS) is to serve these data and information needs.

HPMS data are used extensively in the analysis of highway system condition, performance, and investment needs that make up the biennial Condition and Performance Reports to Congress. These Reports are used by the Congress in establishing both authorization and appropriation legislations, activities that ultimately determine the scope and size of the Federal-aid Highway Program, and determine the level of Federal highway taxation. The data are also used for assessing changes in highway system performance brought about by implementing funded highway system improvement

programs under the GPRA, and for apportioning Federal-aid Highway Funds to individual States under TEA-21.

**Data Source Hierarchy:** The HPMS is an integrated database that was developed in 1978 as a national highway transportation system database. In the data source hierarchy, HPMS is classified as a system.

**Geographic Scope/Coverage:** The HPMS is a nationwide inventory system that includes data for all of the nation's public road mileage as certified by the States' Governors on an annual basis. This includes facilities both on and off State-owned highway systems. Each State is required to furnish annually all data requirements specified in the HPMS Field Manual. The District of Columbia and the Commonwealth of Puerto Rico are considered to be States for HPMS reporting purposes. United States Territories (Guam, the Commonwealth of the Northern Marianas, American Samoa, and the Virgin Islands) are required to report limited HPMS summary data on an annual basis, in addition to public road mileage certifications.[52]

**Format:** HPMS includes limited data on all public roads, more detailed data for a sample of the arterial and collector functional systems, and area wide summary information for urbanized, small urban and rural areas. The HPMS also requires the reporting of supplemental air quality non-attainment area sample data and LRS data for FHWA use in a geographic information system. The following types of data are submitted by States:

- The **statewide summary data** includes information on travel, system length, and vehicle classification by functional system and area type, in addition to land area and population by area type. The area types include rural, small urban, individual urbanized, non-attainment and maintenance areas.
- The term **full extent data** refers to a limited set of data items reported for the entire public road system as individual or grouped length sections.
- **HPMS sample panel data** consists of data items added to the full extent data that are reported for a small portion of the total highway system length. The sampled sections nominally are a fixed sample panel of highway sections that are monitored from year to year and, when expanded, represent the full extent of the systems that are sampled. The more detailed information collected for a sample panel section is used to represent similar conditions on the associated functional system after expansion.

> *A standard sample panel contains the full extent data plus additional data items related to the physical characteristics, condition, performance, use, and operation of the sampled panel sections of a highway. These sample panel data provide detailed information, which is used as the basis for evaluating changes over time, and provides the basic input to the HPMS simulation models (Analytical Process (AP) and Highway Economic Requirements System (HERS)).*

[52] Highway Performance Monitoring System Field Manual, FHWA Office of Highway Policy Information, September 2010, http://www.fhwa.dot.gov/policyinformation/hpms/fieldmanual/.

- The **HPMS LRS data** provides a linear referencing system (LRS) for the full extent and sample data on selected highway functional systems. The represented functional systems include urban and rural principal arterials, rural minor arterials, and all National Highway System (NHS) routes and connectors. This permits the analyses of HPMS data in a geographic information system (GIS) environment.

Roadway information data is submitted as a character separated value (CSV) file in the Section dataset, while LRS data is submitted as a complete GIS file encompassing the base highway network.

**System Requirements:** FHWA developed the HPMS Software System (Version 7.0) to facilitate the submittal and analysis of HPMS data in a microcomputer environment, using a Geodatabase model. The system provides States with the ability to enter data directly into the system, view section-by-section data on the screen, modify section-by-section data, validate data, calculate data values, query data for inconsistencies in coding, and analyze current year data and compare them with prior year's data. The software is used by States to submit the HPMS data annually to FHWA. The data is submitted electronically through the User Profile and Access Control System (UPACS).

The software should be installed on at least a Pentium or Pentium II (preferred) 200 Mhz with MMX extensions or AMD K6 with 3D Now with at least 64 MB of memory (96 MB RAM or 128 MB RAM preferred) and at least a 4 gigabyte hard drive running Windows NT or Windows 98 or better. The resolution of the monitor should be at least 800 x 600 True Colors. The software requires 20 MB of free space on C:\ for installation and approximately five times the size of the current HPMS data file free space for running the system.

HPMS LRS data require use of a geographic information system (GIS).

**Data Quality:** The States submit a Metadata and an Estimates dataset to document any issues or comments related to the reporting of the State's pavement and traffic data.

**Metadata:** See comment above regarding metadata reporting requirements for HPMS.

**Timeliness:** All HPMS data are to represent conditions as of December 31 of the previous calendar year, referred to as the data year. HPMS data is submitted to FHWA annually (due June 15 of the year following the data year).

**Longevity/Future Plans:** The HPMS system undergoes reassessment at various times, due to changes in priorities at the Federal level for information on the nation's transportation system. The most recent reassessment (referred to as HPMS Reassessment 2010+[53]) included a complete redesign of the Federal HPMS data model to utilize a Geographic Information System (GIS) environment. The first year for reporting data under the new HPMS requirements was June 2011 for the 2010 data year. Most States are already modifying existing DOT databases to facilitate the reporting of HPMS using the GIS environment.

**Users:** The HPMS mileage and travel estimates are used for apportionment of Federal-aid highway funds. HPMS data are also used for assessing and reporting highway system performance under FHWA's strategic planning process. HPMS data form the basis of the analyses that support the

[53] HPMS Reassessment 2010+, FHWA Office of Highway Policy Information, September 2008, http://www.fhwa.dot.gov/policyinformation/hpms/2010/hpms2010.pdf.

Conditions and Performance Reports to Congress and are the source for a substantial portion of the information published in *Highway Statistics* and in other FHWA publications and media. Finally, the HPMS data are widely used throughout the transportation community, including other governmental interests, business and industry, institutions of higher learning, the media and public.

**Accessibility:** Access to HPMS Software Version 7.0 is restricted to authorized users. A UPACS account must be established and permission granted through the FHWA Field Division in each State. Some HPMS data for States and MPOs are publicly accessible through the HPMS website.[54]

**Contact:** Ralph Gillmann, HPMS Program Team Leader, 202-366-5042

***National Transportation Atlas Databases***

**Data Owner:** RITA Bureau of Transportation Statistics (BTS)

**Description:** The National Transportation Atlas Databases (NTAD) is a set of nationwide geographic databases of transportation facilities, transportation networks, and associated infrastructure. The datasets include spatial information for transportation modal networks and intermodal terminals, as well as related attribute information for these features. Metadata documentation based on the standards of the Federal Geographic Data Committee is also provided for each database. This is the only multimodal infrastructure database within the Department.

NTAD is compiled and owned by BTS, and they obtain the data from different sources within U.S. DOT (e.g., FHWA, the Federal Railroad Administration, and the Office of Airline Information), the U.S. Army Corps of Engineers, the U.S. Bureau of the Census, the Environmental Protection Agency, the Surface Deployment and Distribution Command – Transportation Engineering Agency, and the National Park Service.

**Purpose:** To support research, analysis, and decision-making across all modes of transportation. Datasets are most useful at the national level, but have major applications at regional, State, and local scales throughout the transportation community.

**Data Source Hierarchy:** NTAD is compiled by BTS from a variety of sources and is classified as a repository.

**Geographic Scope/Coverage:** The datasets include national coverage of the following:

- **Point-based data:** public-use airports, alternative fuel stations, Amtrak stations, automatic traffic recorder stations, U.S. border crossings, highway-rail grade crossings, intermodal terminal facilities, National Bridge Inventory, national populated places, railroad bridges, U.S. Army Corps of Engineers ports, and weigh in motion stations;
- **Polyline data:** Freight Analysis Network, hazardous material routes, Highway Performance Monitoring System, National Highway Planning Network, railway

[54] FHWA Office of Highway Policy Information Highway Performance Monitoring System (HPMS) website, http://www.fhwa.dot.gov/policyinformation/hpms.cfm.

network, public use airport runways, fixed-guideway transit facilities, and U.S. Army Corps of Engineers Navigable Waterway Network; and

- **Polygon data:** Bureau of the Census urbanized area boundaries, the 112th Congressional Districts boundaries, U.S. county boundaries, U.S. county political boundaries, hydrographic features, U.S. military installations, metropolitan planning organizations; core based statistical areas, non-attainment areas, National Park System boundary dataset, U.S. State boundaries, and U.S. State political boundaries.

**Format:** NTAD is distributed in shapefile format. Each shapefile dataset is composed of nine different files:

1. Main file (.shp) – feature geometry;
2. Index file (.shx) – indices of the feature geometry;
3. dBASE file (.dbf) – attribute information for the features;
4. Projection file (.prj) – projection and datum information for the dataset;
5. Spatial index file (.sbx) – spatial index for read-write of shapefiles;
6. Spatial index file (.sbn) – spatial index for read-write of shapefiles;
7. Metadata file (.xml) – xml encoding of shapefile's metadata;
8. Metadata file (.htm) – html encoding of shapefile's metadata; and
9. Metadata file (.txt) – text version of shapefile's metadata.

**System Requirements:** The NTAD databases are designed for use within a geographic information system (GIS); however, the attribute data for each dataset is stored in dBASE format and can be accessed using any database, spreadsheet, or other software package. However, many of the larger dBASE files will not open correctly with spreadsheet software due to spreadsheet limitations.

**Data Quality:** The source agencies are the data stewards and are ultimately responsible for the maintenance and accuracy of the data. BTS performs structured quality evaluations on the data before they are included in NTAD. Further details pertaining to the NTAD compilation process may be found in the metadata file associated with each dataset.

**Metadata:** A data dictionary shows what datasets are included in NTAD. Metadata is provided for each dataset and includes identification information, data quality information, spatial data organization information, spatial reference information, entity and attribute information, distribution information, and metadata reference information.

**Timeliness:** NTAD is updated on an annual basis. Data are available for 1996, 1997, and then yearly from 2000 through 2009.

**Longevity/Future Plans:** NTAD is a primary data product of BTS and will continue to be updated on an annual basis.

**Users:** Customers for NTAD include Federal, State and local agencies with a GIS program, researchers, geographers, and members of Congress who request maps of their districts. For example, the American Recovery and Reinvestment Act (ARRA) is using NTAD to identify project locations and condition of assets.

**Accessibility:** The NTAD databases are accessible to the public and are available for download through the BTS website.[55] Users can also request an NTAD DVD through the BTS bookstore.[56]

**Contact:** Dave Smallen, 202-366-5568; Mark Bradford, 202-366-6810

## Detailed Inventory – Travel Data

***Archived Data User Service (ADUS)***

**Data Owner:** FHWA Office of Highway Policy Information

**Description:** The Archived Data User Service (ADUS) provides the National ITS Architecture with the requirements for archiving and re-use of data collected for ITS operations. It enables transportation agencies to retain ITS-generated data and make them available for analysis.

**Purpose:** The ITS Joint Program Office, the Office of Highway Policy Information, and other U.S. DOT offices sponsor a development track, which focuses on increasing the awareness of and professional capacity to implement ADUS, advancing and expanding the application of the state-of-the-practice of ADUS, and promoting the use of archived data to make better transportation decisions. In addition, the ADUS program strives to streamline Federal reporting systems and performance monitoring.

**Data Source Hierarchy:** ADUS enables the archiving and re-use of traffic volume, lane occupancy, and speed data. It is classified as a raw data source in the data source hierarchy.

**Geographic Scope/Coverage:** Geographic coverage varies by region, but typically includes major freeways that have ITS infrastructure in place.

**Format:** Online data programs allow the exportation of data in different formats (tables, graphs). They generally provide processed, user-friendly data, rather than raw data.

**System Requirements:** Archived Data Online programs allow the extraction of data in formats that are compatible with most PC interfaces (such as .xls format).

**Data Quality:** Maintenance of data quality is performed through mechanisms such as retaining original source data, correcting data at the source, and constructing an audit trail.

**Metadata:** Metadata documents accompany several of the Archived Data On-line Programs.

**Timeliness:** The timeliness of online data programs varies by region. On-line data programs provide archived data of different granularities up to the present year.

**Longevity/Future Plans:** Archived data online programs will continue to be updated on a continuous basis.

**Users:** Archived data on-line programs are generally open to registered users.

[55] RITA BTS National Transportation Atlas Database website, http://www.bts.gov/publications/national_transportation_atlas_database/.

[56] RITA BTS Bookstore website, https://1bts.rita.dot.gov/pdc/user/products/ src/products.xml?p=3389.

**Accessibility:** Documents for ADUS program activities, standards and resources are available to the public and can be accessed through the FHWA ADUS website.[57] In addition, the website provides links to the following archived data online programs:

- California Freeway Performance Measurement Project (PeMS)
- Maryland - Center for Advanced Transportation Technology Laboratory
- Minnesota - Traffic Management Center
- Texas - Houston Transtar Historical Information
- PORTAL: Portland Oregon Regional Transportation Archive Listing
- Virginia - ADMS Virginia, Smart Travel Lab
- Cubeview - Virginia Tech
- Washington Traffic Data Acquisition and Distribution (TDAD)

**Contact:** Tianjia Tang, 202-366-0180; James Pol, ITS Joint Program Office, 202-366-4374, james.pol@fhwa.dot.gov

***Transportation Technology Innovation and Demonstration (TTID)/Traffic.com***

**Data Owner:** FHWA Office of Operations

**Description:** FHWA's Transportation Technology Innovation and Demonstration (TTID) Program funds travel time data collection by Navteq (formerly Traffic.com) in several metropolitan areas. The purpose of this program is to address national, local, and commercial data needs through enhanced surveillance and data management in major metropolitan areas. This involves integration of data from existing surveillance infrastructure and strategic deployment of supplemental surveillance infrastructure to provide real-time and archived roadway system performance data.

The TTID Program is the primary source of loop detector data on many major interstate highways, particularly in smaller metropolitan areas that do not have an ITS infrastructure in place. Polled data is converted into travel information (e.g., congestion levels, speed, etc.), and the data is then shared with local traffic management centers. Sensor data from sensors installed in all of the metropolitan areas are made available to support FHWA research through Intelligent Transportation Infrastructure Program (ITIP) Stakeholder Applications Data Warehouse.

**Purpose:** The purpose of TTID is to share data for real-time operations and planning. At the national level, the goal is to measure the operating performance of the roadway system across the nation. Locally, such roadway system performance data can be used to assist in local system planning, evaluation, and management activities. The same data that is useful to the public transportation agencies also has value for commercial traveler information purposes.

**Data Source Hierarchy:** The TTID program is a source of raw speed and travel time data, while the ITIP Data Warehouse is a repository for archiving and accessing the data.

[57] FHWA Office of Highway Policy Information Archived Data User Service website, http://www.fhwa.dot.gov/policy/ohpi/travel/adus.cfm.

**Geographic Scope/Coverage:** The following 27 cities are part of the programs under which Traffic.com provides traffic data services for FHWA: Atlanta, Baltimore, Boston, Chicago, Cincinnati, Columbus, Detroit, Indianapolis, Las Vegas, Los Angeles, New Orleans, Norfolk/Hampton Road, Oklahoma City, Philadelphia, Phoenix, Pittsburgh, Providence, Raleigh/Durham, Sacramento, Salt Lake City, San Diego, San Francisco, San Jose, Seattle, St. Louis, Tampa, and Washington D.C. Two cities, Philadelphia and Pittsburgh were part of the ITIP Federal project. The remaining 25 cities were part of the Federal TTID program.

Traffic data is available only on freeways where there are sensors. There is no data for arterials.

**Format:** The real-time traffic data is fed back to the TMCs through the Traffic.com Stakeholder Application and via an XML data feed.

**System Requirements:** The real-time applications work best in read-only mode using Internet Explorer or Firefox. However, not all application features are compatible with Firefox.

**Data Quality:** Navteq is responsible for data quality, as dictated through the contract language. The TTID Program has achieved 95 percent availability and accuracy for 27 State agencies. This is higher than that achieved by State agencies on their own (below 60 percent) because they often do not have the funding for maintenance of equipment.

**Metadata:** Sensor Owner, State, Descriptive Station ID, Unique Station ID, Route, Direction, ITS Direction Code, Station Description, Station Milepost, Unique Site ID, Site Milepost (if available), Latitude, Longitude, Unique Sensor ID, Lane Type, Lane Number, Lane Position, Device Number, Sensor Type, ITS Sensor Type Code, Activation Date, Inactivation Date, Sensor Station Status, Sensor Status, and Device Status.

**Timeliness:** A 60-second data collection interval is used for Traffic.com sensors installed for the project. Data is stored in its raw format by minute, and in 5-minute lane-by-lane data, 15-minute, 1 hour and 24 hour data aggregated by direction in CSV format. The raw data is available for 90 days on the stakeholder site and then stored off-line. CSV reports are available for four years then stored off-line. Archived data is currently available from 2007 to the present. TTID data was not archived before 2007, so historical data is not available.

**Longevity/Future Plans:** The TTID program is currently funded through an earmark program that is scheduled to end in 2014. It is likely that the partnership and overall program will continue, and Navteq will likely maintain the data warehouse beyond 2014 once the contract ends.

**Users:** Customers of TTID data include local traffic management centers, FHWA Program Offices, and participating State agencies.

**Accessibility:** Access to the data is provided through the ITIP Stakeholder Applications Data Warehouse,[58] which is available to FHWA and the participating State agencies (currently 27). Each FHWA Program and participating agency is provided with a User ID and password for the data warehouse. The only restriction on use is that users are not allowed to sell the data.

[58] ITIP Stakeholder Applications Data Warehouse website, http://stakeholder.traffic.com.

**Contact:** Bob Rupert, 202-366-2194, Robert.Rupert@dot.gov; Jimmy Chu, 202-366-337, Jimmy.Chu@dot.gov.

***Urban Congestion Report (UCR) Program***

**Data Owner:** FHWA Office of Operations

**Description:** The Urban Congestion Report (UCR) is produced on a quarterly basis and characterizes emerging traffic congestion and reliability trends at the national and city level. The report utilizes archived traffic operations data gathered from State DOTs and a private traffic information company. Performance measures include congested hours, travel time index, planning time index, percent change in VMT, and percent of usable data. The production of this report is a cooperative effort between the Texas Transportation Institute and FHWA.

**Purpose:** The purposes of the Urban Congestion Report are to:

- Provide timely congestion and travel reliability information to State and local agencies;
- Demonstrate the use of archived traffic operations data for performance monitoring; and
- Promote State and local performance monitoring to support transportation decisions.

**Data Source Hierarchy:** The Urban Congestion Report is classified as a report in the data source hierarchy.

**Geographic Scope/Coverage:** The report is currently using data from 23 urban areas in the U.S., although current data sources and availability fluctuates from year to year (e.g., two cities are currently not reporting due to server problems). The Urban Congestion Report includes only those roadways that are instrumented with traffic sensors for the purposes of real-time traffic management and/or traveler information. In many cities, this typically includes the most congested parts of the freeway system. The report currently does not include congestion information on arterial streets.

The congestion information presented in the report may not be representative of the entire roadway system in any particular city. Construction may affect the roadways that are included in the report. The congestion and reliability trends are calculated by comparing the most recent three months of the current year with those of the preceding year. Only instrumented roadways that provided data in both years are included in the Urban Congestion Report.

**Format:** The Urban Congestion Report consists of summary tables and charts that are compiled from 5-minute section level speeds and vehicle-miles of travel (VMT). The report is available in HTML and PDF format.

**System Requirements:** The Urban Congestion Report is designed for use on-line or within Adobe Acrobat Reader.

**Data Quality:** Several automated and manual quality assurance procedures are used in developing the Urban Congestion Report. The automated validity checks and other data processing details are

described in the FHWA report, *Monitoring Urban Freeways in 2003: Current Conditions and Trends from Archived Operations Data.*[59]

**Metadata:** An on-line document[60] describes the performance measures, indicators, and calculation methodology for the Urban Congestion Report.

**Timeliness:** Each quarterly UCR compares data from the most recent three months to the same three months in the previous year. Quarterly reports are currently available from the beginning of Fiscal Year 2009 (October - December 2008) to the beginning of Fiscal Year 2012 (October-December 2011).

**Longevity/Future Plans:** The Office of Operations will continue to publish Urban Congestion Reports for the foreseeable future.

**Users:** The Urban Congestion Reports are used primarily by internal customers (e.g., for performance measures, budgets, reports to congress, etc.).

