This PDF is available at http://nap.edu/23463

Data to Support Transportation Agency Business Needs: A Self-Assessment Guide

DETAILS

BUY THIS BOOK

FIND RELATED TITLES

### AUTHORS

Spy Pond Partners, LLC and Iteris, Inc.

Visit the National Academies Press at **NAP.edu** and login or register to get:

– Access to free PDF downloads of thousands of scientific reports
– 10% off the price of print titles
– Email or social media notifications of new titles related to your interests
– Special offers and discounts

NCHRP REPORT 814

Data to Support Transportation Agency Business Needs: A Self-Assessment Guide

Spy Pond Partners, LLC
Arlington, MA

IN ASSOCIATION WITH

Iteris, Inc.
Santa Ana, CA

TRANSPORTATION RESEARCH BOARD

NATIONAL COOPERATIVE HIGHWAY RESEARCH PROGRAM

Systematic, well-designed research is the most effective way to solve many problems facing highway administrators and engineers. Often, highway problems are of local interest and can best be studied by highway departments individually or in cooperation with their state universities and others. However, the accelerating growth of highway transportation results in increasingly complex problems of wide interest to highway authorities. These problems are best studied through a coordinated program of cooperative research.

Recognizing this need, the leadership of the American Association of State Highway and Transportation Officials (AASHTO) in 1962 initiated an objective national highway research program using modern scientific techniques—the National Cooperative Highway Research Program (NCHRP). NCHRP is supported on a continuing basis by funds from participating member states of AASHTO and receives the full cooperation and support of the Federal Highway Administration, United States Department of Transportation.

The Transportation Research Board (TRB) of the National Academies of Sciences, Engineering, and Medicine was requested by AASHTO to administer the research program because of TRB's recognized objectivity and understanding of modern research practices. TRB is uniquely suited for this purpose for many reasons: TRB maintains an extensive committee structure from which authorities on any highway transportation subject may be drawn; TRB possesses avenues of communications and cooperation with federal, state, and local governmental agencies, universities, and industry; TRB's relationship to the Academies is an insurance of objectivity; and TRB maintains a full-time staff of specialists in highway transportation matters to bring the findings of research directly to those in a position to use them.

The program is developed on the basis of research needs identified by chief administrators and other staff of the highway and transportation departments and by committees of AASHTO. Topics of the highest merit are selected by the AASHTO Standing Committee on Research (SCOR), and each year SCOR's recommendations are proposed to the AASHTO Board of Directors and the Academies. Research projects to address these topics are defined by NCHRP, and qualified research agencies are selected from submitted proposals. Administration and surveillance of research contracts are the responsibilities of the Academies and TRB.

The needs for highway research are many, and NCHRP can make significant contributions to solving highway transportation problems of mutual concern to many responsible groups. The program, however, is intended to complement, rather than to substitute for or duplicate, other highway research programs.

NOTICE

The report was reviewed by the technical panel and accepted for publication according to procedures established and overseen by the Transportation Research Board and approved by the National Academies of Sciences, Engineering, and Medicine.

The opinions and conclusions expressed or implied in this report are those of the researchers who performed the research and are not necessarily those of the Transportation Research Board; the National Academies of Sciences, Engineering, and Medicine; or the program sponsors.

The Transportation Research Board; the National Academies of Sciences, Engineering, and Medicine; and the sponsors of the National Cooperative Highway Research Program do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of the report.

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

The **National Academy of Sciences** was established in 1863 by an Act of Congress, signed by President Lincoln, as a private, non-governmental institution to advise the nation on issues related to science and technology. Members are elected by their peers for outstanding contributions to research. Dr. Ralph J. Cicerone is president.

The **National Academy of Engineering** was established in 1964 under the charter of the National Academy of Sciences to bring the practices of engineering to advising the nation. Members are elected by their peers for extraordinary contributions to engineering. Dr. C. D. Mote, Jr., is president.

The **National Academy of Medicine** (formerly the Institute of Medicine) was established in 1970 under the charter of the National Academy of Sciences to advise the nation on medical and health issues. Members are elected by their peers for distinguished contributions to medicine and health. Dr. Victor J. Dzau is president.

The three Academies work together as the **National Academies of Sciences, Engineering, and Medicine** to provide independent, objective analysis and advice to the nation and conduct other activities to solve complex problems and inform public policy decisions. The Academies also encourage education and research, recognize outstanding contributions to knowledge, and increase public understanding in matters of science, engineering, and medicine.

Learn more about the National Academies of Sciences, Engineering, and Medicine at **www.national-academies.org**.

The **Transportation Research Board** is one of seven major programs of the National Academies of Sciences, Engineering, and Medicine. The mission of the Transportation Research Board is to increase the benefits that transportation contributes to society by providing leadership in transportation innovation and progress through research and information exchange, conducted within a setting that is objective, interdisciplinary, and multimodal. The Board's varied activities annually engage about 7,000 engineers, scientists, and other transportation researchers and practitioners from the public and private sectors and academia, all of whom contribute their expertise in the public interest. The program is supported by state transportation departments, federal agencies including the component administrations of the U.S. Department of Transportation, and other organizations and individuals interested in the development of transportation.

Learn more about the Transportation Research Board at **www.TRB.org**.

C O O P E R A T I V E   R E S E A R C H   P R O G R A M S

FOREWORD

By Andrew C. Lemer
Staff Officer
Transportation Research Board

NCHRP Report 814: Data to Support Transportation Agency Business Needs: A Self-Assessment Guide presents guidance to assist decisionmakers and data practitioners at state departments of transportation (DOTs) in evaluating and improving the value of their data for decision making and their data-management practices. Agency practitioners rely on a wide range of data to support decision making about policy choices, infrastructure investments, and other agency functions. The self-assessment process described here may also be useful to metropolitan planning organizations (MPOs).

Transportation agencies are increasingly reliant on rapidly growing data resources to provide the bases for making critical decisions about how best to allocate available resources to build, operate, and maintain safe, efficient, and sustainable transportation systems to serve their communities. Technologies for collecting, storing, and manipulating data have advanced dramatically, potentially enhancing the timeliness and sophistication of management information while increasing demands on responsible agency staff to ensure that their data are current, accurate, reliable, and available. Data collection and maintenance can be costly, and agencies have sought ways to evaluate their data programs and improve the programs' effectiveness.

The objectives of NCHRP Project 08-92, "Implementing a Transportation Agency Data Self-Assessment" were to (1) test the feasibility of the data self-assessment process proposed in previous NCHRP research and (2) produce a guidebook for transportation agencies undertaking to implement this process. The guidebook was envisioned to be useful to decisionmakers and data practitioners at state departments of transportation (DOTs) for evaluating and improving the value of their data for decision making and their data-management practices. However, the process and guidance may also be useful to metropolitan planning organizations (MPOs), who face many of the same sorts of data challenges as DOTs.

The research was conducted by a team led by Spy Pond Partners, LLC, of Arlington, MA. The research team completed their work in three phases.

Phase 1 comprised a review and refinement of the data self-assessment process originally formulated in NCHRP Project 08-36/Task 100, "Transportation Data Self-Assessment Guide," completed in 2011. Refinements reflected practitioner comments on the earlier work, subsequent work on related topics sponsored by the FHWA and others, and interviews with currently serving senior agency managers, data practitioners, and data customers at DOTs and large MPOs. In addition, Phase 1 included design of case studies (to be performed in subsequent phases) to effectively test the guidance being developed and provide examples of how agencies might efficiently implement the self-assessment process.

In Phase 2, the research team conducted the case studies with two DOTs. The research team documented lessons learned about the feasibility of the data self-assessment methods,

considering such factors as DOT staff time used, meaningfulness of results, and effectiveness of the implementation process.

In Phase 3, the research team used the preceding work to prepare a guidebook to assist agencies with implementation of data self-assessment. The guidebook is aimed at helping agency decisionmakers and data practitioners to evaluate and improve their data to better support effective decision making. This document, *NCHRP Report 814: Data to Support Transportation Agency Business Needs: A Self-Assessment Guide*, is the guide. In addition, the team prepared a final report describing their work and presenting information from their research that may be useful to practitioners and other researchers. The report, published as *NCHRP Web-Only Document 214*, is available for download from the TRB website by searching for NCHRP Web-Only Document 214. Available through the NCHRP Project 08-92 web page are supplemental materials that may be useful to agencies undertaking data self-assessment: (1) three spreadsheet tools (Excel files) implementing the self-assessment process, (2) a slide deck (PowerPoint file) executive presentation on the objectives and steps in data self-assessment, and (3) a flyer (PDF file) succinctly presenting the assessment process to staff.

CONTENTS

SUMMARY

Data to Support Transportation Agency Business Needs: A Self-Assessment Guide

Data Is an Essential Transportation Agency Asset

Transportation agencies are increasingly data-intensive organizations. Data is used internally to drive critical agency decisions about how best to invest available resources to build, operate, and maintain safe, efficient, and sustainable transportation systems. Data about facilities, services, system performance, and expenditures are shared to provide necessary public accountability and meet agency reporting obligations. Real-time data sources have emerged in recent years that provide valuable situational awareness to agency maintenance and operations staff and enable agencies to provide traveler information services that result in more convenient and efficient travel choices.

Data Is Expensive to Collect and Maintain

Collecting, processing, analyzing, reporting, and sharing the varied types of data in a transportation agency require considerable resources—for computer hardware and software, staff time, and vendor support. Data sources, data management best practices, and information technologies are constantly evolving. Agencies can improve decision making by collecting or acquiring new data and by improving data management, analysis, and visualization capabilities. However, many of these improvements require technology investments and specialized staff skills which are in high demand.

A Systematic and Objective Approach for Assessing Transportation Agency Data Is Needed

Transportation agencies need to make prudent investments in data to meet an increasingly complex set of internal and external needs. Even where raw data are available, many agencies find themselves to be "data rich, but information poor." This Guide shows how agencies can systematically assess how well the data they have is working for them. Based on this assessment, agencies can identify steps to better align data with agency business needs and make more effective use of available resources for data collection and management. The Guide recognizes agency resource limitations and the need to look objectively at the value added by existing and potential data resources. The intent of the assessment process is not to develop a long wish list of additional data to collect, but to ensure that investments in data are paying off.

Assessment Approach Is Flexible and Scalable

The approach described in this Guide is intentionally broad and flexible. It can be used to examine the value being added by data from a business user perspective—and the shortcomings in data availability, quality, and usability that are limiting this value. The Guide's approach can also be used to assess the maturity of data management processes used at different levels and functional areas of the agency. Combining perspectives of data users and data managers can provide powerful insights into an agency's critical data improvement needs. These insights can be used to identify immediate specific actions to better meet agency needs for data, as well as more systemic improvements to strengthen the agency's foundational process for data management. The assessment can be scaled to focus on a limited set of data categories or business areas.