**Accessibility:** Urban Congestion Reports are accessible by the public and are available for download through the Urban Congestion Report website.[61]

**Contact:** Rich Taylor, FHWA Office of Transportation Management, 202-366-1327, rich.taylor@dot.gov.

***State Traffic Monitoring Stations***

**Data Owner:** State DOTs

**Description:** State DOTs install and maintain traffic monitoring stations throughout the State to support their traffic monitoring programs. These stations are established at both permanent and portable sites to collect volume, vehicle classification, and truck weight data. Truck weight data are collected at weigh-in-motion (WIM) stations. Special counts are also done for project specific needs, including traffic studies under the direction of the DOT. Data is also collected to provide traffic data for HPMS-designated segments. Speed data is a recent requirement from FHWA and many States are starting to incorporate the collection of speed data into their travel monitoring programs.

**Purpose:** The purpose of traffic monitoring stations is to support the State's travel monitoring program by collecting travel data including volume, classification, and weight data. The classification and weight data are used to monitor truck travel within and through the State.

**Data Source Hierarchy:** State traffic monitoring stations are a primary source of raw traffic volume and classification data. The WIM stations are the primary source of weight data.

[59] Turner, S., R. Margoitta, and T. Lomax. Monitoring Urban Freeways in 2003: Current Conditions and Trends from Archived Operations Data. Prepared for FHWA Office of Operations, FHWA-HOP-05-018, December 2004.

[60] FHWA Urban Congestion Report website, Documentation and Definitions page, http://www.ops.fhwa.dot.gov/perf_measurement/ucr/documentation.htm.

[61] FHWA Urban Congestion Report website, http://www.ops.fhwa.dot.gov/ perf_measurement/ucr/.

**Geographic Scope/Coverage:** Placement of the State traffic monitoring stations are primarily determined by the DOT. However, in some States, the location of traffic stations within MPO boundaries or cities may be determined as a joint effort between the State and local governments in order to facilitate the mutual needs for traffic data from each of these entities.

**Format:** The format of the raw data collected at the traffic stations is usually a function of the equipment used in the collection process. The data is transmitted to a location (part of the State DOT offices) for review and further processing.

**System Requirements:** System requirements depend upon the equipment used to collect the data at the traffic station. Data is routinely polled and transmitted from the data collection device to another office (at the DOT) for further processing.

**Data Quality:** QC of the data collected by traffic stations is done at another location (at offices of the DOT).

**Metadata:** Reporting of metadata varies by State DOT. At a minimum, metadata is required to be reported for traffic data provided as part of a State's HPMS submittal.

**Timeliness:** Traffic monitoring programs collect (at a minimum) hourly volumes by direction (and lane). Continuous counts are taken 365 days a year at a small number of locations. Control or seasonal counts are usually taken from two to twelve times a year, for periods ranging from 24 hours to two weeks. Coverage counts are short duration counts, ranging from six hours to seven days, and are distributed throughout the system to provide point-specific information and area-wide coverage. WIM data are reported annually.

**Longevity/Future Plans:** This varies by State DOT; however, at a minimum the traffic data is collected in accordance with Federal guidelines contained in the Traffic Monitoring Guide (TMG)[62] and the HPMS system. The update of the Traffic Monitoring Guide will also include guidelines for collecting bicycle/pedestrian data.

**Users:** Federal, State, and local governments use traffic monitoring data for travel demand modeling, systems planning, and meeting Federal reporting purposes as required or needed.

**Accessibility:** The State DOT determines how and what traffic data is provided internally and externally. This is usually through a web interface, or FTP server if large quantities of data are requested.

**Contact:** Varies at the State and local government levels.

***National Speed Data Collection and Reporting Program***

**Data Owner:** FHWA Office of Highway Policy Information

[62] FHWA Office of Highway Policy Information Traffic Monitoring Guide website, http://www.fhwa.dot.gov/ohim/tmguide/.

**Description:** With increased emphasis on performance-driven funding and growing public demand for accountability associated with transportation investments, FHWA is evaluating the feasibility of establishing a national speed data collection and reporting program.

**Purpose:** The proposed program would standardize speed data collection procedures, and provide monthly speed trends at the national level to evaluate travel trend and travel conditions.

**Data Source Hierarchy:** TBD

**Geographic Scope/Coverage:** Speed trends would be covered at the National level.

**Format:** TBD

**System Requirements:** TBD

**Data Quality:** TBD

**Metadata:** TBD

**Timeliness:** The program will provide speed trends on a monthly basis.

**Longevity/Future Plans:** Although plans for a data program have not been finalized, this could be a future source of speed data.

**Users:** TBD

**Accessibility:** TBD

**Contact:** Tianjia Tang, 202-366-2236, tianjia.tang@dot.gov

***Traffic Volume Trend (TVT) Reports***

**Data Owner:** FHWA Office of Highway Policy Information

**Description:** The Traffic Volume Trends (TVT) report is published monthly by the Federal Highway Administration (FHWA). The report estimates the vehicle miles traveled (VMT) by State along several functional classes of roads. The estimates are based on two sources of data:[63]

- The Highway Performance Monitoring System (HPMS), and
- Monthly traffic counts from automatic traffic recorders (ATRs).

The HPMS compiles annual data received from the States of the condition and performance of all roads in the United States. HPMS includes the annual average daily traffic (AADT) by road segment. When these AADTs are multiplied by the lengths of their corresponding road segments and summed for all road segments and days of the year, they yield the annual VMT.

[63] FHWA Office of Highway Policy Information, Traffic Volume Trends Frequently Asked Questions website, http://www.fhwa.dot.gov/ohim/tvtw/tvtfaq.cfm

Monthly average daily traffic (MADT) is computed from the ATR traffic counts. Each MADT is compared with the MADT for the same month the previous year to yield a change rate. The change rates are averaged by Roadway Functional Classes. If a State does not provide traffic data on time, their change rates are estimated from the surrounding States. Estimates are re-adjusted annually to match the vehicle miles of travel from the HPMS and are continually updated as data becomes available.

**Purpose:** The purpose of TVT reports is for assessing and reporting highway system performance under FHWA's Strategic Planning Process.

**Data Source Hierarchy:** TVT reports are compiled from monthly traffic volume data for each State, and are therefore classified as reports.

**Geographic Scope/Coverage:** Data to support TVT reports are collected from approximately 4,000 ATR locations nationwide. Coverage is generally dictated by HPMS' sample size requirements. At a minimum, a State should have six permanent sites for each grouping of functional classified roadway.

Data that covers a minimum of 30 States and 70% of the VMT is required for publication of a TVT report. The following month's TVT report will include an update which covers more data.

**Format:** TVT reports are distributed in Adobe PDF and Microsoft Excel format.

**System Requirements:** TVT reports are designed for use with Adobe Acrobat Reader, Microsoft Excel, or MS Excel Viewer.

**Data Quality:** The FHWA runs quality control checks on all data received. If data passes the checks, they are used for the TVT report.

**Metadata:** Metadata are not provided. However, an online Frequently Asked Questions section and notes within the TVT reports provide detail on data sources and calculation methodology.

**Timeliness:** States submit monthly permanent traffic record data to FHWA within 30 days after the closing of the month. TVT reports are published and made available to the public within 60 days after the close of the month. Monthly TVT reports are available for January 2002 through January 2012, and annual archives are available for 1992 through 2002. Historical VMT data by month is available from January 1970 through December 2010.

**Longevity/Future Plans:** TVT reports will continue to be published on a monthly basis.

**Users:** Customers of TVT reports include Federal, State and local agencies, researchers, the White House, Congress, and others with a need for traffic trend data.

**Accessibility:** TVT reports are available online[64] to all users.

**Contact:** Daniel Jenkins, Office of Highway Policy Information, 202-366-1067, daniel.jenkins@dot.gov

[64] FHWA Office of Highway Policy Information, Traffic Volume Trends website, http://www.fhwa.dot.gov/policyinformation/travel_monitoring/tvt.cfm.

### *Travel Monitoring and Analysis System (TMAS)*

**Data Owner:** FHWA Office of Highway Policy Information

**Description:** The Travel Monitoring and Analysis System (TMAS) provides online data submitting capabilities to State Traffic Offices so that travel monitoring data can be submitted to FHWA. These data currently include monthly volume data for the Traffic Volume Trends Report, vehicle classification, and truck weight data. TMAS currently hosts more data records than any other data program within FHWA. It isn't a true FHWA data program, but it is authorized by the Transportation Research Board (TRB) and managed by Turner-Fairbank Highway Research Center (TFHRC).

**Purpose:** The Federal Highway Administration (FHWA) has the responsibility to ensure that adequate highway transportation information is available to support its functions and responsibilities, including those of the United States Congress. The purpose of TMAS is to support States in providing vehicle classification and truck weight data to support these needs.

**Data Source Hierarchy:** TMAS is classified as a system.

**Geographic Scope/Coverage:** This varies with State DOT; however, at a minimum, geographic scope/coverage of TMAS data is established based on Federal guidelines in the Traffic Monitoring Guide (TMG) and the HPMS system.

**Format:** Data format varies with State DOT and the data being reported. At a minimum, data format requirements are established based on Federal guidelines in the Traffic Monitoring Guide (TMG) and the HPMS system.

**System Requirements:** System requirements depend upon the data being reported and the equipment used to collect the data.

**Data Quality:** Data quality requirements are established based on Federal guidelines in the Traffic Monitoring Guide (TMG) and the HPMS system.

**Metadata:** Reporting of metadata varies from one State DOT to the other and the data being reported. At a minimum, metadata requirements are established based on Federal guidelines in the Traffic Monitoring Guide (TMG) and the HPMS system.

**Timeliness:** Timeliness varies by State DOT and the data being reported. At a minimum, reporting requirements are established based on Federal guidelines in the Traffic Monitoring Guide (TMG) and the HPMS system.

**Longevity/Future Plans:** Some States are using the Vehicle Travel Information System (VTRIS) to process vehicle classification and truck weight data, but there are plans to eventually replace VTRIS with TMAS. The FHWA Office of Highway Policy Information is expanding TMAS to include data for all National Parks, bicycle/pedestrian data, and speed data. They also plan to host Long-Term Pavement Performance (LTPP) program data. TMAS could be used to host mobility data, but it has not been discussed for other types of data (e.g., ATRI – not FHWA – hosts interstate speed and travel time data for trucks). A new initiative would grant TMAS access to non-State highway agencies.

**Users:** Users of TMAS include all State agencies and FHWA field offices.

**Accessibility:** TMAS is accessible via the FHWA's User Profile and Access Control System (UPACS) website.[65] Access is limited to State DOT, FHWA Division office personnel, and others with a legitimate reason for accessing the system. User authorization must be obtained through the UPACS Administrator within the FHWA Division office in the individual State. Contact information for all the Division offices is available on the FHWA website.[66]

**Contact:** Daniel Jenkins, Office of Highway Policy Information, 202-366-1067, daniel.jenkins@dot.gov

***Vehicle Travel Information System (VTRIS)***

**Data Owner:** FHWA Office of Highway Policy Information

**Description:** The Vehicle Travel Information System (VTRIS) is a database management system designed to assist States with processing and storing vehicle classification and weigh-in-motion data. For vehicle classification, VTRIS uses detailed vehicle classifications (Class 5 through 13), while systems such as HPMS only distinguishes between single-unit trucks, combination trucks, and all other traffic. The VTRIS system allows users to validate, edit, and summarize data, as well as generate reports on vehicle characteristics. It also maintains the permanent database of station information, vehicle classification, and truck weight measurements.

**Purpose:** The Federal Highway Administration (FHWA) has the responsibility to ensure that adequate highway transportation information is available to support its functions and responsibilities, including those of the United States Congress. The purpose of VTRIS is to support States in providing vehicle classification and truck weight data to support these needs.

**Data Source Hierarchy:** VTRIS is classified as a system, and it stores three hierarchies of data: 1) raw station information, classification, and weight data; 2) aggregated summary data; and 3) reports.

**Geographic Scope/Coverage:** In 2011, 29 States submitted truck WIM data to the FHWA for inclusion in the VTRIS program.

**Format:** VTRIS data consist of the following:

- Station, Classification, and Weight raw data that are validated and loaded into the VTRIS Database (Raw Data Tables).
- Summary data that are developed from VTRIS raw data by averaging and placed into the VTRIS Database (Summary Tables).
    - Report tables that are developed from VTRIS summary data and placed into the VTRIS Database (Report Tables). The tables in the report are designed to provide a standard format for presenting the outcome of vehicle weighing and classification efforts at truck weigh sites. They include the following:
    - Freight Applications: Connected vehicle data could benefit freight applications. There are three freight applications that would benefit from connected vehicle data:
        - ATIS for Freight. Data needs for the ATIS for Freight (FRATIS) application will depend on what comes out of the ConOps. They are looking for information to

[65] FHWA Information Systems – UPACS Login website, https://fhwaapps.fhwa.dot.gov
[66] FHWA Field Offices website, http://www.fhwa.dot.gov/field.html#fieldsites

support more efficient freight movements, including truck parking, weather information, and route guidance that incorporate bridge height/weight restrictions.

- W-2 Table: This table displays a Summary of the Vehicle Counted and the Vehicle Weighed for selected stations by Vehicle Classification. The Vehicle Classification data is averaged for each hour and the 24 hourly averages are added to the average daily count. The W-2 Table breaks the data down by Vehicle Classification 5 through 13 for each Station location.
- W-3 Table: This table contains information about the Average Weights of Empty, Loaded and all Trucks, and their Estimated Average Carried Load. This information is broken down by Vehicle Classification 5 through 13 for each Station location. This process uses weight measurements and the set of breakpoints between empty and loaded trucks to calculate Percent Load, Percent Empty, Average Load Weight and Average Empty Weight.
- W-4 Table: This table is most commonly used in Pavement design since it contains information on Truck Axle loadings and their effect on Flexible and Rigid Pavement based on 18-KIP Equivalent Axle Load. The process builds three different sections:
    I. Equivalency Factors provides the Number of Single, Tandem, Tridem and Quad axles weighed which fall into particular weight ranges and gives the resulting 18-KIP Equivalent Axle Loads on the two type of Pavement
    II. Equivalency Summary contains information on the Summary ESAL (Equivalent Single Axle Load) Design Factors such as ESAL per Vehicle, Percent Distribution for Flexible and Rigid Pavements and Truck percent.
    III. 20 Year ESAL Estimates displays flexible and rigid pavements and Growth Rates for percent Truck. For all sections, information is grouped by Truck type 3 through 13 and can be shown for selected station(s), location(s) and/or highway Functional Classification(s).
- W-5 Table: This table shows the Number of Trucks Weighed in various Gross Weight ranges, Total Average Vehicles Weighed and Total Average Vehicles Counted. This process produces values for Truck types 3 through 13 and shows for selected station(s) and/or Functional Classification(s) of highway. User defines the Ranges of Gross Weight.
- W-6 Table: This table gives the Number and Percent of Vehicle by type exceeding a User specified Axle, Tandem and Gross Weight Limit (i.e. all Violations of the Bridge Formula). A vehicle may have multiple violations of the set criteria. The information can be shown by station(s) and /or Functional Classification(s) of highway and is broken down by Vehicle type 3 through 13. The Number Exceed and the Percent Exceed are calculated for each Axle Type and Gross Weight. This table includes the Violations of the Bridge Gross Weight Formula depending only on Vehicle type. The Bridge Gross Weight Formula provides a standard to control the spacing of Truck axles on the Vehicle that uses highway bridges. The Number of Violations of the Bridge Gross Weight Formula calculated previously is in the Summary process.
- W-7 Table: This table gives the Number and Percent of Average Daily Count for those measurements where Axle Load is above a Limit. The Gross Weight Limits by Vehicle Class (in Kilograms) was assigned by the user. The information can

be shown for each station(s) and/or Functional Classification(s) of highway broken down by Vehicle types 5 through 13.

- Load Log files which keep the information on loads performed during the systems life. They are also considered to be the part of the VTRIS Database.
- Summary Log files keep track of the Summaries created. It is also considered to be part of the VTRIS Database.
- VTRIS System Tables. The data that keep track of various VTRIS activities are read-only.
- VTRIS Temporary Tables. These tables store data that is created for a particular task and then deleted when the task is accomplished.

**System Requirements:** The VTRIS software is distributed on one CD for V.5.0 (or three 1.44MB disks for V.2.6), which contains the forms, programs, system tables, reports, and other components necessary for VTRIS to be properly installed and operational. The installation package is also available from the Office of Highway Policy Information's Internet website. Recommended system requirements for running VTRIS include: Windows 95/NT, Pentium 133 Mhz processor (or higher), 16 MB RAM, 15 MB disk space for installation, and at least 1 GB disk space for data storage.

**Data Quality:** The Load/Validation subsystem is designated to detect errors in uploaded data, provide a facility to correct those errors and eventually load valid data into one of the VTRIS tables. The validated data is placed into STATION, CLASSIFICATION and WEIGHT tables with the codes for each record reflecting the level of reliability of data (fields). The records with errors are placed into the error tables. Data validity includes not only non-violation of coding rules established for STATION, CLASSIFICATION and WEIGHT data, but also includes the support of referential constraint and data integrity within and among the tables.

1. The data to be loaded into the VTRIS tables may arrive as ASCII files, and newly adopted requirements demand that the files be named according to the following convention: SSYY.XXX. Where SS - state code, YY - year of data and XXX - 'STA', 'CLA', or 'WGT' for STATION, CLASSIFICATION and WEIGHT data respectively. Each ASCII record goes through the following validation steps:
2. Determination of the record's type in accordance with FHWA formats set for coding of STATION, CLASSIFICATION and WEIGHT data. The data type, record length, and other record parameters are also checked.
3. Validation of single fields within a record to ensure that they hold a valid field value or are within a specific range of values. In addition, cross-validation within a single record between two or more fields is done to ensure that data fields are not contradicting each other.
4. Checks for duplicates and consistency between the records that are being loaded into the same table. This mostly concerns STATION and CLASSIFICATION data since there should not be duplicate records with the same key value. For WEIGHT data it is not a validation issue, but rather a matter of data maintenance since the specified key may identify an unlimited number of records. This is reflective of the fact that the table contains one record per truck measured.
5. Cross-validation between the fields of the new records and the records from the VTRIS table to prevent duplicates and support referential integrity between different VTRIS tables as well as consistency within a single table. Integrity validation requires that the CLASSIFICATION

and WEIGHT data checks against station data to make sure that the key is valid (e.g., Station-Direction-Lane exists).

The following error levels are established:

1. Junk – those records that are detected at the earliest stage of validation process and result in the records being put into the JUNK file. No further validation is possible for these records until some manual editing is done.
2. Fatal - these records that cannot be admitted “as is” even if User desires to use them. For these errors, an appropriate correction through the ERROR table Browse/Edit facility is required. These are typically errors in the key fields and other significant fields that would violate consistency and referential integrity.
3. Caution - those errors that can be fixed or flagged by User as acceptable and put into the VTRIS tables “as is”. If User accepts and flags them, an appropriate Flag Code will be placed into a VTRIS table along with the record.
4. Out-of Range - for numeric fields whose values are out of a certain valid range. These errors cannot be fixed, since no manual input is allowed for the numeric fields. They can only be flagged "as is" or the whole record must be rejected by user - eliminated or placed into Junk.

**Metadata:** A data dictionary describes the data fields associated with the VTRIS database.

**Timeliness:** VTRIS data are available by hour, day, month, or year.

**Longevity/Future Plans:** VTRIS is being phased out and will eventually be replaced with the Travel Monitoring Analysis System (TMAS).

**Users:** Users of VTRIS include all State agencies and FHWA field offices.

**Accessibility:** Report data that are based on VTRIS summary data are accessible to the public and are available through the VTRIS W-Table website.[67] VTRIS Reports are available to the public between the following hours: Monday- Friday: 7 a.m. to 10 p.m. (EST), Sunday: 12 Noon to 10 p.m. (EST). A User’s Guide consisting of the manual, appendix, and database structure supplementary tables is provided on the FHWA Office of Highway Policy Information website.[68]

**Contact:** Ralph Gillmann, (202) 366-5042; David Jones, (202) 366-5053

***National Transportation Statistics***

**Data Owner:** RITA Bureau of Transportation Statistics

**Description:** National Transportation Statistics (NTS) presents statistics on U.S. transportation system, including its physical components, safety record, economic performance, energy use, and environmental impacts. National Transportation Statistics is a companion document to the *Transportation Statistics Annual Report*, which analyzes NTS data and *State Transportation Statistics*, and presents state-level data on many of the same topics.

[67] VTRIS W Table Generation website, https://fhwaapps.fhwa.dot.gov/vtris-wp/default.aspx
[68] FHWA Office of Highway Policy Information website, VTRIS Users Guide page, http://www.fhwa.dot.gov/ohim/ohimvtis.cfm

This is a large online document comprising more than 260 aggregated data tables plus data source and accuracy statements, glossary, and a list of acronyms and initialisms.

**Purpose:** This data program presents the state of the transportation system and the state of transportation data.

**Data Source Hierarchy:** NTS data is compiled from a variety of sources that are clearly identified for each individual table. Because the data is aggregated and made available in table format, NTS is classified as a report.

**Geographic Scope/Coverage:** The data tables include national coverage of the following:

- The Transportation System: Physical Extent; Vehicle, Aircraft, and Vessel Inventory; Condition; Travel and Goods Movement; and Physical Performance;
- Transportation Safety: Multimodal, Air, Highway, Transit, Railroad, Water, and Pipeline;
- Transportation and the Economy: Transportation and the Total Economy; Transportation and Consumer Expenditures; Transportation Revenues, Employment, and Productivity; and Government Finance; and
- Transportation, Energy, and the Environment: U.S. and Transportation Sector Energy Consumption; Transportation Energy Consumption by Mode; Transportation Energy Intensity and Fuel Efficiency; Air Pollution; and Water Pollution, Noise and Solid Waste.

Notes below each table provide detailed information about its geographic scope/coverage.

**Format:** The aggregated data tables are distributed in HTML, Excel, and CSV format.

**System Requirements:** The NTS data tables are designed for use online or within a spreadsheet or database environment.

**Data Quality:** BTS obtains data to support NTS from many sources, including Federal government agencies, private industry, and associations. Documentation below individual tables and in Appendix E provides detailed information on data sources, definitions, methodologies, and statistical reliability. Some of the data are based on samples and are subject to sampling variability.

**Metadata:** Descriptive data is provided below individual tables and in Appendix E about the coverage, calculation methodology, known data issues or concerns, and data sources.

**Timeliness:** Aggregate data are available from 1960 to 2011. The web edition of the National Transportation Statistics is updated quarterly.

**Longevity/Future Plans:** NTS is a primary data product of BTS and will continue to be updated on a quarterly basis.

**Users:** Customers of NTS include Federal, State and local agencies, researchers, the White House, Congress, OST, and U.S. DOT modal administrations with a need for multimodal and intermodal data and information to support comparative analysis.

**Accessibility:** NTS is accessible to the public and the datasets are available for download on the BTS website.[69]

**Contact:** For additional information, contact BTS Information Services at 800-853-1351 or RITAInfo@dot.gov.

## Detailed Inventory – Climate Data

***Road Weather Information Systems***

**Data Owner:** FHWA Office of Operations – Road Weather Management Program

**Description:** A road weather information system (RWIS) is a combination of technologies that collects, transmits, models, and disseminates weather and road condition information. The component of an RWIS that collects road weather data is the environmental sensor station (ESS). An ESS is a suite of sensors that collects and transmits pavement and meteorological data. Sensors measure a range of weather-related conditions, including pavement temperature and status (wet, dry, snow), subsurface pavement temperature, wind speed and direction, precipitation (amount, occurrence, type), water level conditions, humidity, and visibility.

**Purpose:** Road weather information is used to activate automated warning systems and provide decision support to managers in traffic management centers, road maintenance facilities, and emergency operations centers. Weather service providers also use the data to develop tailored road weather products (for example, pavement temperature forecasts). All of these activities make for safer roadway conditions for motorists.

**Data Source Hierarchy:** RWIS covers multiple levels in the hierarchy. The ESSs provide raw data on road weather conditions that are collected at the source. Road weather condition reports that are distributed to various personnel in the field are classified as reports.

**Geographic Scope/Coverage:** Geographic scope/coverage varies by region. The FHWA Road Weather Management Program has published siting criteria to satisfy road weather monitoring, detection, and prediction requirements.[70]

**Format:** Platforms use varying reporting formats. RWIS is comprised of Environmental Sensor Stations (ESS) in the field, a communication system for data transfer, and central systems to collect field data from numerous ESS. These stations measure atmospheric, pavement and/or water level conditions. Central RWIS hardware and software are used to process observations from ESS to develop now-casts or forecasts, and display or disseminate road weather information in a format that can be easily interpreted by a manager.

There are three types of road weather information: atmospheric data, pavement data, and water level data. Atmospheric data include air temperature and humidity, visibility distance, wind speed and direction, precipitation type and rate, cloud cover, tornado or waterspout occurrence, lightning, storm

[69] BTS National Transportation Statistics website, http://www.bts.gov/publications/national_transportation_statistics/

[70] Road Weather Information System Environmental Sensor Station Siting Guidelines, Version 2.0, FHWA-HOP-05-026, November 2008, pgs. 4-1 - 4-9.

cell location and track, as well as air quality. Pavement data include pavement temperature, pavement freezing point, pavement condition (e.g., wet, icy, flooded), pavement chemical concentration, and subsurface conditions (e.g., soil temperature). Water level data include stream, river, and lake levels near roads, as well as tide levels (i.e., hurricane storm surge).

**System Requirements:** System requirements depend on the type of ESS equipment deployed.

**Data Quality:** Data quality varies based on the platform used.

**Metadata:** The FHWA Office of Operations – Road Weather Management Program has published recommended and optional metadata requirements for ESS sites and their sensors.[71]

**Timeliness:** Timeliness depends on the type of ESS equipment deployed.

**Longevity/Future Plans:** Unknown

**Users:** Transportation managers utilize roadway warning systems, interactive telephone systems (e.g., 511), and web sites to disseminate road weather information to travelers in order to influence their decisions. This information allows travelers to make choices about travel mode, departure time, route selection, vehicle type and equipment, and driving behavior.

Specialized equipment and computer programs monitor weather and pavement condition elements that help users observe how adverse weather is currently affecting the highways and assess future impacts. For example, winter road maintenance managers may benefit from such a system during winter storms by making optimal use of materials and staff, selecting appropriate treatment strategies, utilizing anti-icing techniques, and properly timing maintenance activities. Traffic managers may use road weather observations to modify traffic signal timing, reduce speed limits, and close hazardous roads and bridges.

**Accessibility:** Access to raw sensor data is generally limited to internal State DOT (or other owner agency) staff. States participating in the Clarus initiative make sensor data publicly available through the Clarus System website.[72]

**Contact:** Paul Pisano, Road Weather Management Program, FWHA, (202) 366-1301, paul.pisano@dot.gov

***Clarus***

**Data Owner:** FHWA Office of Operations – Road Weather Management Program

**Description:** Clarus is an initiative to develop and demonstrate a robust data assimilation, quality checking, and data dissemination system that could provide near real-time atmospheric and pavement observations from the collective States' investments in road weather information system,

[71] Road Weather Information System Environmental Sensor Station Siting Guidelines, Version 2.0, FHWA-HOP-05-026, November 2008, pgs. 5-7 - 5-17.

[72] At the time of publication, Clarus is being transitioned to the Meteorological Assimilation Data Ingest System (MADIS). Users can request access to Clarus data through the MADIS Data Application website.

environmental sensor stations (ESS) as well as mobile observations from Automated Vehicle Location (AVL) equipped trucks.[73]

**Purpose:** The U.S. Department of Transportation (DOT) Federal Highway Administration (FHWA) Road Weather Management Program, in conjunction with the Intelligent Transportation Systems (ITS) Joint Program Office established the Clarus Initiative in 2004 to reduce the impact of adverse weather conditions on surface transportation users.

**Data Source Hierarchy:** Because Clarus provides a quality checking algorithm, it is classified as a system in the data source hierarchy.

**Geographic Scope/Coverage:** There are 40 States that currently provide ESS data for Clarus, and there are 2,231 ESS stations and 52,000 individual sensors feeding into the system.

**Format:** The files are formatted as CSV (comma-separated value) files, allowing them to be imported easily into third-party analysis tools.

**System Requirements:** The goal of Clarus is to develop an open, extensible system design that can be implemented by any organization and that can readily accept data from new sources such as VII. A model for the system design is the Linux operating system, in which many readily available 'shrink-wrapped' versions are available. Systems integration challenges always occur, of course, but the idea is to make Clarus implementation as simple as possible so that we can extend its use.

**Data Quality:** Clarus performs comprehensive Quality Checking tests, including sensor range, climate range, time step, like instrument, persistence, inter-quartile range (IQR) spatial, Barnes spatial, dew-point temperature, and sea level pressure tests.[74]

**Metadata:** The Clarus System metadata files contain a textual representation of the metadata contained within the database. The metadata files are available on the Clarus System website.[75]

**Timeliness:** The Clarus System polls its list of contributors for new observations based on each contributor's configured schedule. These observations are stored for seven days. During that time, observations can be retrieved in two ways: either as an on-demand report or as a subscription. In both cases, users define the area from which they want to retrieve observations and the specific observation types in which they are interested. For an on-demand report, users will also specify the time range (one hour at a time) for which they wish to retrieve the observations. Users can request subscription reports at various intervals (e.g., 5, 10, 15, 20, 25, or 30 minutes).[76]

---

[73] RITA ITS Joint Program Office website – ITS Research Success Stories - Clarus, http://www.its.dot.gov/clarus/.

[74] Limber, M., S. Drobot, and T. Fowler. Clarus Quality Checking Algorithm Documentation Report. FHWA Report No. FHWA-JPO-11-075, December 21, 2010. http://ntl.bts.gov/lib/38000/38500/38545/TOPR2_508_FHWA-JPO-11-075.pdf

[75] Clarus System Metadata Files website, http://www.its.dot.gov/clarus/clarus_archive.htm

[76] Mixon/Hill, Inc., Clarus System User Guide. Prepared for USDOT Federal Highway Administration, June 2011. http://www.cmap.illinois.gov/c/document_library/get_file?uuid=98ef52d4-c185-4893-88bb-0a766b21c37a&groupId=20583.

**Longevity/Future Plans:** Nixon Hill currently runs the Clarus contract, and it is in its last year of funding. The long-term plan was to turn it over to the National Weather Service; however, DOT cannot rely on NWS to meet their research needs because they do not conduct the same data quality checks that Clarus does. The Dynamic Mobility Application program will rely on Clarus data, so it will need to continue to be funded to support connected vehicle applications.

**Users:** Clarus data is used for structured demonstrations and made available to researchers, the private sector, and the National Weather Service. NWS uses it to validate forecasts and generate watches and warnings. Users of the data, whether directly or via value-added information products, also include the States, NOAA, and others such as transit and trucking agencies and emergency managers.

**Accessibility:** Clarus data is accessible to the public and is available on the Clarus System website.[77] Subscriptions reports that are fulfilled periodically (such as every 5, 10, 15, 20, 25, or 30 minutes) may require a password.

**Contact:** Paul Pisano, Road Weather Management Program, FWHA, (202) 366-1301, paul.pisano@dot.gov; Ben McKeever, (202) 366-4876, ben.mckeever@dot.gov

## Detailed Inventory – Modal Data

***Border Crossing Data***

**Data Owner:** RITA Bureau of Transportation Statistics

**Description:** The Bureau of Transportation Statistics (BTS) Border Crossing/Entry Data provides summary statistics for incoming crossings at the U.S.-Canadian and the U.S.-Mexican border at the port level. Data are available for trucks, trains, containers, buses, personal vehicles, passengers, and pedestrians. BTS received border crossing data from U.S. Customs & Border Protection. The data only reflect the number of vehicles, containers, passengers or pedestrians entering the United States. Customs and Border Protection does not collect comparable data on outbound crossings.

**Purpose:** BTS uses border crossing/entry data as a core source for analyses on U.S. and North American trade and travel. Such analyses are included in interpretive reports such as the *North American Trade and Travel Trends, U.S. International Travel and Transportation Trends,* and *U.S. International Trade and Freight Transportation Trends*. From a transportation perspective, these data are important because they provide a level of information on the amount of activity at a given port. BTS uses the Border Crossing/Entry Data in conjunction with other datasets (such as the Transborder Freight Data) to assess and track trade and travel flows for North America.

**Data Source Hierarchy:** The North American Border Crossing/Entry Data is classified as a repository that provides access to border crossing/entry data.

[77] Clarus System website. (At the time of publication, Clarus is being transitioned to the Meteorological Assimilation Data Ingest System (MADIS). Users can request access to Clarus data through the MADIS Data Application website.

**Geographic Scope/Coverage:** Border Crossing/Entry data are reported at the U.S. Customs and Border Protection port level. The U.S. Census Bureau maintains a list of Customs districts/ports, port codes and descriptions known as the Schedule D.[78]

**Format:** BTS provides access to the data through an interactive searchable interface called North American Border Crossing/Entry Data Online Data Tool. This allows users to create multi-variable cross-tabulations on port, geography and commodity for all modes of transportation. Search results can be viewed online, printed, or downloaded in CSV format. BTS tabulates and releases data for 12 data elements across five surface modes of transportation. The 12 data elements are listed below with their corresponding definitions, as determined by U.S. Customs and Border Protection.

- **Truck crossings –** Number of arriving trucks; does not include privately owned pick-up trucks.
- **Truck container crossings (loaded and unloaded) –** A container is any conveyance entering the U.S. used for commercial purposes, full or empty. In this case, it is the number of full or empty truck containers arriving at a port. This series includes containers moving as in-bond shipments.
- **Train crossings –** Number of arriving trains at a particular port.
- **Rail container crossings (loaded and unloaded) –** A container is any conveyance entering the U.S. used for commercial purposes, full or empty. In this case, it is the number of full or empty rail containers arriving at a port. This series includes containers moving as in-bond shipments.
- **Passengers crossing in trains –** Number of passengers and crew arriving by train and requiring U.S. Customs processing.
- **Bus crossings –** Number of arriving buses at a particular port, whether or not they are carrying passengers.
- **Passengers crossing in buses –** Number of persons arriving by bus requiring U.S. Customs processing.
- **Privately owned vehicle crossings –** Number of privately owned vehicles (POVs) arriving at a particular port. Includes pick-up trucks, motorcycles, recreational vehicles, taxis, snowmobiles, ambulances, hearses, and other motorized private ground vehicles.
- **Passengers crossings in privately owned vehicles –** Persons entering the United States at a particular port by private automobiles, pick-up trucks, motorcycles, recreational vehicles, taxis, ambulances, hearses, tractors, snow-mobiles and other motorized private ground vehicles.
- **Pedestrian crossings –** The number of persons arriving on foot or by certain conveyance (such as bicycles, mopeds, or wheel chairs) requiring U.S. Customs processing.

[78] U.S. Census Bureau Schedule D – District/Port List, http://www.census.gov/foreign-trade/schedules/d/dist2.txt

**System Requirements:** The North American Border Crossing/Entry Data is designed for use within an online environment. Datasets can be downloaded in CSV format for use within a database or spreadsheet program.

**Data Quality:** In order to provide BTS customers with the most relevant data from a transportation perspective, BTS has developed a system for extracting, processing, and validating Border Crossing Data for all U.S. Customs land ports of entry. BTS extracts data for all major land ports of entry on the northern and southern borders. These data are then organized into monthly and annual tables for each port. Ports are sorted by State, and State totals are aggregated into totals for the northern and southern borders. During this extraction process, BTS reviews the data for consistency and consults with the U.S. Customs Service, the primary source agency, on data quality issues. New data are added approximately six months after the month the data are collected; this processing time is used to validate and revise the data.

**Metadata:** Definitions for border crossing data elements are provided on the BTS website.[79]

**Timeliness:** Data are collected and aggregated monthly, but updates are sent to BTS on a quarterly basis. The Border Crossing/Entry Data are periodically updated. Data become available approximately six months after the end of the month that the data are collected. This allows time for receipt of the data from the Customs Service, as well as BTS quality assurance and validation. Most data elements are available beginning in 1994. Monthly data are available between 1995 and 2008.

**Longevity/Future Plans:** Border crossing data will continue to be reported to BTS for the near future.

**Users:** A wide variety of organizations utilize the Border Crossing/Entry Data, including government agencies (State, local, Federal and non-U.S.), private sector organizations, media, non-profit organizations/associations, and academia. Government agencies include metropolitan planning organizations, State Departments of Transportation, economic development agencies, public health agencies, Federal agencies (including modal administrations within the U.S. Department of Transportation), and others. Private sector organizations include manufacturing firms, transportation service providers, marketing organizations, consulting groups etc. These segments of the public use the dataset for a variety of purposes, including transportation planning, port studies, travel analyses, corridor assessments, and other purposes. In many cases, these data are used in conjunction with other trade and travel data sources.

**Accessibility:** BTS provides access to the data through an interactive searchable interface called North American Border Crossing/Entry Data Online Data Tool.[80] The interface is accessible to the public.

**Contact:** For additional information, contact the BTS Information Services at 800-853-1351 or RITAInfo@dot.gov.

[79] North American Border Crossing/Entry Data Website – Frequently Asked Questions, http://www.bts.gov/programs/international/transborder/TBDR_BC_FAQs.html,

[80] North American Border Crossing/Entry Data Online Data Tool, http://www.bts.gov/programs/international/transborder/TBDR_BC/TBDR_BC_Index.html

### *North American Transborder Freight Data*

**Data Owner:** RITA Bureau of Transportation Statistics

**Description:** The North American Transborder Freight Database contains freight flow data by commodity type and by mode of transportation (rail, truck, pipeline, air, vessel, and other) for U.S. exports to and imports from Canada and Mexico. The database includes two sets of tables; one is commodity based while the other provides geographic detail.

**Purpose:** The purpose of the database is to provide transportation information on North American trade flows. This type of information is being used to monitor freight flows since the signing of the North American Free Trade Agreement (NAFTA) by the United States, Canada and Mexico in December 1992 and its entry into force on January 1, 1994. The database is also being used for trade corridor studies, transportation infrastructure planning, marketing and logistics plans and other purposes. It allows users to analyze movement of merchandise by land, water, and air.[81]

**Data Source Hierarchy:** The North American Transborder Freight Database is classified as a repository that provides access to border crossing/entry data.

**Geographic Scope/Coverage:** Beginning with the 1997 data, the North American Transborder Freight Data officially documents U.S. trade with Canada and Mexico for shipments that entered or exited the United States by surface modes of transport (other than air or maritime vessel). The data from April 1993 to December 1996 included official U.S. trade with Canada and Mexico by surface modes and transshipments that moved from a third country through Canada or Mexico to the United States or from the United States to a third country through Canada or Mexico. For this time period, it was not possible to separate transshipment activity from the official trade activity at the detailed level. Due to customer requests, BTS discontinued the inclusion of transshipment activity in the North American Transborder Freight Data for data after January 1997. This allows customers to perform comparable trade analyses by mode of transportation.

**Format:** The North American Transborder Freight Dataset is extracted from the Census Foreign Trade Statistics Program. Import and export data are captured from administrative records required by the Departments of Commerce and Treasury. Historically, these data were obtained from import and export hard copy documentations that the U.S. Customs Service (Customs) collected at a port of entry or exit. However, with increasing amount of import and export, statistical information is now being captured electronically.

- **Imports –** U.S. imports of merchandise is compiled primarily from automated data submitted through the U.S. Customs' Automated Commercial System. Data are compiled also from import entry summary forms, warehouse withdrawal forms and Foreign Trade Zone documents, as law requires them to be filed with the U.S. Customs and Border Protection.
- **Exports –** U.S. exports of merchandise are compiled primarily from the Automated Export System (AES), paper Shipper's Export Declarations (SEDs), and Canadian data provided by Statistics Canada. The United States is substituting Canadian

[81] North American Transborder Freight Data website, http://www.bts.gov/programs/international/transborder/

import statistics for U.S. exports to Canada in accordance with a 1987 Memorandum of Understanding signed by the Census Bureau, U.S. Customs and Border Protection, Canadian Customs, and Statistics Canada. This data exchange includes only U.S. exports destined for Canada and does not include shipments destined for third countries by routes passing through Canada.

The monthly and annual North American Transborder Freight Data can be downloaded in raw table formats through the Freight Data Interface. Users with a need to customize and manipulate these statistics for various purposes may choose to download these files instead of using the interactive searchable interface. This includes U.S.-Canada and U.S.-Mexico trade data, but excludes volume and speed data.

**System Requirements:** The Freight Data Interface is designed for use within an online environment. Datasets can be downloaded in CSV format for use within a database or spreadsheet program.