Staff Time Is Required but Worth the Investment

Agencies can conduct the self-assessment process using the accompanying tools and following the step-by-step instructions in this Guide. To ensure a productive assessment process, the Guide recommends that an agency (1) designate a champion to drive the activities and (2) enlist the services of suitable in-house facilitators to conduct assessment workshops. Each assessment workshop involves participation from a group of 5 to 10 staff members representing a range of perspectives. The Guide also recommends that agencies designate a planning team (which can be an existing management group) to provide direction for the effort.

The assessment process requires time and effort to obtain a balanced view of data needs and existing processes. However, for most agencies this will be time well spent. Results of the process can help agencies (1) determine additional data-related resource and improvement priorities and (2) set the stage for more effective longer term data management. Added benefits include the relationship building and increased understanding of data uses, needs and management practices that result from dialog among managers, data users, data providers, information technology specialists, and others who participate in the assessment process.

Pay Off: Continuing Improvement to Data Value

In addition to providing a structure for conducting a data self-assessment, the Guide also includes a process that agencies can use to develop and monitor a coordinated plan for continuing data improvements. This process can be followed to ensure that the assessment process yields concrete results for the agency—in the form of better quality, highly accessible and useful data contributing to better agency decision making.

CHAPTER 1

Introduction

Transportation Data: The Invisible Asset

Transportation agencies are responsible for managing various physical assets, such as roads, bridges, and traffic signals. These assets are created and maintained to provide the traveling public with safe, efficient, and convenient travel options. Because of the importance of these assets for supporting the agency mission, transportation agencies strive to ensure that assets are designed to serve traveler needs and are effectively managed throughout their life cycles to provide maximum value.

One asset owned by virtually all transportation agencies—yet often overlooked—is data. Agencies are investing thousands of dollars each year to collect, store, and manage data. Agencies can derive substantial value from their data investments—for example, to target high crash locations for investigation, inform travelers of detours or congestion, efficiently route snowplows, identify urgent bridge replacement needs, and monitor key indicators of agency performance.

However, when data is not treated like an asset, agencies do not derive full value from data investments. Agencies may be data rich, but have difficulties transforming their data into usable information. Treating data like an asset means

- **Inventory:** understanding what data you have and how it is being used;
- **Valuation:** making sure that investments in data are paying off in terms of improved agency decisions, improved customer service, reduced risks, and enhanced accountability;
- **Life Cycle Management:** recognizing that data requires care and feeding throughout its life cycle to be useful and usable; and
- **Accountability:** making sure there are people within the organization who are responsible and accountable for managing data to maximize its value.

Lack of coordinated, agency-wide planning and orchestration of data collection, management, and presentation can lead to both inefficiencies and missed opportunities for transportation agencies:

- Data may be collected, but not updated on a regular cycle, leading to a very short shelf-life.
- Data may be collected but not well used because of lack of sufficient quality, convenient access, or documentation.
- Data may be collected but not well used because the data cannot easily be integrated with other data to produce meaningful information.
- Data may be duplicated across different business units of an agency, resulting in higher than necessary costs for database administration and storage—and potentially leading to conflicting information on different agency reports.
- Some data may continue to be collected that once served a purpose but is no longer adding value—while pressing needs for actionable information in other areas go unmet.
- Staff resources may be overstressed by the need for fire drill-like responses to meet time-critical information requests—without tools to facilitate or automate the process.

Recognizing the benefits that transportation agencies can realize through improved data governance and management, the AASHTO Standing Committee on Planning (SCOP) has adopted a set of core principles for transportation data.[1] These principles articulate the different dimensions of what it means to manage transportation data as an asset:

- **Principle 1—VALUABLE:** Data is an asset—Data is a core business asset that has value and is managed accordingly.
- **Principle 2—AVAILABLE:** Data is open, accessible, transparent, and shared—Access to data is critical to performing duties and functions and data must be open and usable for diverse applications and open to all.
- **Principle 3—RELIABLE:** Data quality and extent is fit for various applications—Data quality is acceptable and meets the needs for which it is intended.
- **Principle 4—AUTHORIZED:** Data is secure and compliant with regulations—Data is trustworthy and safeguarded from unauthorized access, whether malicious, fraudulent, or erroneous.
- **Principle 5—CLEAR:** There is a common vocabulary and data definition—Data dictionaries are developed and metadata are established to maximize consistency and transparency of data across systems.
- **Principle 6—EFFICIENT:** Data is not duplicated—Data is collected once and used many times for many purposes.
- **Principle 7—ACCOUNTABLE:** Decisions maximize the benefit of data—Timely, relevant, high-quality data is essential to maximize the utility of data for decision making.

Transportation agencies adopting these principles—and putting them into action—should realize steady improvements to data value and an increased return on their data investments.

Purpose of the Guide

This Guide builds on the work completed in a prior scoping study conducted under NCHRP 8-36, Task 100. Completed in 2011, the object of this earlier study was to propose a framework and conceptual design to serve as the preliminary thinking for developing a resource to help transportation agencies assess the adequacy, direction, and management of their data programs.

The Guide was developed to refine and operationalize the conceptual framework outlined in the scoping study report and can be used by transportation agencies to operationalize the AASHTO data principles and strengthen management of their data assets to realize greater value. The Guide is intended to help agencies answer the following questions:

- Do we have the right data to make good decisions and meet reporting requirements?
  - What data do we need and why?
- Is our current data good enough?
  - What level of accuracy, timeliness, completeness, and so forth is needed?
- Are we getting full value from the data that we have?
  - Can users access, integrate, and analyze it?
- Are we making the best use of our data collection and management resources?
  - Are we being efficient about how we collect and manage the data?
- What do we need to improve?
  - Spot improvements (e.g., more data, different data, address specific usability issues)
  - Systemic improvements (e.g., improved governance, technical analysis, processes, skill sets, automation)

---

[1] http://planning.transportation.org/Pages/Data.aspx

The Guide features two types of assessment tools that can be used to examine current needs and practices at the agency:

- **Data Value Assessment (Data User Perspective)**—assesses the degree to which *data users* believe that existing data sets are providing value and meeting their information needs, and
- **Data Management Maturity Assessment (Agency Perspective)**—assesses the current level of *agency capabilities* for managing data assets to maximize their value.

These tools can be applied separately or in combination to build an action plan to identify priority improvements for strengthening data programs based on an evaluation of risks, costs, and benefits to the agency. If these tools are applied in combination, the resulting action plan will be strongest because the plan will reflect both user and broader agency perspectives. Table 1 summarizes how the two tools address the core questions listed above.

There is no presumption that agencies should strive for data perfection across the board. The emphasis is on improving data for decision making and making more effective use of existing data.

The *Data Value Assessment* can identify spot improvements to data availability, quality, and usability concerns of data users. Such spot improvements might include collecting additional data, cleaning up existing data sets, and providing new data visualization tools to make data more usable.

The *Data Management Maturity Assessment* can identify systemic improvements to data governance, data architecture and integration, and data quality management that can enhance data value sustainably. Systemic improvements improve agency data management capabilities rather than targeting a specific data availability or quality issue. Such improvements might include reviewing agency databases for duplication or inconsistency, defining data stewardship roles, establishing data standards, or implementing data cleansing tools. In some cases, the actions identified in the data value assessment will suggest the need for systemic improvements—for example, to address a data quality problem, one could search for and correct anomalies (a spot improvement). However, a more systemic approach would involve putting processes in place to define quality standards and validation procedures.

Although the two assessments reflect different perspectives, they are intended to work together. As agency data management capabilities are strengthened, one would anticipate that user satisfaction with data would improve.

Table 1.   Data user and agency perspectives on key questions.

	Data Value: User Perspective	Data Management: Agency Perspective
1. Right Data?	User satisfaction with data availability	Agency's understanding of user needs
2. Good Enough Data?	User satisfaction with data quality	Agency's support for quality assurance
3. Full Value from Data?	User satisfaction with data usability	Agency's efforts to integrate and provide access to data in usable forms
4. Best Use of Resources?	User perspectives on priorities for improvement, given available resources	Agency's efforts to promote data re-use and coordinated data collection across departments
5. Improvement Needs?	Spot improvements to close priority data gaps	Systemic improvements to improve agency data management capabilities

Although user perspectives are critical to consider, agencies cannot meet all of the needs expressed. The purpose of the data assessment is not to create a list that is impossible to deliver. Rather, the assessment is a tool for improving an agency's ability to make informed judgments about data investments and improve the effectiveness of the investments that are made. The assessment process can help agencies determine data priorities and optimize future data resources and investments.

Intended Audiences

This Guide and the accompanying assessment tools were developed based on input from 12 state DOTs and are primarily intended for the DOT audience. However, three metropolitan planning organizations (MPOs) were consulted during the early stages of development, and much of the guidance is sufficiently general to enable adaptation for MPOs, transit agencies, and other public-sector transportation agencies.

Within transportation agencies, the primary audiences for this Guide are

- Senior (division and district-level) managers interested in seeing the agency implement a systematic process for considering, evaluating, and prioritizing data improvement needs;
- Staff with data management responsibilities wanting to review their data products and services systematically and develop an improvement plan; and
- Business line managers (e.g., bureau chiefs) interested in making more effective use of data and wanting to systematically identify current data gaps and develop a strategy to fill these gaps.

Structure of the Guide

Chapter 2 presents an overview of the Data Self-Assessment Process and options available for scaling the process to fit agency priorities and resource availability.

Chapter 3 provides guidance for the Prepare phase of the self-assessment process in which the agency mobilizes staff to direct and facilitate the assessments and decides on an initial set of focus areas to assess.

Chapter 4 provides guidance for the Assess phase of the self-assessment process in which the agency conducts a series of workshops involving representatives of different business areas and/or data management areas. These workshops result in assessment ratings, defined gaps, and lists of candidate improvements to address the gaps.

Chapter 5 provides guidance for the Improve and Monitor phase of the self-assessment process, in which the agency reviews the results of the Assess phase and develops a coordinated plan of data improvements. This phase also includes a quarterly process to monitor the status of improvements that are in planning and implementation stages.

Appendix A provides a glossary of terms used in this Guide.

Appendix B provides templates for compiling an inventory of agency data programs.

Appendix C provides a detailed description of the data management assessment elements.

Appendix D provides a data improvement catalog that includes examples and references.

CHAPTER 2

The Assessment Process

Overview

The assessment process has three phases as shown in Figure 1:

- **Prepare**—mobilization and scoping for the assessment process
- **Assess**—conduct of the assessment using the available tools
- **Improve and Monitor**—development of an action plan based on the assessment findings, and monitoring the implementation of this plan

Each of these phases is important:

- The *Prepare* phase ensures that the entire assessment process will be productive and manageable, scoped appropriately, and with involvement of the right people.
- The *Assess* phase is when various groups in the agency meet to conduct the assessments and agree on ratings and potential actions. This phase produces valuable information on the agency's current data capabilities and gaps.
- The *Improve and Monitor* phase is where the agency decides how to act to get more value from data. This phase also involves monitoring activities to ensure that the identified improvements are implemented. Without the Improve and Monitor phase, the assessment process will have educational value, but will produce no lasting impacts. Agencies need not create new monitoring processes—they can use existing management reporting processes already in place.

Figure 1 illustrates a cyclical process. The data assessment will not be a one-time activity, but repeated annually or bi-annually to track progress and update action plans. Because some parts of the assessment are geared toward application at the level of a particular business unit or function (as opposed to agency-wide), agencies may take a phased approach to the assessment. For example, agencies might spread the assessment of data within six key business areas over a 2-year period—tackling three areas each year.

The following three concepts are reinforced throughout the entire assessment process:

- *One size does not fit all*—Transportation agencies differ in goals, issues, business needs, and the ways they manage data. The scope of the assessment can be tailored to fit with agency priorities, data issues, or other current agency data-related initiatives. These activities can also be scaled to match resource availability and time constraints.
- *Sometimes less is more*—Limiting the number of areas selected for the assessment can help ensure that the process is manageable and sustainable, given competing work activities and agency priorities. Focusing improvements on achievable actions minimizes risk and produces clear value and benefit so as to ensure that the results of the process are not diminished by trying to take on too much.
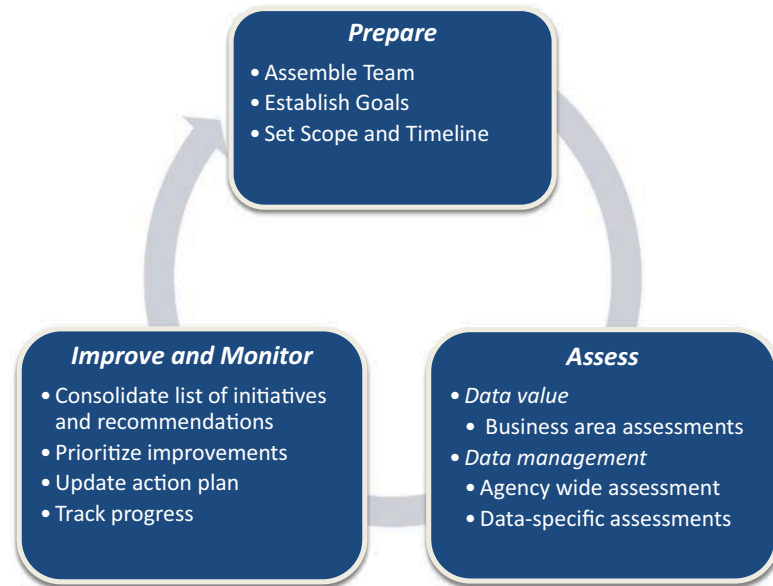
Figure 1.   Data assessment process.

- *The process can be as valuable as the results*—The relationship building, discussions, and increased understanding that occur among data users, data providers, and information technology (IT) personnel can often be as valuable as the assessment results.

Key Elements of the Assessment

The Assess phase is designed to help agencies investigate both data user and manager perspectives. The data value assessment emphasizes the user perspective and considers three distinct elements that together determine whether data is adding value for an agency's business processes:

- **Data Availability** addresses whether or not the agency has the right kinds of data in place, at the right level of detail, and with sufficient coverage to meet its business information needs. Example: if a project manager needs to understand how much of the budget has been expended, but there are no tracking systems in place for this, one would say that expenditure data is not available.
- **Data Quality** addresses whether or not the available data is good enough to meet the agency's information needs. The assessment looks at three aspects of data quality of particular concern to data users: currency, accuracy, and completeness. Example: if a project manager gets budget status reports, but the reports are 1-month old or only include internal staff charges but not contractor costs, one would say that expenditure data is not of sufficient quality. Additional aspects of data quality are considered under the data management assessment.
- **Data Usability** addresses whether or not the agency's data can be easily accessed, integrated, analyzed, and presented in a convenient form for analysis and interpretation. Example: if a project manager gets two sets of monthly reports (one for internal charges and one for contractor charges) and the manager must manually combine the reports to get the full picture, one would say that the expenditure data have poor usability.

Each of these elements must be evaluated within the context of particular business needs. A given data set may be of sufficient quality to meet one need, but not another. For example, a maintenance level of service data set based on a 10% sample of road segments might be sufficient for developing an annual statewide budget, but would not provide a basis for developing work orders or planning equipment needs for a given maintenance area.

To provide meaningful results, separate data value assessments should be applied for specific agency business functions (e.g., planning, maintenance, project scoping, or traffic operations).

The data management assessment considers the following five elements:

- **Data Strategy and Governance** is concerned with how the agency and individual business units make decisions about what data to collect and how best to manage and deliver it. This element includes establishing, enforcing, and sustaining data management strategies, roles, accountability, policies, and processes.
- **Data Architecture and Integration** is concerned with practices to standardize and integrate data. This element includes standardizing spatial referencing and other key linkages across data sets and minimizing data duplication and inconsistencies.
- **Life Cycle Data Management** is concerned with the operational aspects of managing data to ensure that it is adequately maintained, preserved, protected, documented, and delivered to users.
- **Data Collaboration** is concerned with achieving efficiencies through processes to coordinate data collection and management within the agency and partner with external organizations to share data.
- **Data Quality Management** is concerned with practices to define required levels of quality, measure and report data quality, ensure quality as new data is acquired, and improve the quality of existing data.

The data management assessment can be applied to assess agency-wide data management capabilities and an individual data management area or program to examine how one or more specific categories of data (e.g., roadway data, traffic data, and project data) are being managed. In this Guide, "data management area" and "data program" are used interchangeably to refer to an organizational function that is responsible for scoping, collecting, managing, and delivering a particular category or form of data. Sometimes this function resides in a single organizational unit; at other times it is split across business units and IT units. Examples of DOT data programs include GIS, Road Inventory, HPMS, Traffic Monitoring, Crash Records, and Construction Project Data.

Options for the Assessment Process

The assessment was designed to be flexible to meet agency needs. For example, agencies can

- Conduct the data management assessment for the agency as a whole to get a quick read on their data management capability level;
- Conduct the data management assessment for one or more target data management areas (e.g., traffic data or maintenance data);
- Conduct the data value assessment to understand user perceptions of data value in one or more business areas;
- Conduct a combination of data value and data management assessments for a logical cluster of business functions and data types to obtain a balanced perspective (e.g., a data value assessment for preservation program development and a data management assessment for pavement and bridge data);
- Pursue a comprehensive approach covering agency-wide data management and combined data value and data management assessments for priority business areas or data categories.

Further details of these options are included in the following two chapters.

This Guide and accompanying data self-assessment tools can be used to complement and/or supplement any work that agencies have done as part of safety, asset management, operations management or performance management assessments or other data-related self-assessment activities or efforts. These efforts may have produced lists of strategies that can be factored into the Improve and Monitor phase.

CHAPTER 3

Phase 1: Preparing for the Assessment

Overview

This chapter outlines the tasks and activities associated with preparing for an agency data self-assessment and focusing the effort to best meet agency needs.

Key preparation activities include the following:

1. Identify a champion to lead the self-assessment effort
2. Develop an inventory of data programs or management areas
3. Identify goals and candidate focus areas for the self-assessment
4. Obtain management support for the effort
5. Assemble a planning team and finalize the scope
6. Identify assessment team participants
7. Select and prepare facilitators

Step 1: Identify a Lead

New initiatives need good leaders. A single champion should be identified as the main point of contact responsible for leading, coordinating, and conducting the data self-assessment process and ensuring the exercise is worthwhile for the agency. The champion can begin by determining a proposed focus for the data self-assessment and gaining the endorsement and support to move forward from an executive sponsor. Once this is accomplished, this champion can establish a planning team to lead and oversee the effort. An additional staff support team of 1 to 2 individuals should also be identified at this point so they can participate from the start. These individuals will help with the assessment logistics and with compilation and presentation of results.

Responsibilities for the assessment champion include

- Reviewing this Guide and creating a list of potential focus areas for the assessment based on agency needs and priorities.
- Preparing a briefing presentation describing the assessment—what is involved and how it would benefit the agency.
- Setting up an executive briefing and enlisting support from an executive sponsor. The champion should emphasize the relationship of the process to strategic goals and business needs and let the sponsor know that the process can be implemented at different scales or using a phased approach and that it need not necessarily represent a major effort.
- Establishing a planning team to guide the effort.
- Securing and preparing 1 to 2 facilitators to conduct the assessment.
- Convening and chairing planning team meetings.
- Attending assessment sessions.

- Arranging for staff support to assist with meeting arrangements, meeting invitations, note taking and recording of meeting proceedings, documenting results, and preparing the action plan developed during the Improve and Monitor phase of the assessment process.
- Sharing the results of the process and action plan strategies with senior management and the business units that will be involved in implementing improvement strategies and actions.
- Monitoring implementation.
- Scheduling and organizing follow-up assessments.

Champions should have strong leadership and organizational skills, established relationships with the executive team, knowledge of agency data resources, and an appreciation of the importance to the agency of data stewardship and management.

Step 2: Develop an Inventory of Data Programs

A data program (or "data management area") inventory provides a useful resource for the remaining Prepare phase steps. A data program inventory provides a "map of the territory" that can be used to identify specific data management functions or business units that might be covered in the data self-assessment process. The inventory need not be comprehensive, but should include the high-value data categories and the primary business units with responsibilities for data planning, data collection, data quality assurance, and data delivery/reporting for these high-value categories. It is important to include databases that are centrally managed as well as databases managed and operated by individual business areas, districts, or regions if those databases are considered of high value to the agency. The data inventory can serve as a master list of candidate areas to include in the assessment during Step 3. A data inventory initially can be prepared in skeleton form and filled in over time.

A data program inventory is not the same as an inventory of IT applications and centrally managed databases. It is also not the same as a data dictionary or data catalog. These types of data inventories are useful, but can be time-consuming to compile and are, therefore, viewed in this Guide as improvements that the agency may wish to pursue, rather than as a necessary activity conducted as part of the Prepare phase.

Appendix B provides templates for compiling a data program inventory.

Step 3: Identify Goals and Candidate Focus Areas

As described in Chapter 1, the data self-assessment includes two different types of tools, each of which can be applied to different business units. The champion's first task is to review the options carefully and, based on clearly defined goals for what the agency wants to achieve from the assessment, develop a proposed approach to applying the assessment tools in the organization. This proposal can be refined in later steps, but it provides a solid point of departure for the effort.

Specific Pain Points and Opportunities

One way to identify goals for the assessment is to make a list of current visible problem areas related to data. An assessment could document these issues systematically and provide an opportunity to develop balanced plans to address them. Worksheet 1 provides a format for considering possible pain points—specific examples of each type of "data pain" can be documented in the second column. Blank rows are provided at the end for additional entries. The champion can complete this worksheet based on knowledge of the agency and/or discussions with key managers and staff members.