**Data Quality:** Import and export data are a complete enumeration of documents collected by U.S. Customs and Border Protection and are not subject to sampling errors. However, while quality assurance procedures are performed at every stage of collection, processing, and tabulation, the data are still subject to several types of non-sampling errors. The most significant of these include reporting errors, undocumented shipments, timeliness, data capture errors, transiting goods, and underestimation of low-valued transactions.

**Metadata:** A table structure and data dictionary for the North American Transborder Freight database is available on the BTS website.[82]

**Timeliness:** The North American Transborder Freight database is published monthly. Monthly data are available since 1994. Prior to 1993, the Census Bureau only provided mode of transportation information for air, water and other. Details were not available for trade by surface mode of transportation.

**Longevity/Future Plans:** Because of the desire to offer the highest level of timeliness to customers, BTS will continue to release the North American Transborder Freight Data online on a monthly schedule.

**Users:** A wide variety of organizations utilize the North American Transborder Freight Data, including government agencies (State, local, Federal and non-U.S.), private sector organizations, media, non-profit organizations/associations, and academia.

**Accessibility:** BTS provides access to the data through an interactive searchable interface called North American Transborder Freight Data.[83] This allows users to create multi-variable cross-tabulations on port, geography and commodity for all modes of transportation. Search results can be viewed online and then downloaded. This data is available online to all users on the website.

[82] North American Transborder Freight Data Website – Table Structure and Data Fields, http://www.bts.gov/programs/international/transborder/TBDR_DataFields.html

[83] North American Transborder Freight Data Interface website, http://www.bts.gov/programs/international/transborder/TBDR_QA.html.

**Contact:** For additional information, contact the BTS Info Services at 800-853-1351 or RITAInfo@dot.gov.

***Commodity Flow Survey***

**Data Owner:** RITA Bureau of Transportation Statistics

**Description:** The Commodity Flow Survey (CFS) is the primary source of national and state-level data on domestic freight shipments by American establishments in mining, manufacturing, wholesale, auxiliaries, and selected retail commodities. Data are provided on the types, origins and destinations, values, weights, modes of transport, distance shipped, and ton-miles of commodities shipped. It provides a modal picture of national freight flows, and represents the only publicly available source of commodity flow data for the highway mode. The CFS is conducted as a partnership effort between the Bureau of Transportation Statistics and the U.S. Census Bureau.

**Purpose:** The purpose of the CFS is to provide a comprehensive picture of national freight flows, and it is the only available source of data for the highway mode, which carries about 71 percent of the value and 69 percent of the tonnage of freight transported. The primary objective of the 2007 Commodity Flow Survey (CFS) is to estimate shipping volumes (value, tons, and ton-miles) by commodity and mode of transportation at varying levels of geographic detail (i.e., national, State, select MSAs). A secondary objective is to estimate the volume of shipments moving from one geographic area to another (e.g., flows of commodities between States, regions, etc.) by mode and commodity.

**Data Source Hierarchy:** The CFS is classified as a report.

**Geographic Scope/Coverage:** The CFS is a shipper-based survey, and captures data on shipments originating from select types of business establishments located within the 50 States and the District of Columbia. The 2007 survey sampled over 100,000 establishments with paid employees that were located in the United States and were classified, using the 2002 North American Industry Classification System (NAICS) in mining, manufacturing, wholesale trade, and select retail trade industries (electronic shopping, mail-order houses, and fuel dealers). The survey also covers auxiliary establishments (i.e., warehouses and managing offices) of multi-purpose establishment companies, which had non-auxiliary establishments that were in-scope to the CFS or were classified in retail trade.

The CFS does not include establishments classified as forestry, fishing, utilities, construction, transportation, and most retail and services industries. Also excluded were farms and government-owned entities (except government-owned liquor stores). Foreign-based business importing to the United States are also excluded from the survey sample; however, in theory, domestic portions of imported shipments can be captured in the CFS upon arrival at a U.S. based establishment (assuming it is an eligible shipping establishment included in the CFS).

**Format:** The CFS collects information about the commodities shipped by U.S. establishments. Information collected for each outbound shipment includes:

- Shipment ID Number
- Shipment Date (Month, Day)
- Shipment Value

- Shipment Weight in pounds
- Commodity Code from Standard Classification of Transported Goods (SCTG) List
- Commodity Description
- Hazmat Flag (United Nations (UN) or North American (NA) number)
- U.S. Destination (City, State, Zip code) – Gateway for Export Shipment
- Modes of Transport
- Foreign Destination (Exports only – City, Country)
- Export Mode

CFS data are available in table format that provide shipment characteristics (value, tons, ton-miles, average miles) by mode of transport, total modal activity, distance shipped, weight of shipments, industry, and commodity shipped. Shipment characteristics are summarized at the national, State, and 73 select metropolitan statistical area levels; geographic flow tables by commodity, industry and mode; and specialty tabulated tables for hazardous materials and exports. Data tables are available for download in HTML, Excel, or CSV format.

**System Requirements:** CFS data are designed for use within an online environment. Datasets can be downloaded in HTML, Excel, or CSV format for use within a database or spreadsheet program.

**Data Quality:** The estimates presented in the 2007 CFS may differ from the actual, unknown population values. Statisticians define this difference as the total error of the estimate. When describing the accuracy of survey results, it is convenient to discuss total error as the sum of sampling error and non-sampling error. Sampling error is the average difference between the estimate and the result that would be obtained from a complete enumeration of the sampling frame conducted under the same survey conditions. Non-sampling error encompasses all other factors that contribute to the total error of a sample survey estimate.

The sampling error of the estimates in this publication can be estimated from the selected sample because the sample was selected using probability sampling. Common measures related to sampling error are the sampling variance, the standard error, and the coefficient of variation (CV). The sampling variance is the squared difference, averaged over all possible samples of the same size and design, between the estimator and its average value. The standard error is the square root of the sampling variance. The CV expresses the standard error as a percentage of the estimate to which it refers.

Non-sampling errors are difficult to measure and can be introduced through inadequacies in the questionnaire, non-response, inaccurate reporting by respondents, errors in the application of survey procedures, incorrect recording of answers, and errors in data entry and processing. In conducting the 2007 CFS, every effort was made to minimize the effect of non-sampling errors on the estimates. Data users should take into account both the measures of sampling error and the potential effects of non-sampling error when using these estimates.

**Metadata:** An online Frequently Asked Questions page[84] provides background information on the data collected as part of the CFS.

**Timeliness:** The survey is conducted on a five-year cycle as part of the Economic Census. The three previous surveys were conducted in 1993, 1997, and 2002, and most recently in 2007.

**Longevity/Future Plans:** For the 2012 Commodity Flow Survey, BTS has added data fields on temperature controlled and rush delivery shipments.

**Users:** Analysts and researchers in both the public and private sectors use data from the CFS for a variety of purposes, including:

- Analyzing trends in goods movement over time;
- Conducting national, regional and sectoral economic analysis;
- Developing models and analytical tools for policy analyses, management and investment decisions;
- Forecasting future demand for goods movement and associated infrastructure and equipment needs;
- Establishing benchmarks for estimating national accounts; and
- Analyzing and mapping spatial patterns of commodity and vehicle flows.

CFS data are used as the basis for the Federal Highway Administration's Freight Analysis Framework, a model that displays by mode the movement of goods over the national transportation network. In addition, the CFS Hazardous Materials report is the sole source of hazardous materials flow data available for the highway mode.

**Accessibility:** CFS reports and tables are accessible to the public and are available for download on the BTS website.[85]

**Contact:** Contact the Commodity Flow Survey team at cfs@dot.gov

***Intermodal Passenger Connectivity Database***

**Data Owner:** RITA Bureau of Transportation Statistics

**Description:** The Intermodal Passenger Connectivity Database (IPCD) is a nationwide data table of passenger transportation terminals, with data on the availability of connections among the various scheduled public transportation modes at each facility. In addition to geographic data for each terminal, the data elements describe the availability of rail, air, bus, transit, and ferry services. This

[84] Commodity Flow Survey Website – Frequently Asked Questions website, http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/commodity_flow_survey/index.html.

[85] Commodity Flow Survey Website, http://www.bts.gov/publications/ commodity_flow_survey/

data has been collected from various public sources to provide the only nationwide measurement of the degree of connectivity available in the national passenger transportation system.[86]

**Purpose:** The Intermodal Passenger Connectivity Database measures the connectivity between different modes of passenger transportation by counting the number of passenger transportation terminals and the availability of connections among the various scheduled public transportation modes at each facility. A series of analytical reports based on the IPCD data measures the degree of connectivity in the passenger transportation system.

**Data Source Hierarchy:** The Intermodal Passenger Connectivity Database is a source of raw data on passenger transportation terminals and availability of connections among various public transportation modes.

**Geographic Scope/Coverage:** As of April 2012, the database includes 1595 intercity and commuter rail stations, 667 airports, and 288 scheduled passenger ferry terminals; when complete in June 2012 it will also include an estimated 1,800 heavy and light rail stations, and 2,500 intercity bus terminals. The completed database will include passenger terminals within the 50 States and the District of Columbia served by the regularly scheduled service of the following modes: scheduled intercity rail, commuter rail, heavy rail transit, light rail transit, intercity bus, code-share buses and supplemental service buses operated for intercity rail (Amtrak Thruway) and air carriers, intercity ferries, and transit or local ferries. Included in the intercity bus category are intercity airport bus services. All identifiable regularly scheduled stopping locations, whether or not a terminal building is present, are considered terminals for this database. Transit bus service locations are included when they offer connections to other scheduled modes. No attempt will be made to list every possible transit bus stop as in most cities there is usually at least one stop in every block.

**Format:** The database consists of passenger transportation terminals, with data on the availability of connections among the various scheduled public transportation modes at each facility. An online lookup tool allows users to query the database for specific data fields. The database is available for download in CSV format.

**System Requirements:** The database is available online, and it can be downloaded in CSV format for use within a database or spreadsheet program.

**Data Quality:** The source agencies are the data stewards and are ultimately responsible for the maintenance and accuracy of the data. BTS documented its procedure for defining connectivity in an April 2009 Technical Report.[87]

**Metadata:** A data dictionary is provided on the TranStats website.[88] Each record in the database, which represents a specific passenger transportation facility, includes the date that the record was most recently updated and source notes on where information was obtained.

[86] BTS TranStats Website – Intermodal Passenger Connectivity Database Data Profile, http://www.transtats.bts.gov/DatabaseInfo.asp?DB_ID=640&Link=0.

[87] Goldberg, B. The Background, Criteria, and Usage of the Intermodal Passenger Connectivity Database. RITA BTS Special Report TR-007, April 2009. http://ntl.bts.gov/lib/35000/35400/35438/Background.pdf.

**Timeliness:** Data collection has been ongoing since mid-2006. Data is as of the 'updated date' shown in the record for each terminal.

**Longevity/Future Plans:** Data collection began in mid-2006 and has been collected for intercity rail stations, commuter rail stations, airline airports, and ferry terminals only. Data on terminals of other modes is being collected and will be released when it is available. It is anticipated that the database will be complete in June 2012.

**Users:** Customers for the Intermodal Passenger Connectivity Database include Federal, State and local agencies, researchers, intermodal facility owners, and a wide range of other users with a need for intermodal and passenger transportation system data.

**Accessibility:** The database is accessible to the public and is available for download on the TranStats website.[89]

**Contact:** TranStats Customer Support, Phone: (800)853-1351, Email: ritainfo@dot.gov.

***National Transit Database***

**Data Owner:** Federal Transit Administration

**Description:** The National Transit Database (NTD) is the Federal Transit Administration's (FTA) national database of statistics for the transit industry. The NTD is comprised of data reported by more than 600 transit agencies across the U.S., which is then analyzed and compiled into reports published by FTA and made available to the public on the NTD Program website.

**Purpose:** The purpose of the NTD is to maintain a reporting system, using uniform categories to accumulate public transportation financial and operating information, to help meet the needs of individual public transportation systems, the United States Government, State and local governments, and the public for making public sector investment decisions.

**Data Source Hierarchy:** The NTD is classified as a raw data source, while the NTD data tables are classified as a report.

**Geographic Scope/Coverage:** The database includes all modes of public transportation utilized on local and regional routes throughout the country, including private and public buses, heavy and light rail, ferryboats and vanpool service, as well as services for senior citizens and persons with disabilities, and taxi services operated under contract to a public transportation agency.

**Format:** The NTD includes the following types of transit data:

- Operational Characteristics – Vehicle revenue hours and miles, unlinked passenger trips and passenger miles, etc.

[88] BTS TranStats Website – Intermodal Passenger Connectivity Database Data Dictionary, http://www.transtats.bts.gov/Fields.asp?table_id=1180.

[89] BTS TranStats Website – Intermodal Passenger Connectivity Database, http://www.transtats.bts.gov/DL_SelectFields.asp?Table_ID=1180&DB_Short_Name=Transnet.

- Service Characteristics – Service reliability and safety, etc.
- Capital Revenues and Assets – Sources and uses of capital, fleet size and age, and fixed guideways, etc.
- Financial Operating Statistics – Revenues, Federal, State and local funding, costs, etc.

NTD data is available as an annual database (self extracting XLS file), annual data tables (self extracting XLS file), or monthly database tables (self extracting XLS file). The NTD website also provides a search tool to view annual transit profiles for a specific agency based on the agency's NTD ID, Agency Name, City, or State.

**System Requirements:** The database is available online, and it can be downloaded in XLS format for use within Microsoft Excel.

**Data Quality:** Once the monthly data is submitted, it is subjected to review and analysis for data completeness and reasonableness. Additionally, transit authorities may revise their data at any time during a calendar year's reporting cycle, which lasts through March 1 of the subsequent year. These changes may be done unilaterally by the transit authority, as it collects additional data on its operations. These changes would be reflected in subsequent release of the monthly data base.

**Metadata:** A data dictionary shows the datasets included in NTAD. The data dictionary for each year is provided through the website below.

**Timeliness:** After the close of a month, transit agencies have one month to compile and submit data to the NTD. Therefore, there is a 30 days lag from the end of the month to the time the data is submitted. However, some data may be missing for the most recent month if a transit agency failed to report data on time. Annual databases are currently available from 1996 to 2010.

**Longevity/Future Plans:** Reporting requirements change from year to year. Some agencies have stopped reporting data or gone in their own direction because of lack of guidance and standards from FTA. As such, FTA is currently working to improve the quality of NTD data and make it more acceptable.

**Users:** Users include individual public transportation systems, Federal, State and local governments, and the public. NTD data is used for public transportation planning service, the apportionment, Conditions & Performance Report, Government Performance and Results Act Reporting.

**Accessibility:** NTD data is accessible to the public and is available on the NTD Program website.[90]

**Contact:** John D. Giorgis, NTD Program Manager, Federal Transit Administration, 202-366-5430, john.giorgis@dot.gov

***Cross-Town Improvement Project (C-TIP)***

**Data Owner:** Federal Highway Administration - Office of Freight Management

[90] National Transit Data Program Website, http://www.ntdprogram.gov/ntdprogram/ data.htm

**Description:** The Cross-Town Improvement Project (C-TIP) began as the Intermodal Freight Technology Working Group (IFTWG) with a high-level concept that incorporated an intermodal move database for coordinating cross-town traffic to reduce movement of empty containers between terminals. However, it has grown to include the following components:[91]

- Intermodal Move Exchange (IMEX) - facilitates the exchange of load data and availability information between railroads, terminal operators and trucking companies. Its primary function enables collaboration of pickup and delivery schedules by locations in order to maximize the potential for linking moves, thus eliminating bobtail and empty moves.
- Wireless Drayage Updating (WDU) - provides a means to wirelessly and inexpensively exchange information among drivers regarding trip assignments, traffic congestion information, trip status, and location information via a truck-mounted driver interface device (T-MDID).
- Chassis Utilization Tracking (CUT) - provides a means for chassis owners and users to accurately account for asset usage, which is crucial for the allotment of fees, and in order to maintain the proper balance of chassis to support cross-town and other container deliveries.
- Real-Time Traffic Monitoring (RTTM) - provides a means for up-to-the minute information regarding roadway conditions, travel speeds, and predicted travel times to be captured and passed along to the trucking community. Using a combination of traditional roadway sensors, traffic probes (i.e., vehicles that report their progress while traveling on the roadways), and third party providers, RTTM provides the traffic information necessary for drivers and dispatchers to make informed decisions regarding routing and departure times.
- Dynamic Route Guidance (DRG) - the dynamic route guidance (DRG) engine of RTTM will utilize inputs from RTTM and a Geographic Information System (GIS) source along with simulation tools, to act as an intelligence tool to provide real-time visual routing around congested areas.
- Interchange Capacity Management (ICM) – A combination of a simulation tool and a terminal management system that utilizes schedules to better manage container storage and retrieval.

**Purpose:** The purpose of C-TIP is to develop and deploy an information sharing/transfer capability that enables the coordination of moves between parties to maximize loaded moves and minimize unproductive moves in the Intermodal Transportation Network.

**Data Source Hierarchy:** C-TIP is classified as a system.

**Geographic Scope/Coverage:** A Pilot Deployment has been conducted in Kansas City, and the FHWA Office of Freight Management is exploring the C-TIP concept in other environments, including Memphis, Chicago and Los Angeles.

[91] Delcan, SAIC. Cross-Town Improvement Project Concept of Operations. Prepared for the Intermodal Freight Technology Working Group, March 2007.

**Format:** C-TIP is largely about capturing data, formulating information, and distributing it to individuals and organizations that can decipher and act upon it. It is, in its final form, a decision support system. It is based on the premise that, given accurate, timely information, decision makers—industry representatives, independent users or government representatives—are empowered to make choices about behavior that will enhance transportation efficiency, safety, and security within the cross-town environment.

**System Requirements:** TBD

**Data Quality:** TBD

**Metadata:** TBD

**Timeliness:** TBD

**Longevity/Future Plans:** Future work should leverage recent private sector advances in traveler information (e.g., smart phone-based dynamic routing).

**Users:** Cross-Town Improvement Project (C-TIP) data capture efforts will capture information on personal vehicle data, such as travel times and speeds but the information will be gathered for informing the freight industry stakeholders.

**Accessibility:** TBD

**Contact:** Randy Butler, Office of Freight Management – FHWA, 202-366-9215, Randy.butler@dot.gov; Ron Schaefer, SAIC, 618-257-8435, schaeferrl@saic.com

***FHWA Freight Performance Measures Initiative***

**Data Owner:** The FHWA Office of Freight Management and Operations is the Sponsor, but the American Transportation Research Institute (ATRI) administrates the program and is the data owner.

**Description:** The FPM initiative includes a data processing tool that estimates average operating speeds for trucks traveling along the interstate. Speeds are calculated using confidential, onboard data from several hundred thousand trucks including periodic time, location, speed, and anonymous unique identification information. This data is used to produce:[92]

- Average speed, travel time and reliability of truck movement on large transportation networks such as the Interstate Highway System.
- Quantification and ranking of highway bottlenecks, urban congestion and localized system deficiencies on the nation's freight transportation system.
- Crossing time and delay statistics at freight significant U.S./Canadian border crossings.
- Information describing demand for truck routes and highway facilities throughout the U.S.

[92] American Transportation Research Institute website, Freight Performance Measures Program page, http://atri-online.org/2012/02/28/freight-performance-measures/

The goal of the project is to determine how effectively the surface transportation network is accommodating the increasing demand for reliable freight movement and meeting the demands and expectations of its users. Freight performance measures are at Tier 1 and 2.

Performance measurements are produced for this program through the use of real, anonymous, private sector truck data sourced through unique industry partnerships. FHWA has a partnership with ATRI, who contracts and has agreements with satellite vendors and truck fleet operators to collect truck data. ATRI cleans and processes the data, providing FHWA with clean data containing a unique truck identifier (anonymous), latitude/longitude data, and a time/date stamp. The travel time and buffer index are calculated, which are used for the leadership dashboard and DOT scorecard. The buffer index is reported for city pairs, not along the entire route.

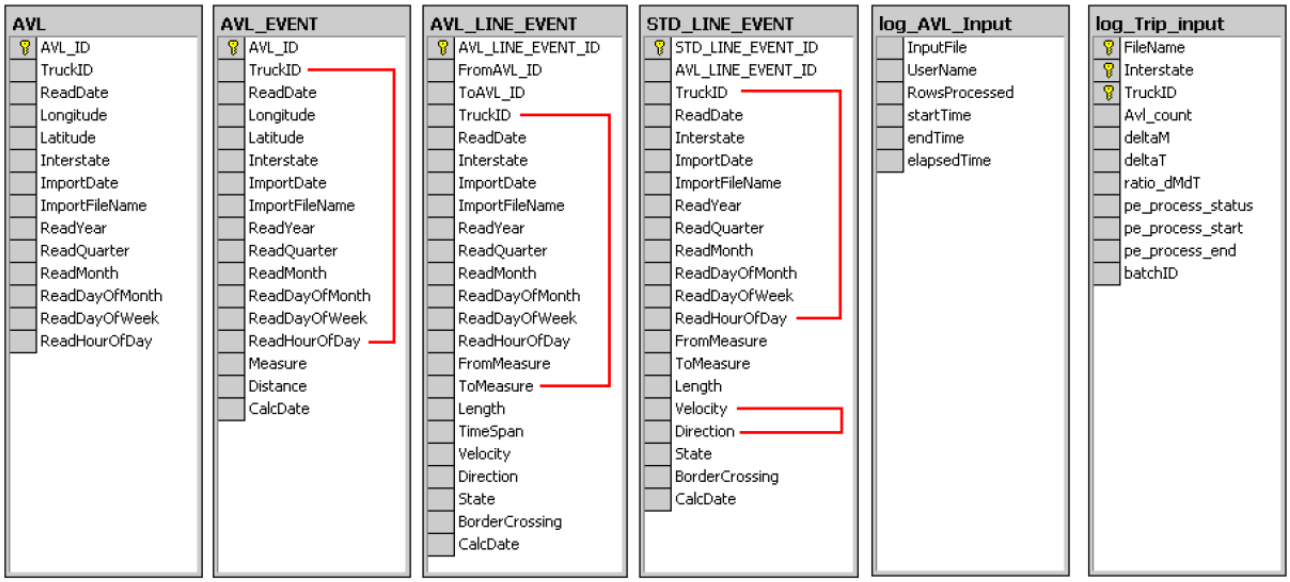
**Purpose:** Transportation data analysts, researchers and other practitioners can determine where, when and how efficiently trucks are moving on selected interstate highways.

**Data Source Hierarchy:** Although the ATRI data has been cleaned and processed, it can be classified as a raw data source for truck speed and travel time data.

**Geographic Scope/Coverage:** This initiative measures operating speeds for trucks at any given place and point in time along twenty-five (25) interstate highways (I-5, I-10, I-15, I-20, I-24, I-25, I-26, I-35, I-40, I-45, I-55, I-65, I-70, I-75, I-76, I-77, I-80, I-81, I-84, I-85, I-87, I-90, I-91, I-94, and I-95) that are considered significant freight routes. Data can be obtained in three-mile segments.

**Format:** The SQL-based FPM system (FPMS II) has a predefined segment size of 3-miles for average speed calculations (the original FPMS I was GIS-based). A data transformation service (DTS) routine residing in the SQL server is used to streamline the truck GPS data processes. Table 2.8 shows the database structure. Raw AVL data is loaded into the AVL table. Projected data with its corresponding linear measures are stored in the AVL_EVENT table. Average measures are calculated, outliers removed, and stored in AVL_LINE_EVENT. The STD_LINE_EVENT table contains the final performance measure calculation after segmentation. The query results can be exported in a CSV file.

**Figure H.8 FPMS II Database Structure** (Source: Liao, C. Freight Performance Measure Systems (FMS) System Evaluation and Data Analysis: Final Report. University of Minnesota Center for Transportation Studies Report No. CTS-08-01, January 2008, http://conservancy.umn.edu/bitstream/5767/1/CTS08_01.pdf)




**System Requirements:** Users must have access to the SQL-based FPM system (FPMS II) or the FPMweb application to access the data.

**Data Quality:** ATRI cleans the data and processes it. Disqualified AVL data is removed (that which cannot be projected to a particular route, performance outliers, etc.).

**Metadata:** Unknown

**Timeliness:** The FPM database contains data spanning more than 7 years. Users can extract annual, daily, or hourly speed data in three-mile segments. Data is not in real-time since this initiative is intended for quarterly and annual reporting.

**Longevity/Future Plans:** Coverage currently includes the interstate system; expanded usage to include the National Highway System is likely. Re-evaluating is underway as to whether spot speed or space mean speed should be used. Volume data is an addition under consideration, but there is a need to coordinate on data definitions from existing systems, that is exuberated by the fact that location data are not matching up.. Another future consideration is the potential linkages between programs (i.e., incorporation into Freight Analysis Framework (FAF), and identifying additional use for these data).

**Users:** Current customers for freight performance measures are primarily within FHWA. Future customers include States and MPOs.

**Accessibility:** Data from the Freight Performance Measures initiative is available on the FPMweb application.[93] Access is limited, and a user must have an account to log into FPMweb. The website currently operates under a "limited user" phase.

**Contact:** Ed Strocko, 202-366-2997, Ed.Strocko@dot.gov

***Freight Analysis Framework (FAF)***

**Data Owner:** Federal Highway Administration (FHWA) Office of Freight Management and Operations

**Description:** The Freight Analysis Framework (FAF) integrates data from a variety of sources to create a comprehensive picture of freight movement among States and major metropolitan areas by all modes of transportation. With data from the 2007 Commodity Flow Survey and additional sources, FAF version 3 (FAF3) provides tonnage estimates and value, by commodity type, mode, origin, and destination for 2007, the most recent year, and forecasts through 2040. Also included are truck flows assigned to the highway network for 2007 and 2040.[94]

The FAF estimates commodity movements by truck and the volume of long distance trucks over specific highways. Models are used to disaggregate interregional flows from the Regional origin, destination, commodity, and mode database into flows among localities and assign the detailed flows to individual highways. These models are based on geographic distributions of economic activity rather than a detailed understanding of local conditions. While FAF provides reasonable estimates for national and multi-state corridor analyses, FAF estimates are not a substitute for local data to support local planning and project development.

**Purpose:** FAF provides a national picture of current freight flows to, from, and within the United States, assigns selected flows to the transportation network, and projects freight flow patterns into the future. FAF is a tool that allows policy makers to make the right decisions, helping to gauge the significance of freight movement across the entire country.

The FAF freight database has gone through several updates and has been used in a variety of transportation analyses including highway capacity and bottleneck assessments, truck size and weight studies, evaluations of the benefits of strategic investments in transportation infrastructure, impacts of changes in road pricing policies, multimodal freight policy analysis, impacts of toll proposals on shipper choice decisions, and the impact on national freight movement of natural and manmade disasters (e.g., the I-40 bridge collapse in Oklahoma in 2002; the I-95 bridge at Bridgeport, Connecticut in 2004; the impact on freight movement due to Katrina in 2005, the Collapse of the I-35W bridge in Minneapolis in 2007; and others).

**Data Source Hierarchy:** FAF provides a complete picture of freight movement to, from and thru the U.S., and relies on numerous data sources (e.g., Commodity Flow Survey (CFS), Foreign Trade Statistics, Air Trade Data, etc.). Modeling is used to fill in the gaps.

[93] Freight Performance Measures Web Application (FPMweb) website, https://www.freightperformance.org/fpmweb/user_login.aspx

[94] Freight Analysis Framework Version 3 website, http://faf.ornl.gov/fafweb/Default.aspx

While the FAF3 modeling process draws from many data sources, the most important is the CFS. The CFS is conducted every five years as part of the U.S. Economic Census, with major funding for the survey provided by the BTS. Data are collected on all shipments from the surveyed establishment for an entire week in each of the four quarters of the census year.

In FAF3, there are Non-CFS or Out-Of-Scope (OOS) freight flows. Their estimation requires a good deal of data collection and integration into the larger flow matrix generation process. The data sources for these OOS flows are for the most part derived from freight carrier reported data sources. These OOS flows represent some 32% of all U.S. freight movements measured on an annual tonnage basis. Developing OOS flow estimates presents considerable effort, with different commodity classes requiring very different, typically multi-step treatments: including the use of both spatial and commodity class crosswalks that convert mode and industry class specific estimates from their native coding categories into FAF3 regional and commodity class breakdowns.

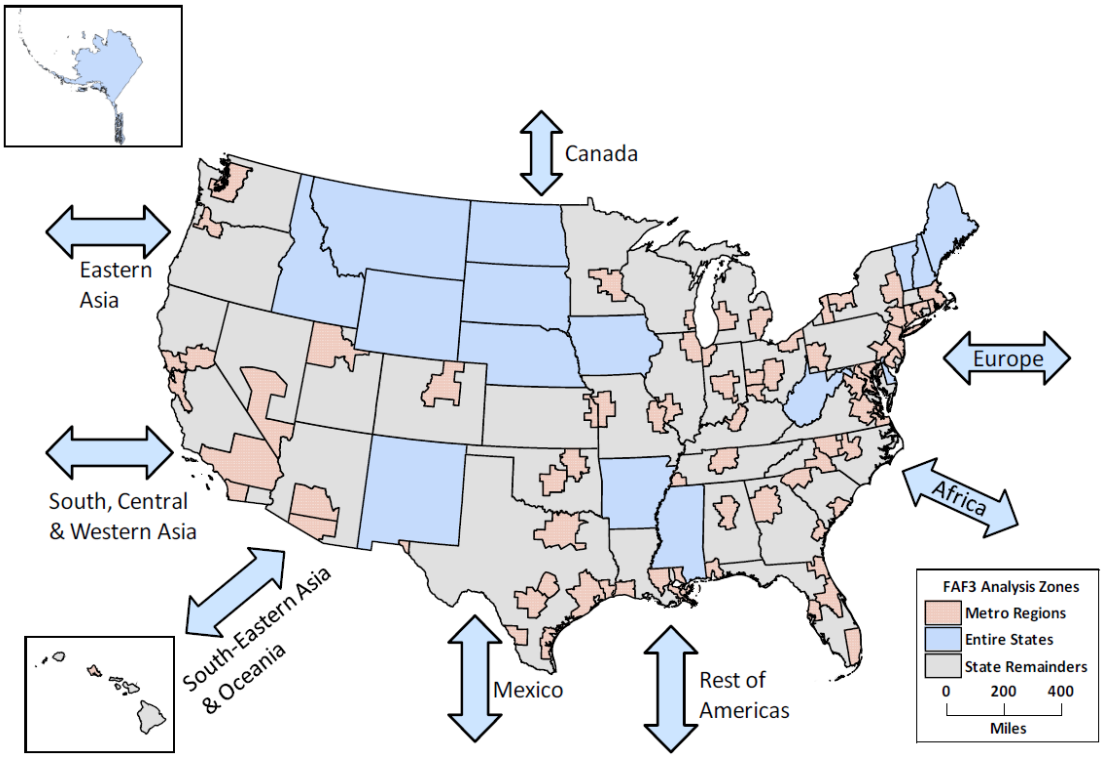
Data sources include the following:[95]

[95] Lambert, B. The Freight Analysis Framework – Overview and Uses Presentation, http://www.slideshare.net/Databaseguys/the-freight-analysis-framework-overview-and-uses.

- BTS/Census Bureau Commodity Flow Survey (All Modes)
- BTS Transborder Freight Statistics (Truck, Rail, Pipeline)
- STB/FRA Rail Waybill Sample (Rail)
- MARAD and US Army Corps data (Water)
- BTS/FAA (Airport)
- Reebie (Truck)
- Census Bureau 5-year Census and Annual Survey of Manufacturers (Truck, Water, Air)
- Private Port Directories (Water)
- DRI Industrial Production Indices (Truck, Water, Air)
- HPMS Database (Truck)
- Truck Configuration estimates (Truck)
- National Highway Network (NHN)
- Agricultural Production Statistics (USDA)
- Trade Association Production and Shipment Reports (Truck, Water, Air)
- US Geological Survey Mineral Industry Reports (Truck, Water)
- Reebie Associates Freight Locator/InfoUSA Street-Address Industrial Employment & Activity (Truck)
- County Population Data (Truck)
- Inter-Industry Trade Patterns (Input/ Output Table) (Truck, Air)
- Motor Carrier Industry Financial & Operating Statistics (Truck)
- Railroad Industry Proprietary Rebill Factors (Truck)
- Vehicle Inventory and Use Survey (Truck)
- Latin American Trade and Transportation Study (LATTS) – All Modes
- State Traffic Count Data (Truck)

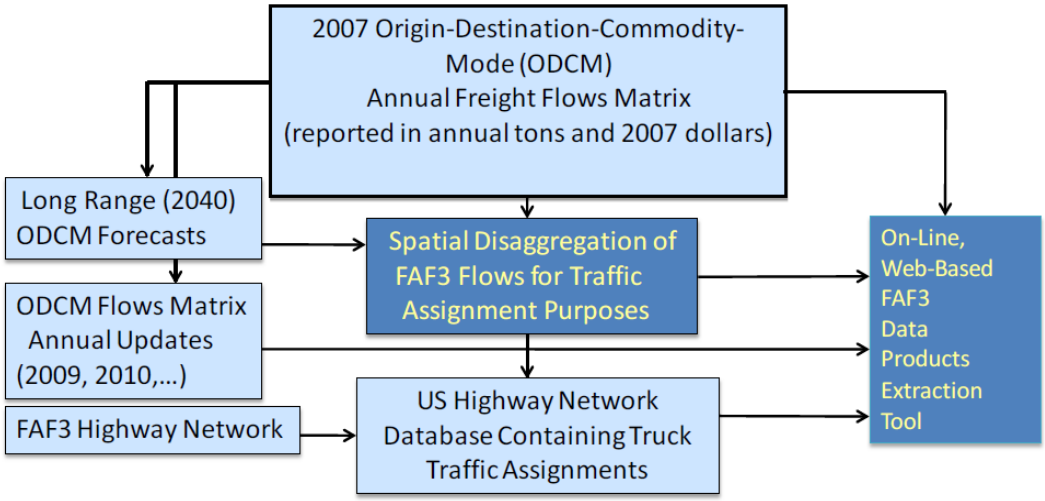
**Geographic Scope/Coverage:** As shown in Figure H.9, the FAF geography is based on 2007 CFS commodity flow tables with 11 additional traffic analysis regions, for a total of 123 domestic regions.

**Figure H.9 FAF Geographic Coverage** (Source: Southworth, F., D. Davidson, H. Hwang, B. Peterson, and S. Chin. The Freight Analysis Framework, Version 3 (FAF3): A Description of the FAF3 Regional Database and How It Is Constructed. Prepared for Federal Highway Administration, June 16, 2011, pg. 4.)




**Format:** The FAF3 data products include both a regional database and network database with highway flow assignment. FAF data includes origin, destination, commodity, and mode data. Their key data products are shown in Figure H.10. The FAF3 freight flows matrix is made up of 131 Origin (O) x 131 Destination (D) x 43 Commodity Class (C) x 8 Modal Category (M) data cells, for each of two reporting metrics, annual tons and annual dollar values.

**Figure H.10** **FAF Data Products** (Source: Southworth, F., D. Davidson, H. Hwang, B. Peterson, and S. Chin. The Freight Analysis Framework, Version 3 (FAF3): A Description of the FAF3 Regional Database and How It Is Constructed. Prepared for Federal Highway Administration, June 16, 2011, pg. 2.)




**System Requirements:** Most of the FAF components are designed for use online. Others involve a spreadsheet (CSV) or database (Microsoft Access) or GIS (ESRI or TransCad) environment.

- FAF3 Regional Databases for 2007, 2010 Provisional Annual Data, and forecasts of 2015 through 2040 in zipped Microsoft Access format [FAF3.2_access03.zip 178MB]
- FAF3 Regional Database for 2007 and forecasts of 2015 through 2040 in zipped CSV format [FAF3.2.zip 116MB]
- FAF3 Provisional Annual Data for 2010 in zipped CSV format [FAF3.2_provisional_2010.zip 16MB]
- FAF3 State Database for 2007, 2010 Provisional Annual Data, and forecasts of 2015 through 2040 in zipped Microsoft Access format [FAF3.2_access03_State.zip 64MB]
- FAF3 State Database for 2007 and forecasts of 2015 through 2040 in zipped CSV format [FAF3.2_State.zip 40MB]
- FAF3 State Provisional Annual Data for 2010 in zipped CSV format [FAF3.2_provisional_2010_State.zip 6MB]
- FAF Output
  - faf3_1_1_data.dbf [29MB]
  - datadict (Data Dictionary) [HTML, PDF 14KB]
- Network
  - Metadata [HTML, PDF 39KB]
  - ESRI Format
    - faf3_1_1_esri.zip [38MB]

    - TransCAD Format
        - faf3_1_1_transcad.zip [35MB]
- FAF3 Regions Boundary Layer
    - FAF3-zone_ESRI.zip [25MB]
    - FAF3-zone_TransCAD.zip [12MB]

**Data Quality:** FAF typically relies on source agencies to supply documentation and metadata (e.g., BTS, BEA, Census Bureau, etc.). Most of the datasets used by FAF are publicly available information, so the statistical validation of data has typically already been done. However, they all have their data quality issues. For example, the CFS involves data suppression (represented with a "S" in the cell) for a combination of reasons including confidentiality, limited sample size, large standard errors, and limitations to the scope of the CFS (across industrial sectors). Another gap related to freight data quality is the need to make HPMS data more reliable.

**Metadata:** FAF relies on source agencies to supply documentation and metadata (e.g., Bureau of Transportation Statistics (BTS), BEA, Census Bureau, etc.). The FAF documentation includes extensive information on the sources and processing of the data. Metadata on the FAF website include:

- FAF3 variables in Microsoft Excel format
- FAF3 variables in HTML format
- Network data
    - http://faf.ornl.gov/fafweb/networkdata/FAF3_MetaData.htm

**Timeliness:** Data is updated annually (not real-time).

**Longevity/Future Plans:** FAF relies on HCM methods and formulas to estimate travel time based on volume and speeds. As such, FHWA is working to incorporate average speed data from FPM to better project travel times, as well as other probe data. FAF3 is nearing the end of the current production cycle, and FAF4 will start next year. Data will be collected in 2012, but it will not likely be processed and available within FAF until 2014 or later.

**Users:** FHWA, other DOT offices (FMCSA, MARAD), State DOTs, MPOs, multi-state coalitions, trade associations, transportation providers, other Federal government users (DOD/Homeland Security – Commerce, USDA, INS, etc.).

**Accessibility:** FAF data products can be downloaded at the FAF3 website including: Data Extraction Tool,[96] National Freight Transportation Maps,[97] and Highway Network Data Files.[98] Users can also

[96] FAF3 Data Extraction Tool website, http://faf.ornl.gov/fafweb/Extraction0.aspx.
[97] FAF3 National Freight Transportation Maps, http://faf.ornl.gov/fafweb/dataMap.aspx.
[98] FAF3 Highway Network Data Files, http://faf.ornl.gov/fafweb/networkdata.aspx.

download the entire FAF3 2007 and 2040 flows database in either Microsoft Access 2003 or CSV format.[99]

**Contact:** Michael Sprung, 202-366-9047, Michael.Sprung@dot.gov

***Motor Carrier Management Information System (MCMIS)***

**Data Owner:** Federal Motor Carrier Safety Administration (FMCSA)

**Description:** The FMCSA maintains the Motor Carrier Management Information System (MCMIS). MCMIS contains information on the safety fitness of commercial motor carriers (truck and bus) and hazardous material (HM) shippers subject to the Federal Motor Carrier Safety Regulations (FMCSR) and the Hazardous Materials Regulations (HMR). There are four elements:

1. Crash data. This data comes from States, who report on all crashes involving a large truck or bus. Crash data is reported for truck crashes involving a fatality (based on the same definition as FARS), injury (where someone was transported from the scene), or a tow-away (where a vehicle towed away due to disabling damage).
2. Carrier Registration (or census file). All motor carriers operating in Interstate commerce, all hazardous carriers, or passenger carriers need to register with FMCSA to secure registration numbers and provide information about their company (e.g., number of units, type of cargo, etc.).
3. Inspection data. This data can come from States, state police officers doing roadside inspection, weigh stations, FMCSA field inspectors, and roadside audits. FMCSA receives approximately 4 million records per year.
4. Company Safety Profile (CSP) reports. FMCSA uses MCMIS data to score carriers for their safety status. CSP reports contain composite extracts regarding a specified carrier from the MCMIS Census, Review, Enforcement, Crash, and Inspection Files -- the last includes the results of roadside driver-vehicle inspections conducted by Federal and State safety agencies. Because it combines information from many sources, the CSP is the most complete and in-depth report on an individual carrier available.