Worksheet 1.   Known data pain points.

	Data Pain Points	Specific Examples
❑	Difficulty compiling the data needed for a major planning effort	
❑	Difficulty gathering and integrating data needed to produce agency performance reports	
❑	Inability to comply with current or emerging external reporting requirements	
❑	Emerging agency policy initiatives or priorities that require new or different information	
❑	Reported data quality problems, including accuracy, currency, completeness, and reliability	
❑	Perception that the agency is behind its peers with respect to data management practices or data availability	
❑	Recognized data problems expressed by users; people aren't getting what they need when they need it, or it is taking too much work to get the data into a useful form	
❑	Risk of data loss associated with informally or unmanaged databases	
❑	Lack of documentation leading to potential misuse of data	
❑	Loss of key staff with specialized knowledge of key data sets	
❑	A large perceived mismatch between money spent on data collection and the value being realized (e.g., pavement data being collected but not used for making any decisions)	
❑	High-value databases are owned and operated by individual business areas (silos) and are not easily integrated, shared, or accessed	
❑	Data is not being analyzed or used to provide actionable information	

A second activity that could help to establish focus areas is to identify upcoming initiatives that could benefit from the data assessment. The assessment process can piggyback on or feed into existing data and IT improvement initiatives. Worksheet 2 provides a format for identifying these opportunities.

Assessment Goals and Tool Selection

Identified pain points and opportunities can be used to formulate goals for the assessment and identify which of the assessment tools to apply to achieve these goals. Worksheet 3 lists possible goals and maps these goals to appropriate assessment tools.

Selecting an Assessment Option

The exercises in Worksheets 1 through 3 provide good preparation for selecting an assessment option. Four basic options are listed below—in order of level of effort required and degree of comprehensiveness. The first two options involve application of the data management assessment only; the third option involves application of the data value assessment only; and the fourth option involves application of both tools:

- **Option 1:** Data Management Assessment for agency-wide practices
- **Option 2:** Data Management Assessment for one or more priority data categories

Worksheet 2.   Upcoming initiatives that a data assessment could inform.

	Type of Initiative	Notes
❑	Major System Implementation or Replacement (e.g., financial, ERP, asset management, traffic monitoring, crash data)	
❑	Data Business Planning or Governance Initiative	
❑	Data Warehouse/Data Integration	
❑	Performance Measurement or Management Initiative	
❑	Asset Management Initiative	
❑	Reorganization of IT Services	
❑	Reassignment or Reconfiguration of GIS Services and Resources	
❑	Enterprise Architecture	
❑	Other:	

- **Option 3:** Data Value Assessment for one or more priority business areas
- **Option 4:** Data Value Assessment for one or more priority business areas combined with data management assessment for related data categories—may also include an agency-wide data management assessment.

Each of these four options is described further below. Each option is illustrated using a version of the Assess block of Figure 1 that highlights applicable portions of the assessment process. For example, given that Option 1 entails an agency-wide Data Management Assessment

Worksheet 3.   Assessment goals and associated tools.

	Assessment Goal	Data Value	Data Management (Agency-Wide)	Data Management (Program Specific)
❑	Get an agency-wide view of how to strengthen data stewardship, data integration, and coordination across business units.		✔	
❑	Look at one or more major data management areas (e.g., crash, inventory, and financial) and identify gaps, risks, and opportunities to improve efficiencies.			✔
❑	Look at one or more business functions (e.g., planning, asset management, and project development) and understand how well existing data is working to meet business needs.	✔		
❑	Better understand employee perceptions of data availability, quality, and usability from a data user perspective.	✔		
❑	Track progress in improving data management against user perceptions of data value over time.	✔	✔	✔

Assess

Data Value

   - Business area assessments

Data Management

   - Agency-wide assessment

   - Data-specific assessments
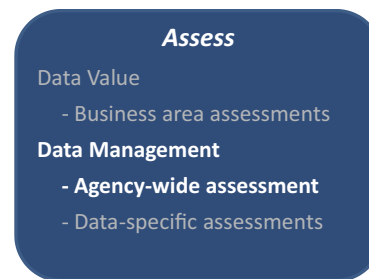
Figure 2.   Option 1: agency-wide data management focus.

only, the data value and data category-specific data management elements of the Assess phase are "greyed out" in Figure 2.

Option 1. Agency-Wide Data Management Assessment (single assessment team)

This first option (illustrated in Figure 2) may require the least effort for most agencies. Using a single assessment team, agencies can improve their agency-wide data management capabilities and assess whether they are managing data as a strategic asset. If an enterprise data governance group has been established, they would be the ideal team for this assessment. If there is no enterprise data governance group, then a team can be composed so as to include representatives of units that provide data services to multiple business units. These would include a GIS group that manages agency-wide spatial data sets, an enterprise data warehouse team, and an enterprise business intelligence or reporting team or a group responsible for managing data collection initiatives serving multiple business units.

Option 2. Data Management Assessment for One or More Data Categories

This option is illustrated in Figure 3. This option can be selected when an agency wants to assess data management practices within one or more specific priority data area(s) (e.g., road inventory, traffic, safety, and real-time operations data). If this option is selected, champions will need to work with their planning team to determine what specific data management areas to focus the data self-assessment effort on (e.g., traffic, safety, pavement, bridge, maintenance, and operations).

Option 3. Data Value Assessment

This third option, illustrated in Figure 4, can be used to focus on how users of data perceive its value—for one or more agency functional areas (e.g., providing traveler information,

Assess

Data Value

   - Business area assessments

Data Management

   - Agency-wide assessment

   - Data-specific assessments
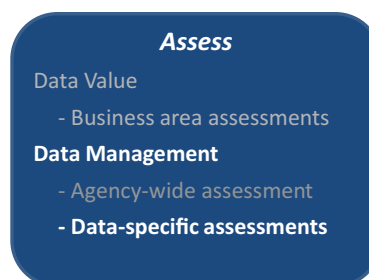
Figure 3.   Option 2: data management focus for specific data category.

> **Assess**
>
> **Data Value**
>   **- Business area assessments**
> **Data Management**
>   - Agency-wide assessment
>   - Data-specific assessments
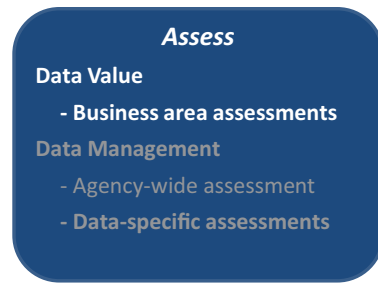
Figure 4.   Option 3: data value only.

project scoping, and maintenance budgeting). This option enables agencies to assess whether or not they have the right or "good enough" data to make decisions, meet business needs, and address reporting requirements. Option 3 can help agencies determine whether or not they are getting full value from the data they have and which supporting data programs need improvement. If this option is selected, champions will need to decide what the specific focus and goals of the assessment will be. Data value assessments are designed to address how data is meeting needs within a specific business functional area. However, such assessments can be tailored to include multiple business units as needed to address a given policy or strategic focus area (e.g., safety, infrastructure preservation, or customer service), a special initiative (e.g., open data), or a compliance need (e.g., MAP-21). Another approach is to choose a type of data (e.g., traffic or road inventory) and involve the major consumers of this type of data to assess its value.

Option 4: Combined Data Value/Data Management Maturity Assessment

This option, illustrated in Figure 5, can be used to examine a cluster of business areas and their supporting data programs from both user and management perspectives. This option may include doing the agency-wide data management assessment in addition to a set of more focused data management assessments. For example

- **Agency Functions:** Maintenance Budgeting and Maintenance Work Tracking
- **Supporting Data Categories:** Road Inventory, Maintenance Work Orders, Maintenance Level of Service, and Maintenance Budgets and Expenditures.

This option requires the most effort and provides a more comprehensive view of both data management and data quality perspectives.

> **Assess**
>
> **Data Value**
>   **- Business area assessments**
> **Data Management**
>   **- Agency-wide assessment**
>   **- Data-specific assessments**
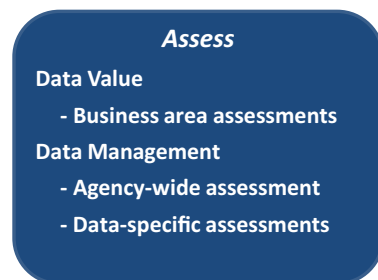
Figure 5.   Option 4: combined data value and management assessment for selected agency business areas and supporting data.

Worksheet 4.   Data self-assessment option.

	Option	Candidate Focus Areas
❏	1. Data Management Assessment: Agency-Wide	NA
❏	2. Data Management Assessment: Selected Programs	Use Worksheet 5 to identify potential data programs
❏	3. Data Value Assessment	Use Worksheet 6 to identify potential business functions
❏	4.Comprehensive Assessment—both data value and Data Management	Use Worksheets 5 and 6 to identify data programs and business functions

Identifying Candidate Focus Areas

If Option 2, 3, or 4 is selected, you will need to identify specific focus areas for the data value and/or data management assessments. Selected focus areas should be manageable. It is important to establish realistic expectations for the assessment process and limit the scope to what can reasonably be accomplished given competing work priorities. It is also advisable to identify several different candidate areas and finalize the selections pending further discussions with the agency leadership and the planning team. Worksheet 4 provides a format for recording your recommended assessment option and candidate focus areas.

Some agencies may have already undertaken data self-assessments as part of comprehensive highway safety planning, including the Crash Data Improvement Program (CDIP) and the Roadway Data Improvement Program (RDIP). Agencies may find this Guide and the accompanying tools useful to provide a higher level assessment of agency data programs while relying on existing safety assessment tools for more detailed assessments and diagnostics to identify safety-related data management improvements. Alternatively, agencies could use the completed safety assessment to inform portions of the assessment tools included in this Guide.

Step 4: Obtain Management Support

Before proceeding with the self-assessment, it is important to identify an executive sponsor and ensure that the effort has strong management support and is considered a priority activity. Management support is necessary to make the necessary staff resources available to participate in the assessment. If management is not interested in seeing and acting on the results of the assessment, the opportunity to use the effort to make a positive and lasting effect on the organization will be diminished.

Management support can be secured through a combination of one-on-one conversations and more formal briefings that describe the purpose and expected outcomes of the assessment, the timeframe, and the expected level of effort that will be required. Managers need to have confidence that the process will produce objective and balanced results, rather than a long list of new data to collect. Ideally, a briefing would include members of the senior management team—particularly those whose responsibilities include data programs or business areas that have been identified as assessment candidates. Managers would be asked to endorse proceeding with the effort and weigh in on the selection of focus areas and planning team members. Such managers can also be asked to send an email (drafted by the champion) to prospective planning team members requesting them to participate in the effort and attend an initial meeting.

Worksheet 5.   Select candidate DOT data programs for data management assessment.