**Purpose:** The Federal Motor Carrier Safety Administration (FMCSA), within the Department of Transportation (DOT), has been given the responsibility to reduce crashes, injuries, and fatalities involving large trucks and buses. In carrying out its safety mandate, the FMCSA:

- Develops and enforces data-driven regulations that balance motor carrier (truck and bus companies) safety with industry efficiency;
- Harnesses safety information systems to focus on higher risk carriers in enforcing the safety regulations; and
- Targets educational messages to carriers, commercial drivers, and the public.

To meet these goals, the FMCSA partners with stakeholders, including Federal, State, and local enforcement agencies, the motor carrier industry, safety groups, and organized labor on efforts to reduce bus and truck-related crashes. Since a first step to reduce accidents is to understand them,

[99] FAF3 2007 and 2040 flows database website, http://ops.fhwa.dot.gov/freight/freight_analysis/faf/index.htm.

FMCSA collects and maintains commercial vehicle safety data, as well as a national inventory of motor carriers and shippers subject to the Federal Motor Carrier Safety Regulations and Hazardous Materials Regulations.

**Data Source Hierarchy:** MCMIS is classified as a system in the data source hierarchy.

**Geographic Scope/Coverage:** MCMIS coverage is national.

**Format:** The FMCSA Crash File contains data from State police crash reports involving drivers and vehicles of motor carriers operating in the U.S. Each report contains about 80 data elements pertaining to the motor carrier, driver, vehicles, and circumstances of a crash.

MCMIS Census File contains records for a steadily growing number of active entities, i.e., motor carriers, hazardous materials shippers, entities that are carrier and a shipper, or registrants (entities who register vehicles but are not carriers). Data contained in the MCMIS Census File originates from the following sources:

- Federal Form MCS-150. A form used by the Federal Motor Carrier Safety Administration to collect general identification data from entities.
- Federal Form MCS-151 Part A. The first page of the Safety/Compliance/Educational Contact Review. This page provides identifying data about the carrier/shipper (i.e., name, address, number of trucks, etc.). A review contains information regarding a Federally-conducted or State-conducted review of a carrier's/shipper's compliance with the FMCSR and HMR. These reviews result in the carrier receiving a Federal safety rating of satisfactory, unsatisfactory or conditional. A safety rating is generated by MCMIS after the review has been added to the MCMIS Safety/Compliance/Educational Contact Review File.
- Licensing and Insurance Division (Former Interstate Commerce Commission). Applications for Permanent, Emergency and Temporary Operating Authority.
- Other Users of MCMIS. Department of Defense, American Trucking Association, etc.
- SAFETYNET. State Inspection and Accident Reports.
- Telephone contacts and correspondence concerning entities.
- Federal Form MCS-63. Federally-conducted Driver/Vehicle Inspection Reports.

The FMCSA Inspection File contains data from State and Federal inspection actions involving motor carriers, shippers of hazardous materials, and transporters of hazardous materials operating in the U.S. The majority of the inspections are conducted at the roadside by State personnel under the Motor Carrier Safety Assistance Program (MCSAP). Severe violations of the FMCSR may result in a vehicle and/or driver being placed "out-of-service."

The CSP provides Federal and State investigative personnel with national information on interstate carriers and some intrastate carriers. This information is used to assist in carrying out a specific action on the carrier. For example, a CSP may be used as an adjunct to a Federally-conducted or State-conducted compliance review, a complaint investigation, an enforcement action or any similar investigation. Data contained in the CSP reports are obtained through various sources:

- Motor Carrier Identification Report (MCS-150).
- Motor Carrier Safety/Compliance Reviews and Educational contacts (MCS-151, Parts A-D).
- Enforcement Reports (MCS-32A).
- State-contributed crash reports.
- Motor Carrier Driver/Vehicle Inspection Report (MCS-63) and State-submitted inspection reports.

**System Requirements:** The crash data collected are entered into a microcomputer-based system called SAFETYNET, which allows States to do analysis on all motor carriers in the State (interstate and intrastate) and to transmit these data to MCMIS via telephone lines.

**Data Quality:** The MCMIS system provides internal data edit checks on all data submitted to MCMIS. FMCSA data entry contractors conduct a verification process to ensure that accurate information is entered in MCMIS. The Federal and State system where the crash, inspection and compliance review data are entered contains data quality edit checks before the data is submitted to MCMIS.

**Metadata:** The following metadata documentation is available: crash file data elements,[100] census file data elements,[101] inspection file data elements,[102] and company safety profile data elements.[103]

**Timeliness:** Crash data is reported within 90 days, and inspection data is reported within 10 days. Some States have automated systems, so FMCSA is getting their data much quicker. Currently, the state-reported crash file elements are available from years 1989 to present.

**Longevity/Future Plans:** FMCSA will continue to operate MCMIS for the foreseeable future.

**Users:** MCMIS data are used by: 1) State agencies for targeting motor carrier safety enforcement and for developing safety programs; 2) safety organizations to evaluate safety trends, promote safety programs, and evaluate the effectiveness of existing and proposed safety guidelines, enforcement standards, and rules, 3) insurance companies for evaluating potential clients; and 4) the general public to choose safe companies for household moving and bus transportation.

**Accessibility:** Most information is available to the public through the MCMIS Data Dissemination Program.[104] Personally identifiable information (PII) can only be accessed by designated individuals, through a written request that the FMCSA Information System Division Office reviews and approves. Responsibility for disseminating the data, and the authority to charge prescribed fees for the data, have been delegated to a government contractor, Computing Technologies, Inc.

[100] MCMIS Crash File data elements website, http://mcmiscatalog.fmcsa.dot.gov/ d_crash2.asp
[101] MCMIS Census file data elements website, http://mcmiscatalog.fmcsa.dot.gov/ d_census_recordLayout.asp.
[102] MCMIS Inspection file data elements website, http://mcmiscatalog.fmcsa.dot.gov/ d_inspection2.asp.
[103] MCMIS Company Safety Profile data elements website, http://mcmiscatalog.fmcsa.dot.gov/d_inspection2.asp.
[104] MCMIS Catalog and Documentation website, http://mcmiscatalog.fmcsa.dot.gov.

**Contact:** Bill Bannister, William.Bannister@dot.gov

***Licensing and Insurance (L&I) Database***

**Data Owner:** Federal Motor Carrier Safety Administration

**Description:** For-hire and passenger carriers are required to have minimum levels of liability insurance, and this information is collected through the Licensing and Insurance (L&I) Database.

**Purpose:** The L&I database is designed to assist persons who are interested in interstate surface transportation businesses.

**Data Source Hierarchy:** The L&I database is classified as a system that provides carrier licensing and insurance information.

**Geographic Scope/Coverage:** National

**Format:** The L&I database has on-line tools for searching blanket companies and their process agents, carrier search to access information about interstate for hire companies, FMCSA register search to access all decisions and notices for a specific date, insurance filing option to view file certificates and cancellations and print confirmation, acceptance, and reject reports, Mexican application status to access information about carriers with OP-1[MX] or OP-2 applications that are currently being processed, and out of service (OOS) carriers that are currently placed OOS or who have been placed OOS and had the order rescinded within the last 15 days. Data format varies based on the type of search conducted, but are generally available as an HTML table or PDF report.

**System Requirements:** The L&I database is available online.

**Data Quality:** Unknown

**Metadata:** The FMCSA Registration and Licensing Forms website provides links to common registration and licensing forms, along with instructions and definitions for data elements.[105]

**Timeliness:** The biennial update rule requires that all motor carriers file a biennial update every two years.

**Longevity/Future Plans:** In terms of connected vehicle applications, if someone wanted information about a vehicle on the roadway (e.g., information on the vehicle or its company), it is possible to use the L&I database to look up the information.

**Users:** Users include for-hire and passenger carriers, process agents, brokers, representatives of insurance companies, surety companies, and financial institutions, and other users with a need for L&I information.

**Accessibility:** Some L&I information is accessible to the public and is available on the public page of the L&I public facing website.[106] Representatives of insurance companies, surety companies, and

[105] FMCSA Registration and Licensing Forms website, http://www.fmcsa.dot.gov/ forms/print/r-l-forms.htm.

financial institutions can register with the FMCSA as an electronic insurance filer to view file certifications and cancellations, last transmissions, and print confirmation, acceptance, and reject reports through the L&I Internal Information System.[107] Insurance filers can request a user name and password by calling (202) 385-2423 M-F 8:00 a.m. - 4:30 p.m. EST or (202) 385-2400 after hours or holidays to leave a message.

**Contact:** 1-800-832-5660

***International Freight Data System***

**Data Owner:** RITA Bureau of Transportation Statistics

**Description:** The International Freight Data System (IFDS) is a Federal government information technology initiative to coordinate, standardize, and ultimately simplify the Federal border clearance and other international trade and transportation processes. It is a jointly shared database that will interface with the U.S. Customs and Border Protection's (CBP) Automated Commercial Environment (ACE) program.

The IFDS examines transportation trade data and how long it takes goods to move between origin-destination pairs. It will provide a one-stop shop for international trade data for all DOT modal administrations. The system is in development and is scheduled to come online in 2012.

**Purpose:** The goal of the IFDS is to create and implement an integrated Federal system for the electronic collection, use, and sharing of international trade and transportation data. Implementation of the IFDS will bring DOT into compliance with the requirements of Section 405 of the Security and Accountability for Every Port Act (Pub.L. 109-347, 10/13/06), which mandates participation by all Federal agencies that require documentation for clearing or licensing the importation or exportation of cargo.

**Data Source Hierarchy:** The IFDS will be classified as a system.

**Geographic Scope/Coverage:** National border clearance and international trade data with all countries.

**Format:** In addition to transactional-level shipment and conveyance information on these movements into, from or across the United States, the IFDS will contain the following information.[108]

- Corporate and individual names and addresses, and corporate contact information including corporate telephone numbers and titles;
- License numbers and the issuing entities;

[106] FMCSA L&I Public website, http://li-public.fmcsa.dot.gov/LIVIEW/pkg_html. prc_limain

[107] Licensing & Information Internal Information System website, http://li-public.fmcsa.dot.gov/LIVIEW/pkg_menu.prc_menu

[108] RITA, Privacy Impact Assessment: International Freight Data System (IFDS), Interfacing with U.S. Customs and Border Protection's Automated Commercial Environment and International Trade Data System (ACE/ITDS), February 10, 2009, http://www.dot.gov/citizens/privacy/pia-international-freight-data-system-ifds.

- Electronic signatures and signatories' name, title, related information and date of signature;
- Travel document numbers and issuing countries;
- Nationalities of commercial parties (e.g., conveyance owners and/or operators);
- Dates of births for commercial operators and/or crew members; and
- Corporate and individual license/registration/certificate numbers and types of licenses/ registrations/certificates.

The data will be in a similar format to that available in the Transborder Freight Data (e.g., monthly and annual data, by country, by port district, by commodity, by mode). The exact data elements to be released have not been determined.

**System Requirements:** Unknown

**Data Quality:** Because IFDS will not collect data directly from the public or corporations, data cannot be independently verified for accuracy. For data that are downloaded from ACE, the IFDS will deploy normal data edits and verification processes (currently under development) to help ensure and improve the accuracy of the information that is received. IFDS also will receive periodic updates of the ACE information; this information will be forwarded to participating OAs, so that their databases can be refreshed.

**Metadata:** Unknown

**Timeliness:** For DOT, information from ACE will be made available automatically on a monthly, weekly, or daily basis into the IFDS for sharing among the participating DOT OAs. The information will improve the ability of certain participating OAs to carry out their enforcement responsibilities for:

- The safety of international shipments of hazardous materials; and
- The compliance of imported motor vehicles and motor vehicle equipment with applicable safety standards.

IFDS has an application pending with the National Records and Archives Administration. If the application is approved, IFDS plans to keep information stored in the database for 6 to 10 years before destroying.

**Longevity/Future Plans:** Unknown

**Users:** DOT's Research and Innovative Technology Administration (RITA) is leading the development and management of the IFDS interface. Other participating Operating Administrations (OAs) include: the Maritime Administration (MARAD), the Federal Aviation Administration (FAA), the Pipeline and Hazardous Materials Safety Administration (PHMSA), the Federal Motor Carrier Safety Administration (FMCSA), and the Federal Highway Administration (FHWA).

**Accessibility:** Access (review and use the information) to the IFDS will be limited to the statutory authority of each DOT OA. Information will be automatically and securely downloaded from ACE/ITDS to the IFDS database managed by RITA. The IFDS database will then make the information available

to participating DOT OAs. Only those DOT OAs that have statutory authority to view the data in order to execute their international missions will be given access to the IFDS.

**Contact:** Long Nguyen, RITA BTS, long.nguyen@dot.gov

## Detailed Inventory – Travel Behavior Data

***National Household Travel Survey***

**Data Owner:** RITA Bureau of Transportation Statistics

**Description:** The NHTS is the authoritative source of national data on the travel behavior of the American public. The NHTS, previously called the NPTS (National Personal Travel Survey), has been conducted since 1969 and is the nation's inventory of daily and long-distance travel. The dataset allows analysis of daily travel by all modes, including characteristics of the people traveling, their household, and their vehicles. The NHTS can be used to study travel forecasting, policy implications and travel behavior trends.

**Purpose:** The National Household Travel Survey (NHTS) provides information to assist transportation planners and policy makers who need comprehensive data on travel and transportation patterns in the United States. NHTS data are used to:

- Quantify travel behavior,
- Analyze changes in travel characteristics over time,
- Relate travel behavior to the demographics of the traveler, and
- Study the relationship of demographics and travel over time.

The NHTS data are used primarily for gaining a better understanding of travel behavior. The data enable DOT officials to assess program initiatives, review programs and policies, study current mobility issues, and plan for the future.

The NHTS is a tool in the urban transportation planning process; it provides data on personal travel behavior, trends in travel over time, trip generation rates, national data to use as a benchmark in reviewing local data, and data for various other planning and modeling applications.

**Data Source Hierarchy:** The NHTS is classified as a source of raw data on travel behavior. In addition, there is a summary report on travel trends based on NHTS data.

**Geographic Scope/Coverage:** Data for the survey are from 150,147 randomly selected landline telephone households. An additional small sample of 1,254 cell phone only households was conducted as a test of methods and responses, but data from this group was not incorporated into the final dataset.

The NHTS serves as the nation's inventory of daily travel. Data is collected on daily trips taken in a 24-hour period, and includes:

Purpose of the trip (work, shopping, etc.);

- Means of transportation used (car, bus, subway, walk, etc.);
- How long the trip took, i.e., travel time;
- Time of day when the trip took place;
- Day of week when the trip took place; and
- If a private vehicle trip:
    - Number of people in the vehicle , i.e., vehicle occupancy;
    - Driver characteristics (age, sex, worker status, education level, etc.); and
    - Vehicle attributes (make, model, model year, amount of miles driven in a year).

These data are collected for:

- All trips;
- All modes;
- All purposes;
- All trip lengths; and
- All areas of the country, (i.e. both urban and rural).

**Format:** Publicly available datasets are available in SAS Windows Binary, SAS Transport, Dbase, and ASCII format. A roster file and trip chaining datasets are available in SAS Windows binary format, SAS Proc Format. The datasets are a microdata which contains a record of each interview (with identifying information deleted).

The 2009 NHTS data set includes, but is not limited to:

- Household data on the relationship of household members, education level, income, housing characteristics, and other demographic information;
- Information on each household vehicle, including year, make, model, and estimates of annual miles traveled;
- Data about drivers, including information about work related travels;
- Data about one-way trips taken during a designated 24-hour period (the household's travel day) including the time the trip began and ended, length of the trip, composition of the travel party, mode of transportation, purpose of the trip, and the specific vehicle used (if a household vehicle);
- Information to describe characteristics of the geographic area in which the sample household and workplace of sample persons are located;
- Data on telecommuting;
- Public perceptions of the transportation system;
- Data on Internet usage; and

- The typical number of transit, walk and bike trips made over a period longer than the 24-hour travel day.

A summary of content is provided in Figure H-11.

**System Requirements:** The datasets are designed for use within any statistical software package such as Statistical Analysis Software (SAS) or SPSS predictive analytics software. The data is also stored in Dbase and ASCII format so it can be used with any database or spreadsheet software. There is also an Online Analysis Engine on the website that allows users to create customized tables from the NHTS data without having to download the dataset and code the table specifications.

**Data Quality:** Data quality for the NHTS is conducted by performing a series of edits designed to check for data consistency and logic. These edits are performed at various stages of data collection and include both programmed edits and manual reviews. Detailed procedures are documented in the NHTS User's Guide.[109]

**Metadata:** A data dictionary describes the data fields associated with the NHTS datasets.

**Timeliness:** Surveys are conducted approximately every 5 to 7 years. The surveys were conducted in 1969, 1977, 1983, 1990, 1995, 2001 and 2009.

[109] 2009 National Household Travel Survey User's Guide (Version 1), U.S. DOT Federal Highway Administration, February 2011, http://nhts.ornl.gov/2009/pub/ UsersGuideV2.pdf

**Figure H-11 2009 NHTS Summary of Content** (Source: FHWA, 2009 National Household Travel Survey User's Guide, Version 2, Updated October 2011. http://nhts.ornl.gov/2009/pub/UsersGuideV2.pdf)

| 2009 NHTS summary of content | |
|---|---|
| **For Each Household:** | **For Each Vehicle:** |
| *Number of people, drivers, workers and vehicles* | *Make/Model/Age (year)* |
| *Income* | *Annual miles driven* |
| Housing Type | Commercially licensed** |
| Owned or rented | How long owned* |
| Number of cell phones* | Odometer reading |
| Number of other phones | Alternative Fuel** |
| *Race of reference person* | Primary Driver |
| *Hispanic status of reference person* | |
| Tract and Block Group characteristics | **Daily Travel Data:** |
| Internet Use & Delivery to households** | Origin and Destination address (for Add-ons) |
| | *Time trip started and ended* |
| **For Each Person:** | *Distance* |
| *Age/Sex/Relation to reference person* | *Means of transportation:* |
| *Driver status* | *vehicle type* |
| *Worker status/Primary activity* | *if household vehicle, which one* |
| Internet use* | if transit, wait time |
| Home deliveries from Internet shopping** | if transit, access and egress mode* |
| Travel Disability* | Interstate Use** |
| Effect of disability on mobility* | Tolls Paid** |
| *Education level* | *Trip Purpose* |
| Immigrant status* | Detailed purpose* |
| Views on transportation | *Travel Party Size* |
| *Annual miles driven* | Last time of travel* |
| Incidence of public transit use in past month | |
| Incidence of motorcycle use in last month | |
| Incidence of walk and bike trips in past week | |
| School travel (children)** | |
| **For Each Worker:** | |
| *Full or part-time* | |
| More than one job | |
| Occupation (four categories)* | |
| *Workplace location* | |
| *Usual mode to work* | |
| *Drive alone or Carpool* | |
| *Usual distance to work* | |
| Usual time to work** | |
| Work from home | |
| Usual arrival time at work | |
| Flexibility in work arrival time** | |

* added in 2001

** added in 2009

**Longevity/Future Plans:** The future of the NHTS is uncertain. For the 2009 update, there was a threat of withdrawal of BTS support, and the Federal contribution - $3 million out of $24 million – was a small portion of the total cost. The bulk of the funding was provided by State and MPO add-ons. The uneven cycle of updating the NHTS further contributes to its uncertainty, particularly since it is dependent on funding from a multitude of sources.[110]

**Users:** The transportation research community, including academics, consultants and government, use the NHTS extensively to examine:

- Travel behavior at the individual and household level;
- The characteristics of travel, such as trip chaining, use of the various modes, amount and purpose of travel by time of day and day of week, vehicle occupancy, and a host of other attributes;
- The relationship between demographics and travel; and
- The public's perceptions of the transportation system.

People in various fields outside of transportation use the NHTS data to connect the role of transportation with other aspects of our lives. Medical researchers use the data to determine crash exposure rates of drivers and passengers, including the elderly, who have heightened morbidity and mortality rates. Safety specialists study the accident risk of school-age children, particularly when they are traveling on their own by walking or biking. Social service agencies need to know more about how low-income households currently meet their travel needs.