Candidate Programs
❑ GIS/LRS—management of base maps, road centerlines, linear referencing system (LRS), area boundaries, and feature locations ❑ Road Inventory—management of roadway characteristics (may be combined with GIS/LRS) ❑ HPMS—management of data required for federal HPMS reporting (may be combined with above elements) ❑ Traffic Monitoring—management of traffic count, AADT, and vehicle miles traveled (VMT) data ❑ System Operations—management of information about travel time, delay, reliability, and incidents (may be combined with Traffic Monitoring) ❑ Travel Demand—management of travel demand model inputs and outputs ❑ Active Transportation—management of information about pedestrian and bicycle travel and pedestrian and bicycle facilities ❑ Freight—management of information about freight transportation supply and demand ❑ Crash—management of crash and fatality data ❑ Pavement—management of pavement inventory and condition data ❑ Bridge—management of bridge inventory and condition data ❑ Financial—management of financial data (e.g., revenues, obligations, budgets, expenditures, vouchers, and payments) ❑ Capital Program—management of information about scope, schedule, budget, and funding allocations for capital projects (e.g., S/TIP data) ❑ Construction Information—management of design and as-built plans and related documents ❑ Facilities—management of facility inventory and condition information (e.g., buildings, park-and-ride lots, rest areas, airports, ferry terminals, and rail stations) ❑ Maintenance—management of information about planned and completed maintenance activities and associated equipment, labor, and materials ❑ _____ ❑ _____ ❑ _____ ❑ _____ ❑ _____ ❑ _____

Step 5: Assemble the Planning Team and Finalize Scope

The next step is to assemble a planning team of 3 to 6 individuals to help set direction, review findings and results, and oversee action planning. The planning team can refine assessment goals, identify priorities and specific issues or questions to be addressed, and set the scope and timeframe for the assessment process. They can help the champion identify individuals to include in the Assess phase and serve as a sounding board for validating assessment results and action plan recommendations.

Ideally, the members of the planning team could allocate resources to data improvement strategy implementation. Members of the planning team could include a mix of

- Business area managers or supervisors in the areas picked for the assessment focus
- Data program managers or supervisors
- IT managers or supervisors
- Principal data system or application owners or stewards

Worksheet 6.   Select candidate DOT business functions for data value assessment.

Candidate DOT Business Functions
Planning and Programming
❑ Agency Performance Management
❑ Long-Range/Corridor Planning
❑ Freight Planning
❑ Safety Planning
❑ Pavement Management
❑ Bridge Management
❑ Asset Management
❑ Program Development and Management (STIP)
❑ Bicycle/Pedestrian Planning
❑ _____
Maintenance and Operations
❑ Maintenance Budgeting
❑ Maintenance Management
❑ Equipment Management
❑ Facilities Management
❑ Traffic Operations Management
❑ _____
Project Development & Delivery
❑ Project Scoping
❑ Right-of-Way and Utilities
❑ Environmental
❑ Survey and Roadway Design
❑ Traffic Engineering
❑ Pavement Design
❑ Bridge Design
❑ Construction Management
❑ Materials/Testing
❑ Federal Reporting
❑ _____
Modal Operations
❑ Aviation
❑ Transit/Rail
❑ Marine
❑ _____
❑ _____
Other
❑ Public Affairs and Communication
❑ Driver Licensing and Vehicle Registration
❑ _____

Note: A group with a similar membership composition may already exist (e.g., a data governance group or an IT investment review board). If this group is willing to take on the responsibilities of the planning team, it is not necessary to form a new group for the assessment. In fact, it is preferable to piggyback on established decision-making groups.

Once the members of the planning team are identified and agree to participate, their first order of business is to review and refine the assessment objectives and prioritize the candidate

business areas and/or data programs from among the candidates identified by the champion. Several factors may be considered in this process:

- Alignment with Agency Priority—the degree to which the data program or business area supports a priority area for the agency and the degree to which data improvements are critical for making progress toward established objectives
- Impact—the degree to which an assessment would be likely to bear fruit (e.g., support actions to improve agency efficiency and effectiveness)
- Willingness and ability of business units/data program managers to participate (given workload and other initiatives)
- Stability of current business processes and systems—if the unit or its main data systems are undergoing transition, it may not be the best time for an assessment
- Staff experience—if there has been recent staff turnover and current staff is very new, it may be difficult for them to provide a valid assessment of current practice or needs

If the goal is to get a broad view across the agency covering several different business units and/or data programs, one option is to develop a staggered schedule for the assessments—covering three to four assessments each quarter. Either the champion or members of the planning team can then secure agreement to participate from the managers of each of the selected groups and schedule them for a period that fits with their other activities.

Decisions about how many areas to assess should be informed by an estimate of time and resource requirements for each assessment. Worksheet 7 can be used to develop these estimates (this does not consider startup time for the planning team or the facilitators). Estimated hours shown are for each assessment conducted, with the exception of the last one—the summary presentation can cover results from multiple assessments. This worksheet assumes that a support staff person helps with scheduling meetings and conference rooms, preparing handouts, compiling questionnaire results, and preparing summary presentations.

Members of each assessment team would be expected to allocate between 1 and 2 days of time, spread across three workshop sessions. The facilitator and support staff would spend an additional 2 to 3 days preparing and synthesizing results between workshops. Further information on the activities listed can be found in Chapter 4 of this Guide.

Worksheet 7.   Assessment time and resource requirements.

Activity	Staff Involvement	# Hours	Meeting Room?
PHASE 2: Assess			
Assessment Preparation Meeting	Champion + Staff	4-8	No
Workshop 1—Assessment	Full Assessment Team	4-8	Yes
Preparation Meeting for Workshop 2	Facilitator, Champion + Staff	2-4	Yes
Workshop 2—Gaps and Candidate Actions	Full Assessment Team	2-4	Yes
Summary of Assessment Results	Facilitator + Staff	2-4	No
PHASE 3: Implement and Monitor			
Workshop 3—Action Planning	Full Assessment Team	2-4	Yes
Action Synthesis	Facilitator, Champion + Staff	2-4	No
Results Presentation to Planning Team and Executive Sponsor (covering results from multiple assessment teams)	Champion and Facilitator	1-2	Yes

Step 6: Identify Assessment Participants

Once business areas and/or data programs have been identified and scheduled, assessment participants can be identified. Each group should have a minimum of 4 and a maximum of 15 participants. Fewer than 4 participants may not yield a sufficiently balanced perspective, but more than 15 participants may be unwieldy to manage and it may be difficult to achieve a consensus set of assessment ratings.

The following guidelines can be used to select participants:

- **Data Value Assessment:** The objective is to obtain a balanced set of perspectives about data availability, quality, and usability to support a particular business function. Therefore, there should be representation from
    - Senior staff who request information (decision makers) and more junior staff involved in hands-on data gathering and preparation activities (data analysts)
    - Staff responsible for different aspects of the business area that depend on (or could benefit from) quality data
    - Staff in both central office and field offices (where applicable)
    - IT and data program managers—their participation will enable them to learn more about how data is being used and answer any questions on quality and availability
- **Data Management Assessment—Program Specific:** The objective is to obtain an objective perspective about the extent to which different data management practices are being carried out. There should be representation from
    - Managers responsible for overseeing the data program
    - Data program staff responsible for hands-on data management tasks (e.g., quality assurance, data loading, report development, and data request fulfillment)
    - IT or GIS group staff who support data management tasks (as appropriate)
    - Key customers of the data program (as appropriate)
- **Data Management Assessment—Agency-wide:** Participants should be similar to those for the program-specific data management assessment—with managers, staff and customers of agency-wide functions (where they exist) for data governance, metadata management, data architecture, data security, data warehousing, data integration, reporting/business intelligence, and GIS.

Step 7: Select and Prepare Facilitators

This step—selecting the right people to facilitate the assessment and ensuring that they are well prepared—is critical to the success and value of the assessment. The following are essential characteristics of a good facilitator:

General Facilitation Skills
- Someone who participants will trust and view as neutral, with no specific agenda other than to achieve an objective assessment;
- Ability to work through an agenda within the allotted time, and ability to keep the discussion focused on the topic at hand and diplomatically steer the discussion back if it gets off track;
- Ability to ensure that perspectives are heard from all of the participants—not just the senior managers or the most naturally vocal people in the group;
- Ability to ensure that all participants have a consistent understanding of what each person is saying—by asking speakers to provide clarification or examples and by restating comments in different words and asking for confirmation;
- Ability to guide a group to a consensus opinion—by crystalizing and restating differences in perspectives and clarifying assessment criteria definitions so the group can reach agreement.

Substantive Knowledge

- Understanding of agency business processes and data—enough so that participants do not need to spend a lot of time explaining the basics of their program or business function to the facilitator.
- Sufficient experience with data management practices and concepts to be able to understand the terminology used in the assessment and explain it in terms that participants will understand (this is particularly important for data management assessments, but helpful for data value assessments as well).

Ideally, agencies would identify two facilitators to share the load, provide more scheduling flexibility, and ensure that a backup facilitator is available if one leaves the agency.

Once facilitators are identified, they should review the material in this Guide carefully and ensure that they understand and can clarify all of the terminology and distinguish across the different assessment rating scales (i.e., maturity levels for the data management assessment; good-fair-poor ratings for the data value assessment).

CHAPTER 4

Phase 2: Conducting the Assessment

This chapter provides specific guidance for the champion and the facilitators on conducting data value and data management assessments. It assumes that agreement has been reached on a set of assessments to conduct and participants have been identified for each assessment team based on the guidance in Chapter 3. Following a summary of each assessment tool, step-by-step guidance is provided for conducting assessments and summarizing the results that will be required for Phase 3.

Data Value Assessment Tool Overview

The Data Value Assessment Tool takes the assessment team through a process of rating the availability, quality, and usability of data required to meet a defined set of business functions. As described in Chapter 2:

- **Data Availability** addresses whether the agency has the right kinds of data in place, at the right level of detail, with sufficient coverage to meet its information needs. Example: if a project manager needs to understand how much of their budget has been expended, but there are no tracking systems in place for this, one would say that expenditure data is not available.
- **Data Quality** addresses whether the data available are good enough to meet the agency's information needs. The data value assessment looks at three aspects of data quality of particular concern to data users—currency, accuracy, and completeness. Example: if a project manager gets budget status reports, but they are 1-month old they may not be sufficiently current or timely. If reports only include internal staff charges but not contractor costs, one might say that the expenditure data is not sufficiently complete. If there are known errors or inconsistencies between reports, the data might not be sufficiently accurate or reliable to meet agency needs. Additional aspects of data quality are considered under the data management assessment.
- **Data Usability** addresses whether the agency's data is being provided in a convenient form for analysis and interpretation. Data usability includes consideration of how easily data can be accessed and how well it is integrated, analyzed, and presented in a convenient form for users and customers. Example: if a project manager gets two sets of monthly reports—one for internal charges and one for contractor charges, and the manager must combine them manually to get the full picture, one would say that the expenditure data has poor usability.

Availability is assessed with respect to specific business activities. **Quality** and **Usability** are assessed for each of the major data sources used for performing the selected business activities. The assessment categories ratings are Excellent, Good, Fair, and Poor. To calculate weighted ratings across various data sources, the assessment also asks team members to rate the importance of each data source to each business activity. Figure 6 summarizes the different elements and ratings of the data value assessment.

22

Element	Rating
Data availability—is data available at the right level of detail, with sufficient coverage?	🔴 **Poor.** Little or no data available to support this activity 🟠 **Fair.** Limited data available—large gaps remain 🟢 **Good.** Basic data is available—some gaps remain 🟢 **Excellent.** Sufficient data is available to meet needs
Data quality—are data sufficiently accurate, credible, complete, and current to support decision making?	🔴 **Poor.** Quality not sufficient—data not useful 🟠 **Fair.** Lack of currency, accuracy, or completeness limits value 🟢 **Good.** Acceptable but needs improvement 🟢 **Excellent.** Sufficient to meet needs
Usability—can data be easily accessed, integrated, analyzed, and presented as needed to support decision making?	🔴 **Poor.** Requires substantial effort to get data into usable form 🟠 **Fair.** Requires moderate effort to get data into usable form 🟢 **Good.** In usable form but reporting improvements helpful 🟢 **Excellent.** In usable form and no improvement is needed

Figure 6.   Data value assessment elements and ratings.

Once an agency has applied the Data Value Assessment Tool for all of the key major functional areas, a high-level view, such as that shown in Figure 7, can be developed to highlight variations in how well data is working to meet agency needs. Each data value assessment provides one row of this chart.

Data Management Assessment Tool Overview

Overview of the Tools

Two data management tools are available—one for an agency-wide assessment and one for an assessment of a specific data program or data category. These tools are similar in structure, but vary in the assessment elements and criteria. The data management tools take the agency-wide or data program management assessment team through a process of rating current data management processes. As described in Chapter 2, the following elements are considered:

- **Data Strategy and Governance** is concerned with how the agency and individual business units decide what data to collect and how best to manage and deliver it. This area of concern

Business Area	Availability	Quality	Usability
Maintenance Management	Good	Fair	Fair
Pavement Management	Excellent	Good	Good
Safety Planning	Excellent	Good	Fair
Performance Management	Fair	Fair	Good
Project Scoping	Good	Fair	Good
Construction Management	Good	Good	Fair
Corridor Planning	Good	Good	Poor

Figure 7.   Sample data value assessment summary.

includes establishing, enforcing, and sustaining data management strategies, roles, account-ability, policies, and processes.

- **Data Architecture and Integration** is concerned with practices to standardize and integrate data and includes standardizing spatial referencing and other key linkages across data sets and minimizing data duplication and inconsistencies.
- **Life Cycle Data Management** is concerned with the operational aspects of managing data to ensure that it is adequately maintained, preserved, protected, documented, and delivered to users.
- **Data Collaboration** is concerned with achieving efficiencies through processes to coordinate data collection and management within the agency and to partner with external organizations to share data.
- **Data Quality Management** is concerned with practices to define required levels of quality, measure and report data quality, ensure quality as new data is acquired, and improve the quality of existing data.

Each of these areas is broken into a set of assessment sub-elements. For each sub-element, different maturity levels are defined that characterize a progression from an ad hoc approach to data management to an approach that is well-defined, documented, and institutionalized within the agency or data program.

The complete set of assessment elements and sub-elements is provided in Appendix C, along with the criteria for each maturity level, a mapping of each sub-element to the AASHTO data principles, a discussion of the benefits of moving up the maturity scale, and a listing of relevant improvement actions that can be considered. Table 2 summarizes the different data management assessment elements, sub-elements, and maturity levels. Two sub-elements are only applicable for the agency-wide assessment; one is only applicable for data-program-specific assessments.

Once an agency has completed a data management assessment for agency-wide functions and/or specific data programs or categories, a high-level view such as that shown in Figure 8 can be developed to highlight variations in how well data is working to meet agency needs. Each data management assessment provides one row of this chart.

Table 2.   Data management assessment elements and maturity levels.

Element/Sub-element	Description	Maturity Levels
1. Data Strategy and Governance	Leadership and management practices to manage data as a strategic agency asset	**1: Initial:** Ad hoc and event driven, success due to heroic efforts of individuals
1.1 Strategy and Direction	Leadership commitment and strategic planning to maximize value of data to meet agency goals	**2: Developing:** Recognized need for improvement, pilot initiatives under way
1.2 Roles and Accountability	Clear roles, accountability, and decision-making authority for data quality, value, and appropriate use	**3: Defined:** Defined and documented processes not yet stabilized or widely socialized
1.3 Policies and Processes (Agency-wide Only)	Adoption of principles, policies, and business processes for managing data as a strategic agency asset	**4: Functioning:** Implemented processes—operating and adding value
1.4 Data Asset Inventory and Value	Tracking of agency data assets and their value added	**5: Sustained:** Evaluated and improved processes, sustained over time
1.5 Relationships with Data Customers (Program Specific Only)	Connections between data producers and users	

Table 2.   (Continued).

Element/Sub-element	Description	Maturity Levels
1.6 Data Management Sustainability	Continuity of data management knowledge and expertise through staff transitions	
2. Data Life Cycle Management	Practices for managing data throughout its life cycle from collection to archiving or deletion	
2.1 Data Updating	Well-defined and coordinated data update cycles	
2.2 Data Access Control	Well-defined policies and guidelines for managing access to data sets	
2.3 Data Findability and Documentation	Availability of data catalogs and dictionaries that enable discovery and understanding of available agency data assets	
2.4 Data Backups and Archiving	Guidelines and procedures for protection and long-term preservation of data assets	
2.5 Data Change Management	Processes to minimize unanticipated downstream impacts of data changes	
2.6 Data Delivery	Delivery of data to users in various convenient, useful, and usable forms	
3. Data Architecture and Integration	Technical standards, processes, tools, and coordination mechanisms to maximize data integration and minimize duplication	
3.1 Location Referencing	Common location referencing methods across agency data sets	
3.2 Geospatial Data Management (Agency-wide Assessment Only)	Standardized approach to managing geospatial data	
3.3 Data Consistency and Integration	Standards and practices to ensure use of consistent coding and common linkages so that different data sets can be combined to meet business information needs	**1: Initial:** Ad Hoc and event driven, success due to heroic efforts of individuals
3.4 Temporal Data Management	Standardization of date-time data elements to enable trend analysis and integration across data sets collected and updated on varying cycles	**2: Developing:** Recognized need for improvement, pilot initiatives underway
4. Data Collaboration	Internal and external collaboration to maximize data sharing and avoid duplication of effort	**3: Defined:** Defined & documented processes not yet stabilized or widely socialized
4.1 Internal Agency Collaboration	Collaboration across agency business units to leverage opportunities for efficiencies in data collection and management	**4: Functioning:** Implemented processes—operating and adding value
4.2 External Agency Collaboration	Partnerships with external entities to share data and avoid duplication	**5: Sustained:** Evaluated & improved processes, sustained over time
5. Data Quality	Standards and practices to ensure that data is of sufficient quality to meet user needs	
5.1 Data Quality Measurement and Reporting	Metrics and reporting to ensure user understanding of current data quality	
5.2 Data Quality Assurance and Improvement	Practices for improving the quality of existing data and ensuring the quality of newly acquired data	

Data Program	Strategy & Governance	Life Cycle Mgt.	Arch. & Integration	Collab-oration	Quality	Overall Level
Agency-wide	2-Developing	3-Defined	2-Developing	2-Developing	Not Assessed	2-Developing
Traffic Monitoring	3-Defined	4-Functioning	3-Defined	5-Sustained	4-Functioning	4-Functioning
Crash Data	5-Sustained	4-Functioning	3-Defined	5-Sustained	4-Functioning	4-Functioning
Pavement Inspection	1-Initial	4-Functioning	3-Defined	1-Initial	5-Sustained	3-Defined
STIP/Capital Projects	3-Defined	5-Sustained	1-Initial	2-Developing	2-Developing	3-Defined
Financial	5-Sustained	5-Sustained	4-Functioning	Not Assessed	4-Functioning	4-Functioning

Figure 8.   Sample data management assessment summary.

Assessment Phase Guidance

Purpose of the Assessment

The purpose of the assessment is to identify opportunities to improve decision making and make more effective use of existing data. Agencies do not need to strive for data perfection across the board. Such data perfection would neither be feasible, given resource constraints, or necessarily desirable from a benefit-cost perspective. The assessment tools provide a framework within which agencies can identify the current state of data and the current state of data management practices. This provides a baseline for discussion about potential improvements. Although application of the assessment tools will suggest potential improvements, the agency assessment teams need to evaluate whether or not each type of improvement makes business sense. For example, higher levels of data management maturity are typically characterized by formal documented processes and procedures. These can require considerable investments in staff time to create, maintain, and operationalize within the agency. In some cases, these investments are worth it (e.g., where an undocumented, chaotic process creates unacceptable risks of providing inaccurate performance data to the state legislature). In other cases, formalizing processes may not be appropriate (e.g., where experimentation is being encouraged for a new type of data). In deciding whether to maintain the status quo or take steps to improve, agencies can weigh the risks of doing nothing and the likely returns from moving forward. In developing strategies for improvement, the concept of diminishing returns is useful. Agencies can strive to invest in data improvements until the marginal cost of making (and sustaining) the improvement is equal to the marginal benefits gained.

Overview of Activities

Both the data value and the data management assessment feature two workshops during the Assessment Phase. Assessment Phase activities are summarized in Worksheet 8. The worksheet identifies participants and inputs and outputs for each activity. Activities 1 through 4 are conducted separately for each assessment. Activity 5 is conducted at the end of Phase 2 to summarize the results from all of the individual assessments.

Worksheet 8.　Assessment Phase activity checklist.