**Accessibility:** All of the data for 1969, 1977, 1983, 1990, 1995, 2001 and 2008 are accessible to the public and are available for download on the NHTS website.[111] The website serves as a central facility to obtain information on the NHTS, status updates on the dataset or publications, publications using NHTS data, and the disaggregate NHTS data which is available for download and analysis by the user community.

**Contact:** Adella Santos, NHTS Program Manager, 202-366-5021, adella.santos@dot.gov.

***Omnibus Household Survey***

**Data Owner:** RITA Bureau of Transportation Statistics

**Description:** The Omnibus Household Survey (OHS) is conducted to collect information about specific transportation issues (e.g., cell phones and distracted driving) and to assess public satisfaction and attitudes about the transportation system and its interaction with DOT agencies (e.g., airline passenger opinions on security screening procedures).

**Purpose:** The OHS serves as an information source for the U.S. DOT modal administrators to support congressional requests and to gauge internal DOT performance. Overall, OHS supports the collection of information on a wide range of transportation-related topics.

[110] How We Travel: A Sustainable National Program for Travel Data, Transportation Research Board Special Report 304, 2011.
[111] NHTS website, http://nhts.ornl.gov/

**Data Source Hierarchy:** The OHS is classified as a source of raw data on traveler satisfaction with the transportation system, while the OmniStats publications based on OHS data are classified as a report.

**Geographic Scope/Coverage:** Data for the survey are from approximately 1,000 randomly selected telephone households and are weighted to allow inferences about the non-institutionalized population aged 18 years or older who is currently living in the United States.

**Format:** OHS data is distributed in Excel, comma-separated values (CSV), and Statistical Analysis Software (SAS) format.

**System Requirements:** The datasets are designed for use within Statistical Analysis Software (SAS). The data is also stored in CSV and Excel format so it can be used with any database or spreadsheet software.

**Data Quality:** Standard procedures are implemented for data quality control. Data is reviewed and examined by senior analysts for outliers, entry errors, or missing variables. Each variable is examined to ensure that responses fall within expected parameters. Potentially invalid responses and outliers are further investigated. Variables are crosschecked to each other to ensure internal consistency between responses to interrelated variables.

When inconsistencies or outliers are found, related call sheet logs, notes and the actual recordings of the interview in question are reviewed to determine if data had been incorrectly interpreted or entered by the interviewer. If the survey is still in the field, callbacks are made by supervisors to respondents for cases that could not be reconciled through a review of the logs or recordings. Once the survey interval ends, these cases are flagged and reported.

**Metadata:** A data dictionary describes the data fields associated with the OHS dataset.

**Timeliness:** The OHS had been conducted annually, but data have not been collected since 2009, pending survey redesign and funding.

**Longevity/Future Plans:** The OHS was cited by DBP stakeholders as a potential source for collecting speed data in the future.

**Users:** Customers for OHS data include U.S. DOT modal administrators, members of Congress, and other transportation researchers.

**Accessibility:** All OHS data for 2000, 2001, 2002, 2003, 2005, and 2009 are accessible to the public and are available for download on the BTS website.[112] The website serves as a central facility to obtain information on the OHS, sample design and data collection procedures, OmniStats publications using OHS data, and the raw OHS data which is available for download and analysis by the user community.

**Contact:** Joy Sharp, RITA BTS, joy.sharp@dot.gov.

[112] Omnibus Household Survey website, http://www.bts.gov/programs/omnibus_surveys/household_survey/.

# Detailed Inventory – Other Data Programs

***Data.gov***

**Data Owner:** Data.gov is hosted by the U.S. General Services Administration. Departments/Agencies submitting datasets to Data.gov retain ownership of the data.

**Description:** The Data.gov is a data repository for Federal datasets and tools. It contains searchable catalogs that provide access to raw datasets and various tools from different Federal agencies. Data.gov increases the ability of the public to easily find, download, and use datasets that are generated and held by the Federal Government. It is the closest thing to having a requirement in place to make data available.

Data.gov recently transitioned to a cloud-based Open Data platform that allows citizens, developers and government agencies to interactively discover, explore, share and contribute to data. The Next Generation Data.gov platform provides a new data catalog that makes it easier to search datasets and tools by category, agency, keyword, or data format. It allows interactive exploration of data and the ability to sort, search, and analyze data easily, as well as visualize it through charts or maps. Users can disseminate data through social media, blogs, and forums. The Next Generation Data.gov also supports open data federation where an agency's datasets will be automatically uploaded on Data.gov.

**Purpose:** The purpose of Data.gov is to increase public access to high value, machine readable datasets generated and hosted by the Federal government.

**Data Source Hierarchy:** Data.gov is classified as a data repository.

**Geographic Scope/Coverage:** Geographic coverage varies for each dataset.

**Format:** Data.gov provides access to datasets, external datasets, files and documents, filtered views, charts, maps, calendars, and forms. Data formats could include data in XML, Text/CSV, KML/KMZ, Feeds, XLS, or ESRI Shapefile formats, while a catalog of tools link users to sites that include data mining and extraction tools and widgets. The following data catalogs are available:

- Raw Data Catalog: Data.gov features a catalog with instant view/download of platform-independent, machine-readable data (e.g., XML, CSV, KMZ/KML, or shape file formats), as well as a link to a metadata page specific to the respective dataset. The metadata page will have additional links to authoritative source information from the sponsoring agency's website including any pertinent agency technical documentation regarding the dataset.
- Tools Catalog: Data.gov features a tool catalog that provides the public with simple, application-driven access to Federal data with hyperlinks. This catalog features widgets, data mining and extraction tools, applications, and other services.
- Widgets: are interactive virtual tools that provide single-purpose services such as showing the user the latest news, the current weather, the time, calendar, dictionary, map program, calculator, desktop notes, photo viewers, or even a language translator, among other things.

- Data Mining and Extraction Tools: offer applications that allow users to either produce maps, tables, or charts of the subset of data that are specific to the user's interests or to build their own dataset extracted from a data source. Many of these sites also offer downloadable data.
- RSS Feeds: RSS (also known as Really Simple Syndication) is a format of data feed some agencies use to provide frequently updated content. When an agency feed is syndicated it means an individual can subscribe to it and get automatically updated content from the aggregator of the particular feed.
- Geodata Catalog: Data.gov features a geodata catalog that includes trusted, authoritative, and Federal geospatial. This catalog includes links to download the datasets and a metadata page with details on the datasets, as well as links to more detailed Federal Geographic Data Committee (FGDC) metadata information. The geospatial datasets available on Data.gov are provided in up to three open file formats: Keyhole Markup Language (KML), Compressed Keyhole Markup Language (KMZ) and ESRI Shapefile. These datasets are all viewable via many commercial and freely available applications.

**System Requirements:** The Raw Data Catalog provides an instant download of machine-readable, platform-independent datasets. The Tools Catalog provides hyperlinks which lead to agency tools or web pages that allow users to mine datasets. The Geodata Catalog requires use of a geographic information system or a free application such as ArcGIS Explorer or ArcReader to view datasets.

**Data Quality:** All information accessed through Data.gov is subject to the Information Quality Act (P.L. 106-554). For all data accessed through Data.gov, each agency has confirmed that the data being provided through this site meets the agency's Information Quality Guidelines.

As the authoritative source of the information, submitting Departments/Agencies retain version control of datasets accessed through Data.gov in compliance with record retention requirements outlined by the National Archives and Records Administration (NARA).

**Metadata:** Data.gov provides descriptions of the Federal datasets, metadata, information on how to access the datasets, and tools that leverage government datasets. Metadata standards for Data.gov are based on an adaptation of the Dublin Core metadata standard. A Glossary of Terms provides information on the various file formats. A metadata template is available through the Data.gov Dataset Management System (DMS) website.[113]

**Timeliness:** Timeliness varies with each dataset.

**Longevity/Future Plans:** Data.gov was created as part of the Open Government Initiative of the Obama Administration. Future plans/longevity of the program are unclear.

**Users:** In 2011, there were more than 1.5 million downloads of datasets and tools from Data.gov. The user base is broad and could include citizens, journalists, researchers, educators, transparency advocates, commercial entities, entrepreneurs, Federal, State, and local government agencies, and numerous other types of users.

[113] Data.gov Dataset Management System website, https://dms.data.gov/dms/webroot/flex/datagov.html.

**Accessibility:** Data.gov datasets are accessible to the public and are available for download through the Data.gov website.[114] The Dataset Management System (DMS) website provides an automated process for departments/agencies to submit datasets to the Data.gov catalog.

**Contact:** Contact information varies. For each dataset, a link is provided to contact the dataset owner. These messages are then sent to the relevant agency for review.

***Integrated Corridor Management (ICM) Program***

**Data Owner:** RITA ITS Joint Program Office

**Description:** The U.S. DOT is partnering with eight "Pioneer Sites" in a 5-year initiative to develop, deploy and evaluate integrated corridor management (ICM) concepts in eight of our nation's busiest corridors. The ICM Initiative aims to advance the state of the practice in transportation corridor operations to manage congestion. This initiative will provide the institutional guidance, operational capabilities, Intelligent Transportation Systems (ITS) technology and technical methods needed for effective ICM systems.

**Purpose:** The objective of the ICM initiative is to demonstrate how intelligent transportation systems (ITS) technologies can efficiently and proactively manage the movement of people and goods in major transportation corridors.

**Data Source Hierarchy:** ICM is classified as a raw data source of freeway and arterial speed and volume data, HOV lane data, ITS asset locations, signal timing plans and location data, transit data (lines, frequencies, stop or terminal locations, existing diversion plans), 511 system data, regional MPO travel demand model roadway networks and socioeconomic data, and other corridor data needed to support the ICM modeling effort.

**Geographic Scope/Coverage:** Geographic scope/coverage is regional. Currently, there are three Stage 2 analysis, modeling, and simulation (AMS) evaluation sites under the United States Department of Transportation (U.S. DOT) Integrated Corridor Management Program. These sites are Minneapolis, San Diego, and Dallas. Caltrans has twenty-six (26) Corridor System Management Plans (CSMP) under development or deployed. There also are related studies underway in Atlanta, New York, and other areas.

**Format:** Data format varies based on ICM corridor and simulation method used.

**System Requirements:** System requirements depend on the general type or levels of simulation analysis tools used for integrated corridor modeling and simulation – microscopic, mesoscopic, and macroscopic. Microscopic simulation models focus on the movement of every vehicle every second or a tenth of a second. Microscopic simulation models are useful for assessing traffic control strategies, such as ramp metering and arterial traffic signal coordination. Mesoscopic simulation models provide detailed analysis for sub-areas. Mesoscopic models are used to analyze traveler information, tolling, high-occupancy toll (HOT) lanes, congestion pricing, and regional diversion. Macroscopic travel demand models represent the traditional travel demand models that are used to assess regional travel patterns and mode shift.

[114] Data.gov website, http://www.data.gov

**Data Quality:** Data quality varies by corridor.

**Metadata:** Varies by corridor.

**Timeliness:** Varies by corridor.

**Longevity/Future Plans:** The project is a five-year initiative.

**Users:** Users include a variety of transportation stakeholders within each ICM corridor that would benefit from better transportation operations decision-making.

**Accessibility:** Currently, accessibility to corridor data is limited to ICM team members.

**Contact:** Brian Cronin, (202) 366-8841, brian.cronin@dot.gov; Steven Mortensen, (202) 493-0459, steven.mortensen@dot.gov; Robert Sheehan, (202) 366-6817, robert.sheehan@dot.gov; Dale Thompson, 202-493-3420, dale.thompson@dot.gov.

***I-95 Corridor Coalition Data***

**Data Owner:** I-95 Corridor Coalition

**Description:** The I-95 Corridor Coalition's Vehicle Probe project is a groundbreaking initiative and collaborative effort among the Coalition, University of Maryland and INRIX providing comprehensive and continuous real-time travel information to members.

**Purpose:** The purpose of this multi-state project is to acquire travel times and speeds on freeway and arterials using probe technology to provide regional management of traffic and traveler information useful to long-distance travelers.

**Data Source Hierarchy:** The I-95 Traffic Monitoring website is considered a repository for accessing the Inrix data stream.

**Geographic Scope/Coverage:** Coverage includes 1,917 miles of Interstate-95 along the eastern seaboard from Maine to Florida.

**Format:** Real-time traffic data is delivered in XML format or traffic tiles that allow overlay of traffic data onto a map for visualization, while historical data is delivered in CSV format.

**System Requirements:** Inrix delivers files to customers in XML format via a Web Services Application Programming Interface (API). Archived speed data is not available via API, but rather by State and road type within a CSV file that is returned after a request is made on the I-95 Traffic Monitoring website.

**Data Quality:** Data quality has been independently verified in large-scale testing. Documented quality levels are above 95% accuracy and above 99.9% availability.

**Metadata:** Metadata is provided in the form of data quality indicators such as a score to indicate whether speed is based on real-time or historical data, and a confidence value (C-value) to provide supplemental information on the type and confidence of real-time data.

**Timeliness:** The project started in October 2006. July 1, 2011 marked the start of the fourth year (Coalition Project Year 19) that Vehicle Probe Project (VPP) data is being provided to the Coalition's members. Traffic data includes speed, travel time, average speed, and the 85th percentile reference speed for each road segment, provided 24 hours per day, 7 days per week, and a maximum data latency of 5 minutes.

**Longevity/Future Plans:** The project has been ongoing since October 2006. Coverage expansions continue to occur throughout the years.

**Users:** I-95 Coalition members include key decision makers that have or will influence the operation of the Corridor including State and Local Departments of Transportation, Transportation Authorities, Transit and Rail Agencies, Port Authorities, Motor Vehicle Agencies, State Police/Law Enforcement, U.S. Department of Transportation, Canadian Province Department of Transportation, Intercity Passenger and Freight Transportation Providers, and Transportation Industry Associates.

**Accessibility:** Access is limited to I-95 Coalition member agencies and their contractors. Access to data is granted once a member agency has signed a Data Use Agreement through the University of Maryland. The I-95 Traffic Monitoring website[115] provides a visualization of traffic conditions on a map and access to archived data.

**Contact:** Bill Stoeckert, I-95 Corridor Coalition, 774-207-0367, wstoeckert@yahoo.com; Phil Tarnoff, University of Maryland, 301-403-4619, tarnoff@eng.umd.edu; Stan Young, University of Maryland, 301-403-4593, seyoung@umd.edu

***VII Data Use Analysis and Processing (DUAP) Project***

**Data Owner:** Michigan Department of Transportation (MDOT)

**Description:** The Vehicle-Infrastructure Integration (VII) Data Use Analysis and Processing (DUAP) program was initiated in 2006 by MDOT to complement research initiatives by the Federal government and automobile manufacturing industry. The focus is on supporting the collection, aggregation, storage, and use of data in applications of interest to MDOT. The program generally covers processes aimed at collecting information from vehicles traveling within a road network. The expectation is that a DUAP system will draw data from envisioned connected vehicle systems, existing MDOT data sources, relevant external sources, and various MDOT projects that may be executed. The resulting probe vehicle data collection system would be used not only to enrich existing MDOT applications, but also to facilitate the development of new applications within MDOT, as well as applications outside MDOT relying on data managed by the agency. Proof of concept (POC) testing of the probe vehicle data collection system was conducted in 2008.

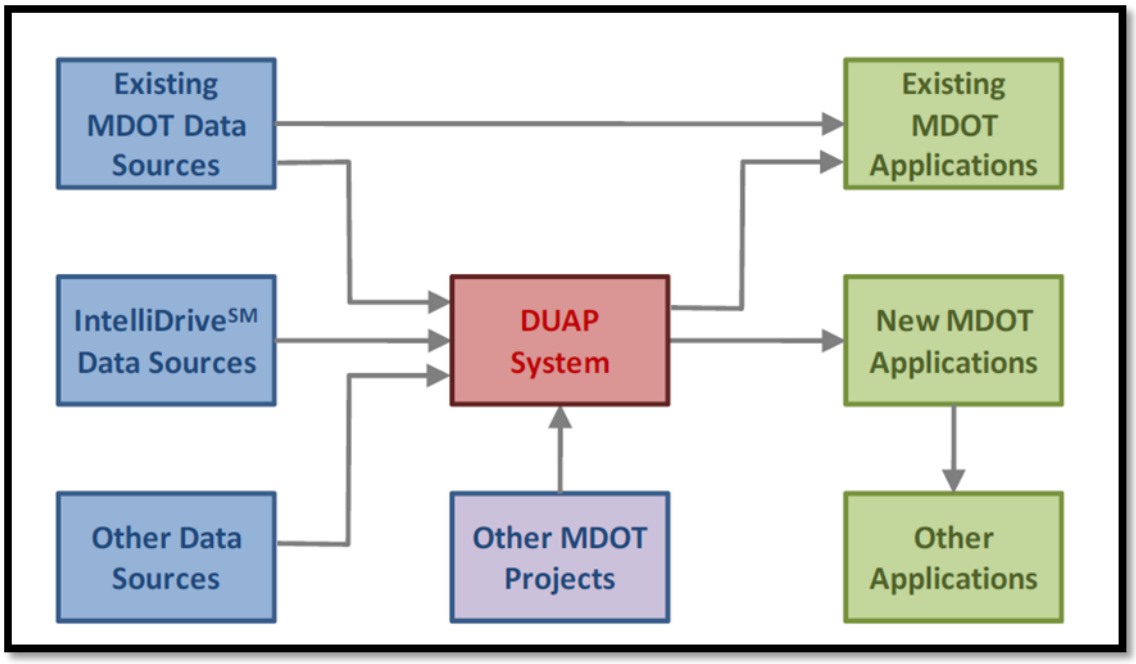
**Purpose:** The primary purpose of MDOT's Data Use Analysis and Processing (DUAP) project is to study how new connected vehicle data sources may be combined with other data sources, and how use of these new data sources may impact MDOT's current transportation system operations through improved methods and processes. Given that the vision of connected vehicles is to get real-time probe data from most, if not all, of the vehicles on the roadway, these projects seek to answer the following questions:

[115] I-95 Traffic Monitoring website, http://i95.inrix.com.

- How can probe vehicle data be used to transform MDOT's business practices and better achieve MDOT's organizational goals?
- How can this data be used to provide significant improvements in a road network, situational awareness, congestion management, asset management and safety?
- In what other ways can MDOT benefit from probe data?

**Data Source Hierarchy:** The probe vehicle data collection system is a source of raw data that is derived from existing MDOT data sources, connected vehicle data sources, and other data sources, as shown in Figure H.12.

**Figure H.12 DUAP Evaluation Framework** (Source: University of Michigan Transportation Research institute, VII Data Use Analysis and Processing (DUAP), Final Project Report (Phase II), October 2010)




**Geographic Scope/Coverage:** The DUAP program is a regional project, with probe data collection conducted in and around Novi, Michigan.

**Format:** Probe data collection within the POC test followed recommended data format standards at the time the test was developed. The primary standard was the SAE J2735 Dedicated Short Range Communications (DSRC) Message Set Dictionary, which specifies standard message sets, data frames and data elements for use by applications intended to utilize the 5.9 GHz DSRC communication systems.

Data within the probe vehicle data collection system consists of three basic types of snapshots:

- Periodic snapshots, meant to record the status of various vehicle systems at specific intervals
- Stop/start event snapshots, meant to record when a vehicle stops and starts moving after having made a stop
- Special event snapshots, meant to record when specific changes in specific vehicle status occur, such as when brakes are applied, wipers or headlights turned on, etc.

Detailed samples of each of these data formats are provided in the source document.[116]

**System Requirements:** The VII POC **input service** uses the X-031 Probe Data Service interface as described by the VII Network User to Service Delivery Node (SDN) Subsystem document. The **dynamic data services** are a caching service intended to store DUAP data for fast access over a relatively short duration. **Persistent data services** provide longer-term storage within DUAP of both data collected by input services and metadata needed by DUAP services to sustain their operations. The purpose of the DUAP system's **computational services** is to apply logical algorithms to incoming vehicle and traffic observations in order to transform those observations into data that is directly applicable to transportation management and operations processes. **Output services** subscribe to analyzed data and format it for use by other services both within the DUAP system and external to it. This could be as simple as structuring data for presentation as a report on a DUAP Web page. It could also build informative messages to be sent to VII-participating vehicles. This interaction could be achieved using the X-032 interface to the Advisory Message Distribution Service as described in the VII National System Specification. **Presentation Services** support human interpretation of DUAP data. **Administrative services** exist within the DUAP system to configure other services.