	Activity	Participants and Inputs	Outputs (Results)
❏	1. Conduct Assessment Preparation Meeting (one per assessment team)	Champion, Facilitator, Staff, Managers of selected business units	Assessment Rosters Tool Configuration Workshop Agenda Workshop Invitations—from executive sponsor, champion, and/or group managers
❏	2. Conduct Assessment Workshop (one per assessment team)	Facilitator, Champion (optional), Assessment Team—bring laptops with assessment tool + projector for group assessment exercise	Consensus ratings and participant comments on key gaps
❏	3. Preparation Meeting for Gaps and Candidate Actions Workshop (one per assessment team)	Champion, Facilitator, Staff, Managers of selected business units	Summary assessment results Handouts for Gaps Workshop
❏	4. Gaps and Candidate Actions Workshop (one per assessment team)	Facilitator, Champion (optional), Assessment Team–Projector for group gap validation and action identification exercise	Validated gaps with business impacts List of candidate actions to close the gaps
❏	5. Assessment Results Analysis and Summary (a single meeting for all assessments combined)	Champion, Facilitator, Staff	Summary workshop results for presentation and synthesis with other assessment areas

Step 1: Conduct Assessment Preparation Meeting

For Data Value Teams

Tool Configuration

The Data Value Assessment Tool is generic because it is intended to be applicable for any DOT business function that depends on data. Therefore, the tool must be configured for each business area in which it is to be applied. Configuration consists of three steps—each of which can be accomplished at the Assessment Workshop Planning Meeting:

1. Specify the business area to be assessed by recording the selected area in the space provided on the tool's **Configuration** tab
2. Break your selected business area into specific activities
3. Identify types of data needed to perform these activities

Selecting Business Activities.　Breaking business areas into activities enables the assessment team to focus on specific ways that data is used (or could be used). A comprehensive breakdown of all activities for the business area is not necessary. Criteria for identifying (and describing) activities are as follows:

- Each activity should be important to the success of the overall business area.
- Each activity should be clearly and consistently understood by different members of the assessment team.

- Activities should not be fundamentally about data collection or processing—they should be activities where data is or could be used to make better decisions or respond to information needs.
- Activities can include areas that could benefit from greater data availability, not just areas where data needs are met.
- Activities can be those that the agency performs or those that the agency would like to perform in the future.

The following generic set of activities that require data can be used as a starting point for any DOT business area—these can be tailored as needed to represent activities specific to the agency:

- Monitor results or performance against established objectives
- Track expenditures, resources used, and accomplishments
- Assess future needs for budgeting or lining up new/different resources
- Diagnose root causes for limited performance or inefficiencies
- Plan, prioritize, or schedule actions to be taken

In the Data Value Assessment Tool, sample activities for different business areas are included on the **Example Lists** tab. Figure 9 shows examples of sample business areas and associated activities included in the **Example Lists** tab of the tool.

Sample Business Areas and Associated Activities					
Asset Management	**Maintenance Management**	**Project Scoping**	**Traffic Operations Management**	**Safety Planning**	**Corridor Planning**
Pavement Needs and Risk Assessment	Maintenance Budgeting	Current Conditions Assessment	Incident Management	Network Screening	Current Conditions Assessment
Pavement Resource Allocation and Treatment Selection	Maintenance Activity Tracking	Scope Development	Traveler Information	Counter-measure Analysis	Future Demand Analysis
Bridge Needs and Risk Assessment	Equipment Management	Schedule and Budget Development	Signal Timing and Coordination	Counter-measure Design	Alternatives Evaluation
Bridge Resource Allocation and Treatment Selection	Materials Management		Active Traffic Management	Counter-measure Evaluation	Strategy Prioritization
Other Asset Needs Assessment and Budgeting	Winter Maintenance—Snow Route Planning, Snow Plow Tracking			Project Prioritization	
Cross-Asset Tradeoffs					

Figure 9. Data value tool—Example Lists tab sample business areas and associated activities.

Figure 10 shows the **Configuration** tab. To use one of the standard business areas and associated activities from the **Example Lists** tab, copy them from the Example Lists and paste them into the **Configuration** tab. Otherwise, enter the business area and activities in the spaces provided.

Identifying Data Types.   The third configuration step is to identify the major types of data applicable to the collection of activities that you have defined. The idea here is not to identify all data types used for the activities comprehensively, but rather the ones that either add the most value to the results or that could (with improvement) add value to the results.

Configuration

| Apply | Reset |

Step 1: Define the **business area** under consideration

| Maintenance Management |

Step 2: List up to 20 **activities** related to this business area below (leave unfilled cells blank)

| Maintenance Budgeting |
| Maintenance Activity Tracking |
| Equipment Management |
| Materials Management |
| Winter Maintenance – Snow Route Planning, Snow Plow Tracking |
| |
| |
| |
| |
| |
| |
| |
| |
| |
| |
| |
| |
| |
| |
| |

Step 3: List up to 20 **data sources** that are most important for this collection of activities (leave unfilled cells blank)

| Maintenance Work Orders, resource utilization, accomplishment |
| Maintenance level of service/maintenance feature quality |
| Fleet/Equipment Inventory and Utilization Data |
| Snow Plow Routes |
| Materials Inventory and Utilization Data |
| Incident Data – type, status, response time |
| Road Weather Data |
| Oversize/Overweight Permitting Data |
| Outdoor Advertising Data |
| Permits (access, right-of-way, utility, detour, etc.) |
| |
| |
| |
| |
| |
| |
| |
| |
| |
| |

Step 4: Press **Apply** button

Figure 10.   Data value tool—configuring business areas, activities, and data sources.

A sample list of data types is included on the **Example Lists** tab. You can copy and paste from this list or enter your own. Criteria for identifying data types to include are as follows:

- Each type of data should be important to at least one of the defined business activities, but not necessarily to all of them.
- Each type of data should be sufficiently well-defined that members of the assessment team have a common understanding of what type is included. (For example, using the term "traffic data" could lead to confusion about whether this includes both real-time data and count data.)
- Each type of data should be specific enough so that assessment team members can rate its quality and usability. Although it is OK to include multiple data sets under a single "data type" (e.g., "environmental data" encompasses data from multiple sources), this may make it more challenging to assign meaningful and consistent ratings if there is wide variation in data availability, quality, or usability across the data sets included in the single data type.
- At least some members of the assessment team should have sufficient familiarity with each type of data to be able to rate its quality and usability.
- Up to 20 data types can be selected, but the amount of time needed to conduct the assessment process will be directly affected by the number of data types, because each data type is assessed individually. In general, selecting between 5 and 10 data types is recommended.

Once the business areas and data types have been configured, click the **Apply** button at the top of the **Configuration** tab. This updates the other tabs of the worksheet.

See Table 3 for examples of Data Value Assessment Tool configurations.

Agenda Development

A full day should be allocated for the data value assessment workshop, although it may take less time when the assessment team is smaller and when the tool is configured to include relatively few business activities and data types. A sample agenda is shown in Figure 11.

For Data Management Teams

The data management tools allow for configuration of parameters affecting how maturity levels for each element are calculated. Figure 12 shows the **Configuration** tab for the Agency-wide tool—the **Configuration** tab for the Program-specific tool is similar, but shows a slightly different set of sub-elements.

Agencies can tailor the data management elements to meet their own needs, priorities, or areas of focus. If a particular data management element is not applicable or relevant to the assessment discussion, its weight can be set to 0%. Then, any elements weighted at 0% will not affect overall assessment results.

There are two types of configuration for this tool: (1) weights on elements and sub-elements and (2) selection of a threshold value.

Adjusting weights. Sub-element weights are used in the tool to calculate maturity levels for elements based on sub-element maturity levels. Element weights are used to calculate overall data management maturity levels based on element maturity levels. By default, all sub-elements in an element are weighted equally and each of the five elements has an equal weight in calculating the overall maturity level. Weights can be adjusted as desired, but all of the element weights need to sum to 100%, and the sub-element weights for each element also need to sum to 100%.

One reason to adjust weights is if a particular element or sub-element is not applicable for your agency or the particular data program you are assessing. In that case, set the element or sub-element weight to 0 and adjust the other weights so that they sum to 100.

Table 3.   Example Data Value Assessment Tool configurations.

Business Area	Business Activities	Data Types
Maintenance Management	Track maintenance level of service Track maintenance expenditures, resources used, and accomplishments (outputs) Develop future-year maintenance budget requests Identify opportunities for improved efficiency Plan, prioritize, and schedule work	Road Inventory Maintenance Feature Inventory Maintenance Feature Condition/Performance Maintenance Work Orders Budget Allocations
Project Scoping and Design	Project Management/Project Control Prepare Design Plans Environmental Review Review Existing Conditions/Identify Needs Create Concept Reports	Road Geometry Traffic Counts, AADT, Classification Asset Inventory Environmental Data Crash and Fatality Data Construction Project Status Data Right-of-Way Data
Facilities Management	Track facility inventory and condition (Includes both buildings and system/components) Track facility capital and maintenance expenditures and work accomplishment Identify candidate projects for rehabilitation, replacement, and expansion/addition Diagnose causes of high maintenance costs or inefficiencies Prioritize candidate projects and develop resource-constrained improvement program	Facility Inventory Facility Inspection/Condition Maintenance Records Facility Improvement Program Budgets and Expenditures
Congestion/Mobility Improvement	Transportation system performance monitoring Scoping and design of candidate projects Corridor and long-range planning, multi-modal planning Real-time traffic and incident management Improvement program development/prioritization	Road Inventory Traffic Counts, AADT, Classification Bike Routes and Paths, Non-Motorized Travel Counts Real-Time Traffic Volume/Occupancy, Travel Time Construction Project Scope and Status

A second reason to adjust weights is to give priority to certain elements. For example, if you think that under data collaboration that internal collaboration is more important than external collaboration, you might assign internal agency collaboration a weight of 80% and external collaboration a weight of 20%.

Adjusting the threshold value. The threshold value affects how maturity levels are assigned to each sub-element. To assess the maturity level for each sub-element, the data management tools present five descriptions of current practice—corresponding to the five maturity levels. Because there may be situations where it is difficult to place agency practice in a single maturity level, members of the assessment team are asked to indicate the extent to which they think that each description reflects the current state of agency practice. Options are as follows:

- 1-Totally Disagree
- 2-Somewhat Disagree

Data Value Assessment Workshop Agenda	
9:00 AM	**Background** • Introductions • Assessment Purpose • Assessment Steps and Schedule
9:45 AM	**Assessment Content** • Business Activities and Data Types • Assessment Results and Definitions
10:30 AM	**Importance Ratings** • Instructions • Individual Ratings • Consensus Ratings
11:30 AM	**Availability Ratings** • Instructions • Individual Ratings • Consensus Ratings
12:30 PM	**Lunch**
1:30 PM	**Quality Ratings** • Instructions • Individual Ratings • Consensus Ratings
2:30 PM	**Usability Ratings** • Instructions • Individual Ratings • Consensus Ratings
3:30 PM	**Results Review** • Results and their Derivations • Discussion and Adjustment
4:00 PM	**Wrap-Up** • Feedback • Next Steps

Figure 11.   Data value assessment workshop sample agenda.

- 3-Somewhat Agree
- 4-Totally Agree

The tool assigns a maturity level to the practice description representing **the highest maturity level that receives a rating equal to or higher than a configurable threshold value of either 3-Somewhat Agree or 4-Totally Agree**. The default value of this threshold value is 3-Somewhat Agree. However, this can be changed to 4-Totally Agree if you only want to assign a maturity level when the assessment team indicates this strongest level of concurrence with the description of a particular level.

Figure 13 shows an example of a sub-element rating.