**Data Quality:** The DUAP system evaluation framework considered the following issues that impact data quality:

- Data sampling rate
- Ability to collect data from every link in a network
- Potential for data bias due to snapshot generation protocols
- Potential data losses due to full memory buffer
- Impacts of privacy rules on quantity of data collected
- Ability to effectively track vehicles over short distances and across intersections
- Data latency

**Metadata:** Metadata characterizes the context in which the data was collected. Metadata for probe vehicle data include information about the weather and traffic conditions on the day or period the data were collected, data precision, instrumentation used to generate the data, or the various processes that were applied in the processing of data. In addition, a data dictionary summarizes the data fields captured in the probe vehicle dataset. Table H-2 provides the POC probe message header data fields.

[116] Dion, F. and R. Robinson. VII Data Use Analysis and Processing (DUAP) Final Project Report, Phase II. University of Michigan Transportation Research Institute Report No. UMTRI-2010-28, October 2010. http://deepblue.lib.umich.edu/bitstream/2027.42/ 78569/1/102726.pdf

**Table H-2. POC Probe Message Header Data Fields** (Source: University of Michigan Transportation Research institute, VII Data Use Analysis and Processing (DUAP), Final Project Report (Phase II), October 2010)

| Parameter | | Description | Units | Precision | Value recorded in snapshot |
|---|---|---|---|---|---|
| Message ID | | Unique probe message ID | n/a | n/a | Message ID |
| SDN | Year | Year data was recorded | Calendar year | n/a | Year in yyyy format |
| | Month | Month data was recorded | Month name | n/a | Month of year (1 – 12) |
| | Day | Day data was recorded | Days of week | n/a | Day of month (0 – 31) |
| | Hour | Hour data was recorded | Hour of day | 1 hour | Hour of day (0 – 24) |
| | Min | Minute data was recorded | Minutes within hour | 1 min | Minutes from start of hour (0 – 59) |
| | Sec | Second data was recorded | Seconds within minute | 0.001 s | Milliseconds from start of minute (0 – 59999) |
| | Msg | Message count | Messages | n/a | Number of messages sent from a specific vehicle during a given transmission |
| RSE | Year | Year data was recorded | Calendar year | n/a | Year in yyyy format |
| | Month | Month data was recorded | Month name | n/a | Month of year (1 – 12) |
| | Day | Day data was recorded | Days of week | n/a | Day of month (0 – 31) |
| | Hour | Hour data was recorded | Hour of day | 1 hour | Hour of day (0 – 24) |
| | Min | Minute data was recorded | Minutes within hour | 1 min | Minutes from start of hour (0 – 59) |
| | Sec | Second data was recorded | Seconds within minute | 0.001 s | Milliseconds from start of minute (0 – 59999) |
| | Long | Geographical longitude | Degrees | 0.000125 | -7,200,000,000 to 7,200,000,000, representing a range of $-90^{o}$ to $+90^{o}$ |
| | Lat | Geographical latitude | Degrees | 0.000125 | -14,400,000,000 to 14,400,000,000, representing a range of $-180^{o}$ to $+180^{o}$ |
| | Elev | Elevation of vehicle | Meters | 0.1 m | Measured elevation minus 1000 m |
| | Heading | Heading of vehicle | Degrees | 0.00459 | Heading measurement ($0\text{-}360^{0}$) multiplied by 182.0417 |
| | Speed | Speed of vehicle | Meters per second | 0.01 m/s | Speed measurement multiplied by 100. Should be 0 for any fixed RSE |
| Snapshot Count | | Number of snapshots in message | Snapshots | n/a | Number of snapshots, between 1 and 4 |

**Timeliness:** Data were collected between May 6, 2008 and October 5, 2008. Periodic snapshot data was generated every 20 seconds for vehicles traveling at speeds greater than 60 mph, every 4 seconds for speeds lower than 20 mph, and at intervals linearly interpolated between 4 and 20 seconds for speeds between 20 and 60 mph. Stop events were recorded after a vehicle has been immobilized for five seconds, when at least 15 seconds had elapsed since the last recorded stop. A start event was recorded when the speed of a vehicle under consideration attains a speed of 10 mph. Special event snapshots were generated when specific changes in vehicle status were observed, such as when brakes were applied, headlights turned on or off, etc.

**Longevity/Future Plans:** The final project report recommended the development of test beds and that MDOT initiates the development and deployment of a real connected vehicle system.

**Users:** Users of DUAP program data include transportation agencies, automobile manufacturers, researchers, and others interested in probe vehicle data.

**Accessibility:** Access to data is limited to MDOT and its associated stakeholders.

**Contact:** Michigan DOT 517-373-2090

## Detailed Inventory – Real-Time Data Capture and Management Program

***TFHRC Data Resources Testbed***

**Data Owner:** Turner Fairbanks Highway Research Center (TFHRC)

**Description:** TFHRC recently established a Data Resources Testbed (DRT), which is part of the Transportation Operations Laboratory (TOL). The overall lab is comprised of: 1) a Data Resources Testbed (DRT); 2) a Concepts and Analysis Testbed (CAT); and 3) a Cooperative Vehicle-Highway Testbed (CVHT). The three testbeds within the TOL will help FHWA fulfill their operations research missions.[117]

It is envisioned that the DRT will host real-time and archived transportation data from a variety of sources to support performance measurement and transportation system management applications. Data sources will include traditional roadway sensors, data from vehicle probes including detailed data available from vehicles equipped with connected vehicles technologies, data from personal mobile devices, and transit and freight operators. The DRT will capture internal test data and would most likely be linked to external data from PeMS, RITIS, the Research Data Exchange, and other sources as they become available. The Testbed will assemble and archive data, host traffic data sets, analyze operations and performance, provide advanced visualization tools to improve situational awareness, and aid strategic program and tactical operations decision-making.

The Testbed will also provide the data environment and resources needed by internal and external researchers to develop and test new cooperative services under the intelligent transportation systems (ITS) research initiatives, including the ability to conduct V2V and V2I testing onsite.

[117] Turner-Fairbank Highway Research Center website, http://www.fhwa.dot.gov/research/tfhrc/labs/operations/.

**Purpose:** The laboratory will enable FHWA to validate and refine new transportation services and technologies before committing to larger scale research, development, testing, and deployment phases. Eventually, the laboratory will also provide a unique venue for outreach of the innovative service concepts and technologies, in that technology transfer can be supported through first-hand knowledge of the technologies and physical examination of prototype systems and simulations of concepts of operations.

**Data Source Hierarchy:** The multi-source data includes data from traditional roadway sensors, data from vehicle probes including the detailed data available from vehicles equipped with connected vehicles types of technologies, data from personal mobile devices, and data from transit and freight operators. Once developed, the testbed could be classified as a source of raw data.

**Geographic Scope/Coverage:** TBD

**Format:** The DRT has been established. However, guidelines and procedures are not currently in place. Potential research concepts include data integration, visualization, and other interesting concepts involving data. TFHRC is interested in partnering with other offices and vendors on potential research ideas and funding sources.

**System Requirements:** TBD

**Data Quality:** TBD

**Metadata:** TBD

**Timeliness:** TBD

**Longevity/Future Plans:** The Testbed was built under and is currently operating under the Ops IDIQ. It will be re-competed as a task order work contract so it can be a resource for other offices such as operations, planning for operations, etc.

**Users:** Internal and external researchers

**Accessibility:** Open to both internal and external researchers

**Contact:** Carol Tran, 202-493-3315, carol.tan@dot.gov

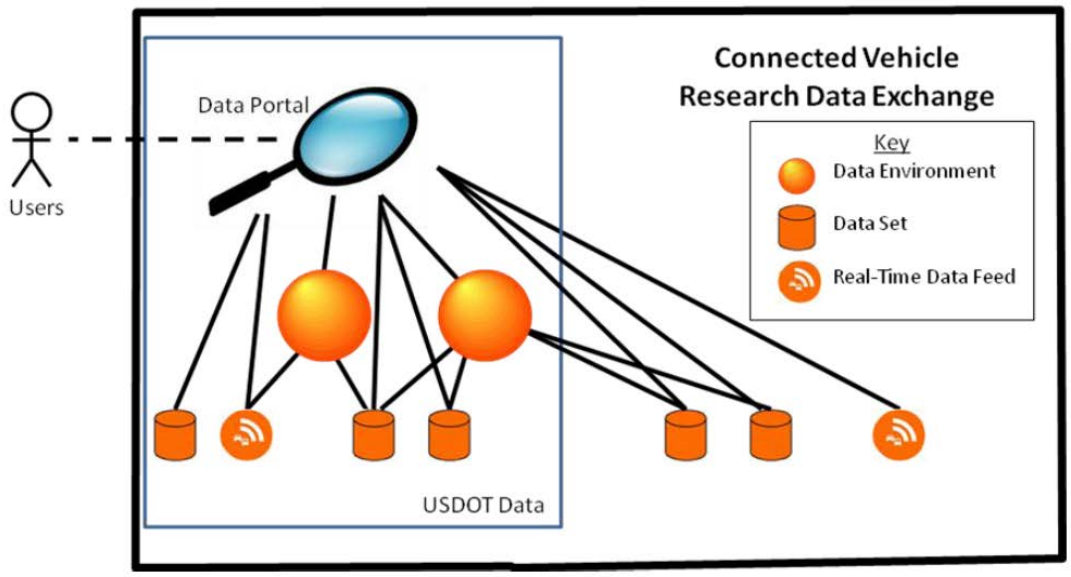
***DCM Program Research Data Exchange***

**Data Owner:** The RITA ITS Joint Program Office owns the Research Data Exchange, and it is administered by Noblis. Ownership of data sets and data environments will remain with the source.

**Description:** The Research Data Exchange (RDE) is a new decentralized, Internet-based, federated system of systems. The RDE serves as a research repository for the Real-Time Data Capture and Management Program, and it contains data environments (collections of related data sets and possibly real-time data feeds), individual data sets, and real-time data feeds. A common web-based Data Portal will provide access to all the contents of the Research Data Exchange (Data Environments, metadata, data sets, and real-time data feeds) regardless of which constituent Data Management System the information resides in. It will also provide mechanisms for users to link to

and obtain data sets or data feeds from the associated Data Management Systems.[118] Figure H.13 provides an overview of the Research Data Exchange.

**Figure H.13 Overview of the Research Data Exchange** (Source: McGurrin, M., A. Corrigan, J. Larkin, and K. Wunderlich. Concept of Operations: Data Capture and Management Research Data Exchange (Draft Report). RITA ITS Joint Program Office, August 1, 2011)




**Purpose:** The purpose of the RDE is to serve the needs of mobility application developers and researchers who want access to connected vehicle data. DOT wants to make sure data is available through the RDE for all researchers to use.

**Data Source Hierarchy:** The Research Data Exchange serves as a repository for data environments, individual data sets, and real-time data feeds.

**Geographic Scope/Coverage:** The geographic scope will vary for each data environment. The prototype data environment currently includes proof of concept datasets from the V2V and V2I Technology Test Bed, as well as a live feed from probe data.

**Format:** The RDE will store information about data content, projects, status, and summary information on access policies for all of the Data Management Systems associated with the Research Data Exchange. This information comes in two forms: one is the metadata detailing the data sets, Data Environments, and real-time data feeds, including data type, location, frequency, time period, and access rules; and the other is in the form of news information describing changes in the Research Data Exchange, including current status of existing data sets and Data Environments, new data sets

[118] McGurrin, M., A. Corrigan, J. Larkin, and K. Wunderlich. Concept of Operations: Data Capture and Management Research Data Exchange (Draft Report). Prepared for the RITA ITS Joint Program Office, August 1, 2011. http://www.its.dot.gov/meetings/pdf/ResearchDataExchangeConceptofOperations_August2011DRAFT.pdf

and Data Environments, and corrections to the data. Data format will adhere to open data and open source software standards.[119]

The Data Portal will provide a means to register new users and allow registered users to post information about how they are using the data. It will also allow users the ability to search and review metadata from all data sets, Data Environments, and real-time data feeds in the Research Data Exchange. These capabilities will allow members and the administrator of the system to know who else is using the data and for what purpose.

**System Requirements:** The Research Data Exchange is accessible through a common web-based Data Portal, which serves as a central location for general information about the Data Capture and Management Program. Users register with the Data Portal by completing a user profile located on the website.

**Data Quality:** Data providers are expected to perform a data quality analysis before disclosing their data set to the public. Completeness, accuracy, and logical consistency are adopted as the key metadata quality measures.

**Metadata:** All data posted on the RDE must have accompanying metadata per ASTM E 2468-05, Standard Practice for Metadata to Support Archived Data Management Systems.[120] Metadata must be updated whenever there are any changes to data definitions in the associated data set. The ITS JPO anticipates that the open data licensing and metadata requirements will be included in the Appendix of all procurement documents going forward. The RDE may also accept volunteered data, although requirements would still be enforced.

**Timeliness:** Timeliness will vary. Data will be captured and stored at once or at regular intervals from one or more sources.

**Longevity/Future Plans:** Phase I involved developing the prototype data environment, which demonstrated the need for more than a static archive. The prototype data environment will go offline and be replaced by the RDE by May 2012. The first step is to migrate current capabilities – i.e., all functionality and content from the prototype data environment will be made available in the RDE. The test datasets and data from other projects will also be added to the RDE. For example, the Weather Program is testing probe data from snowplows, and these archived datasets will be added. They want to link to existing data sources (e.g., SunGuide, PeMS, etc.) rather than re-create them. It is possible for the RDE to accept data from other Federal programs, as long as they follow the same data requirements.

The RDE is currently funded through 2015. As part of the sunset plan, it is assumed that data will migrate out of the RDE into other collegiate places, or be taken over by others and run.

[119] Brecher, A., J. Hassol, and S. Sloan. Policy Analysis and Recommendations for the Data Capture and Management Program: Implementation of Open Data Policies and System Policies for the Research Data Exchange and Data Environments (Draft Report). Prepared for the RITA ITS Joint Program Office, February 2012.

[120] Jung, S., R. Glassco, K. Wunderlich. Metadata Guidelines for the Research Data Exchange, Final Report. Prepared for the RITA ITS Joint Program Office, December 30, 2011.

**Users:** Users include system developers who will create and support the Research Data Exchange based on the user needs and system concepts; connected vehicle mobility stakeholders who will determine whether their needs and desires have been adequately captured for use by the Research Data Exchange developer(s); and analysts, researchers, and mobility application developers requiring access to a research data for application development.

**Accessibility:** The Research Data Exchange is accessible through a common web-based Data Portal.[121] Users must register with the system to obtain access to data and participate in the community resources provided by the Research Data Exchange.

**Contact:** Dale Thompson, dale.thompson@dot.gov; Gene McHale, gene.mchale@dot.gov

***DCM Program Test Data Sets***

**Data Owner:** RITA ITS Joint Program Office

**Description:** The Real-Time Data Capture and Management Program issued a Request for Procurement for well-documented, quality test data sets available from recent or ongoing operations, field tests, or simulations of emerging technologies support mobility, environment, transit, freight, weather, and other surface transportation research.

**Purpose:** Test data sets will provide tangible and broadly accessible examples of the core data environment concept. Data sets will also be available to support applications research and development. These prototypes and test data sets will further serve as precursors to resolve cross-cutting issues such as hosting, aggregation, access, and intellectual property rights. The test data sets will elicit feedback on data needs from early Dynamic Mobility Applications development.

**Data Source Hierarchy:** The test data sets are classified as a raw source of data.

**Geographic Scope/Coverage:** Geographic scope/coverage varies for each test data sets. The following test data sets are available:

- Pasadena Test Data Set. Mygistics (formerly PTV) modeled the City of Pasadena. The dataset will include 2 months of simulated, real-time and forecast data. This includes AirSage data, live detector information, and 15- and 30-minute projections of traffic conditions. This dataset will assess the accuracy of 30-minutes forecasts. The dataset will be delivered by the end of January 2012.
- Seattle Test Data Set. The University of Washington is submitting a Seattle I-5 data set, including freeway, arterial, freight, and transit data. Data sources include loop detectors, arterial data from automated license plate, Bluetooth MAC readers, traffic signal timing plans, and freight GPS data. This dataset will include 6 months of data and will be delivered by the end of December 2011.
- Portland Test Data Set. Portland State University is submitting a multimodal dataset for Portland, Oregon. It will include multimodal, freeway, arterial (second-by-second data from signals), transit, and incident data, and the dataset will be delivered by the end of December 2011. The test data sets will include transit data from Tri-Met.

[121] U.S. DOT Research Data Exchange (RDE) Website, https://www.its-rde.net/

- San Diego Test Data Set. Berkeley Transportation Systems is submitting a San Diego dataset. It will include incident, loop, GPS data, and weather data for a section of freeway in San Diego. This dataset will be delivered by the end of December 2011.

**Format:** Test data sets comprise of observed data or a combination of observed, derived, and/or simulated data.

**System Requirements:** The data sets meet open data principles, e.g., the Open Data Commons Attribution License.[122]

**Data Quality:** Data quality varies for each test data set. In general, data cleanup procedures were conducted to identify and correct errors, flag suspect data values, and identify periods of missing data.

**Metadata:** All test data sets now available on the RDE include ASTM-compliant metadata that describe data quality, relationship to other data, and information about where and when each value was captured.

**Timeliness:** Timeliness varies for each test data set.

**Longevity/Future Plans:** The test data sets will be shared with the public in the future as part of one or more Real-Time Data Capture and Management data environments.

**Users:** Users include researchers, application developers, and system operators.

**Accessibility:** Test data sets and resulting documentation will be made available through the Research Data Exchange.[123]

**Contact:** Dale Thompson, (202) 493-0259, dale.thompson@dot.gov

***Dynamic Mobility Application Datasets***

**Data Owner:** RITA ITS Joint Program Office

**Description:** Dynamic Mobility Application Datasets will consist of integrated data sets from relevant data sources such as wirelessly connected vehicles, travelers, and roadside/wayside infrastructure. The resulting well-organized data and associated descriptors (data environments) will include data from field tests and advanced simulation models and be made broadly available.

**Purpose:** To support the development of high priority arterial, freeway, regional (information), and corridor (control) applications.

**Data Source Hierarchy:** Dynamic Mobility Application datasets will be a source of raw data.

**Geographic Scope/Coverage:** TBD

**Format:** TBD

[122] Open Data Commons Attribution License website, http://www.opendatacommons.org/licenses/by/
[123] U.S. DOT Research Data Exchange (RDE) Website, https://www.its-rde.net/

**System Requirements:** TBD

**Data Quality:** TBD

**Metadata:** TBD

**Timeliness:** TBD

**Longevity/Future Plans:** TBD

**Users:** Users will include researchers, application developers, and system operators.

**Accessibility:** Dynamic Mobility Application data sets and resulting documentation will be made available through the Research Data Exchange.[124]

**Contact:** Katherine K. Hartman, (202) 366-2742, kate.hartman@dot.gov; Robert Rupert, (202) 366-2194, robert.rupert@dot.gov; Steven Mortensen, (202) 493-0459, steven.mortensen@dot.gov; Randy Butler, (202) 366-9215, randy.butler@dot.gov.

# H.3 Recommendations and Next Steps

Next steps for the Data Business Plan project include the following:

- Identify the relationship of data sets to the Data Business Plan goals and objectives. This will be accomplished through the development of a diagram showing the relationship between core national data sets, the Research Data Exchange, the Real-Time Data Capture and Management Program, and FHWA data programs; showing how current data sets and data programs currently (and in the future will) support the RDE and the Real-Time Data Capture and Management Program; and identifying gaps and overlaps.
- Develop White Paper (which includes the Data Governance Evaluation Framework for the Data Business Plan). The white paper will be accomplished in the context of the Data Governance Plan documented by the Office of Highway Policy Information. It will include recommendations for enhancement to ensure mobility data and connected vehicle data needs (Goals 1 and 2) are incorporated; creation of a Mobility Data Forum; and creation of appropriate governance structures for the focus areas of the forum.
- To maintain usefulness of the Data Inventory Report as a resource, it is recommended that FHWA consider committing resources to periodically review and update its contents.

[124] U.S. DOT Research Data Exchange (RDE) Website, https://www.its-rde.net/.

U.S. Department of Transportation
ITS Joint Program Office-HOIT
1200 New Jersey Avenue, SE
Washington, DC 20590

Toll-Free "Help Line" 866-367-7487
www.its.dot.gov

FHWA-JPO-13-084

U.S. Department of Transportation