In this example, by default, a maturity level of 4-Functioning would be assigned for the data updating sub-element because this is the highest maturity level description that received

Figure 12.   Data management tool configuration tab.

2.1.	Data Updating: well-defined and coordinated data update cycles.	
2.1.1	Data updating cycles and business rules for data updates have not been defined	1-Totally Disagree
2.1.2	Updating cycles have been established but have not been documented	2-Somewhat Disagree
2.1.3	Updating cycles have been documented Business rules have been defined for how key data entities are added, updated, and deleted	4-Totally Agree
2.1.4	Business rules for data updating are embedded in and enforced by applications	3-Somewhat Agree
2.1.5	Data updating methods are periodically reviewed to identify opportunities for improved efficiencies	2-Somewhat Disagree

Figure 13.   Example of data management tool rating method.

a rating of at least 3-Somewhat Agree. Responses in the example indicate that the agency has defined updating cycles and business rules for data updates and has documented these updating cycles. They have to some extent embedded business rules for data updating within applications.

If the tool was configured so as to change the threshold from 3-Somewhat Agree to 4-Totally Agree, a maturity level of 3-Defined would be assigned—corresponding to the practice description in 2.1.3.

Assessment Workshop Planning Meeting: Agenda Development

A full day should be allocated for the data management assessment workshop, although it may take less time when the assessment team is smaller and when the tool is configured to exclude multiple sub-elements. A sample agenda is shown in Figure 14.

Step 2: Assessment Workshop

Separate guidance is provided for data value and data management assessment teams. However, the facilitator should be prepared for both types of teams to address the following sets of assessment challenges:

- Understanding the context for the assessment. Ensure that team participants understand that the results will not be used to judge individuals or business units, but to identify improvements to benefit the agency.
- Understanding terminology. Before asking participants to select ratings for any element, be sure that they are clear on both the assessment elements and the definitions of the ratings.
- Varying ratings within an assessment element. That there may be variation within a given element can make it difficult to assign a single rating. For example, for a data value assessment that identifies "travel data" as a source, participants may state that data for vehicular travel has excellent availability but data for pedestrian travel has poor availability. In this situation, an overall rating should be assigned that reflects the relative importance of these two varieties of travel data for the business activities included. However, the comments area should be used to identify areas of specific weakness so that data gaps can be identified and addressed. For the data management assessment, there will likely be situations—particularly with the agency-wide assessment—where practices for some types of data are more mature than for others. Again, ratings should be assigned reflecting the predominant situation, but particular areas of weakness should be noted as gaps.

For Data Value Teams

At this workshop, the facilitator will lead members of the assessment team in completing the Data Value Assessment Tool. The recommended approach at this workshop follows.

Background: Why Are We Here?
- Describe why the agency is conducting the data self-assessment and how it plans to use the results. Identify the executive sponsor.
- Describe why this business area was selected.
- Summarize the schedule of meetings—for both the Assessment and (if available) Implementation and Monitoring phases.
- Provide an opportunity for questions.

Data Management Assessment Workshop Agenda	
9:00 AM	**Background** • Introductions • Assessment Purpose • Assessment Steps and Schedule
9:45 AM	**Assessment Content** • Element Overview • Maturity Level Descriptions • Rating Procedure
10:30 AM	**Strategy & Governance Ratings** • Sub-Element Descriptions • Individual Ratings • Consensus Ratings
11:15 AM	**Life Cycle Management Ratings** • Instructions • Individual Ratings • Consensus Ratings
12:00 PM	**Lunch**
1:00 PM	**Architecture and Integration Ratings** • Instructions • Individual Ratings • Consensus Ratings
1:30 PM	**Collaboration Ratings** • Instructions • Individual Ratings • Consensus Ratings
2:00 PM	**Data Quality Ratings** • Instructions • Individual Ratings • Consensus Ratings
2:30 PM	**Results Review** • Results and their Derivations • Discussion and Adjustment
3:15 PM	**Wrap-Up** • Feedback • Next Steps

Figure 14.   Data management assessment workshop sample agenda.

Activities and Data Types

- Describe how the activities for the business area were selected.
- Ask assessment team members to describe each activity; provide clarification if there are differences in how the different activities are understood.
- Describe how the data types were selected and clarify what they include.
- Ask for a show of hands from the assessment team members about who has used each type of data and feels qualified to rate its quality and usability.

Assessment Results and Definitions

- Talk about the result of the data value assessment—show the sample in Figure 7. Explain that this process will allow the agency to take a broad look at data needs across different business areas—integrating the perspectives of people who work with data on a daily basis as well as people making decisions based on data.
- Present the definitions of data availability, quality, and usability.
- Ask members of the assessment team to provide their own examples for each—to make sure that the group understands these concepts.

Importance Tab

- Describe the first activity: to identify the importance of each data type for each of the specific activities.
- Explain that this information will be used in calculating final ratings. Provide an example: let's say that both traffic data and pedestrian data are used for project scoping, and the quality of traffic data is very good but the quality of pedestrian data is low. If traffic data is rated as having "High Importance" and pedestrian data is rated as having "Low Importance," the overall data quality score for project scoping will be higher than if both of these data types were rated as having "High Importance."
- Discuss the different importance ratings:
  - High Importance—Essential, can't perform this activity without it
  - Medium Importance—Valuable, could do without it but it would affect value or credibility of results
  - Low Importance—Helpful, but could do without it
  - NA—Not helpful or relevant for this activity
- Discuss how there may be some data types that aren't being used for a given activity but still may be important for that activity. For example, it may be that information on mainte-nance activity costs is not available or not reliable, and therefore it is not used for budgeting. However, if high-quality cost information were available, it would be important for the bud-geting activity. Therefore, this type of data should be rated as being of Medium Importance—because budgeting is happening without it, but the credibility of budgeting results is suffering from the lack of good cost data.
- Ask each team member to individually complete ratings for the first activity. When they are finished, ask if they had difficulty assigning ratings. Provide clarification as needed to address their difficulties.
- Ask each team member to complete the remaining ratings on the **Importance** Tab. Ask them to provide a brief 1- to 2-sentence comment on why they selected the ratings they did.
- When everyone has completed the exercise, open a copy of the Data Value Assessment Tool, and project it on a screen.
- Select a member of the assessment team and ask them to state how they rated each of the data types for the first activity. Ask the group if anyone selected anything different. Discuss reasons for variation in each rating and enter the consensus rating for the group. If it is difficult to

achieve consensus in a reasonable amount of time, use either a "majority rules" approach or appoint one member of the team to have the final say.
- Continue through the other activities using the same process.
- Make sure to stay within the allotted time. The objective is not perfection, but to obtain the general sense of the group on the relative importance of each data source for the different activities.

Availability Tab

- Describe the second activity: to identify the availability of data for each of the activities. The purpose of this activity is to provide a general assessment of whether or not the agency has sufficient data to support the business area (i.e., answer the question: "Do we have the right data?").
- Explain that this activity is being done only for the activities (not for the data types) to highlight cases where there is an unmet need for a type of data that the agency doesn't collect—and therefore might not be included on the selected list of data types.
- Present the different availability ratings:
  - Poor—Little or no data to support this activity
  - Fair—Limited data and large gaps remain
  - Good—Basic data available, but some gaps remain
  - Excellent—Sufficient data is available to meet needs and there are no gaps
- Provide examples for each rating (tailored to the agency if possible):
  - Poor Availability: Developing a bike/pedestrian plan—with little or no information on pedestrian/bike facilities or current travel patterns.
  - Fair Availability: Conducting network screening to identify candidate locations for countermeasures—using crash data but very limited road inventory data.
  - Good Availability: Projecting future pavement condition—based on 5 years of trend information on pavement deterioration—but some gaps in understanding of how deterioration rates vary by pavement type and traffic level.
  - Excellent Availability: Developing a maintenance budget—based on regularly updated unit costs for labor, equipment, and materials.
- Ask each team member to individually complete the Availability Ratings. When they are finished, ask if they had difficulty assigning ratings. Provide clarification as needed to address their difficulties.
- Select a member of the assessment team and ask them to state how they rated the first activity. Ask the group if anyone selected anything different. Discuss reasons for variation in each rating and enter the consensus rating for the group. If it is difficult to achieve consensus in a reasonable amount of time, use either a "majority rules" approach or appoint one member of the team to have the final say.
- Continue through the other activities using the same process. For ratings other than "Excellent Availability," record a comment on the master assessment tool that documents the gaps in data availability.

Quality Tab

- Describe the third activity: to rate the quality of each data type. The purpose of this activity is to provide a general assessment of whether or not the agency's data is of sufficient quality to support the business area (i.e., answer the question: "Is our data good enough?").
- Define the three different dimensions of data quality:
  - Currency—the extent to which the data represents current conditions.
  - Accuracy—the degree to which the data represents actual conditions as they existed at the time of measurement.

- Completeness—the degree to which the data provides sufficient coverage and includes values for all required data elements. For example, a data set may be considered incomplete because it is missing coverage of some portion of the road network or some time periods or some classes of travelers.
- Stress that data quality should be rated relative to what the needs are. For example, data for planning purposes can be less accurate than data for design purposes. Data for traveler information needs to be more current (i.e., real time) than data for monthly or annual performance reporting.
- Explain that quality ratings are to be assigned for each data type—considering the most demanding needs across the identified activities.
- Present the different quality ratings:
  - Poor: Data is not current, accurate, or complete enough to be useful
  - Fair: Data is useful but lack of currency, accuracy, or completeness limits value
  - Good: Data quality is acceptable, but should be improved
  - Excellent: Data quality is sufficient for this activity—no improvements are needed
  - NA: Don't know—not enough information
- Ask each team member to individually complete the three sets of Quality Ratings for each data type, and provide a comment about why they assigned the ratings they did (i.e., what quality issues exist with each data source). When they are finished, ask if they had difficulty assigning ratings. Provide clarification as needed to address their difficulties.
- Select a member of the assessment team and ask them to state how they rated the first data type. Ask the group if anyone selected anything different. Discuss reasons for variation in each rating and enter the consensus rating for the group. If it is difficult to achieve consensus in a reasonable amount of time, use either a "majority rules" approach or appoint one member of the team to have the final say.
- Continue through the other activities using the same process. For ratings other than "Excellent," record a comment on the master assessment tool that documents the gaps in data quality.

Usability Tab

- Describe the final activity: to rate the usability of each data type. The purpose of this activity is to provide a general assessment of whether or not the agency's data is provided in a convenient form to support the business area.
- Provide examples of what makes data more (and less) usable:
  - Effort required to find the appropriate data (e.g., a web page or catalog)
  - Effort required to access the data (e.g., self-service versus special request)
  - Available tools and effort required to report and visualize data
  - Effort required to combine different data tables or data sets as needed (e.g., integrating separate data tables for rigid and flexible pavements or integrating data tables that use different spatial referencing methods)
  - Effort required to understand what the data means and how it was collected (e.g., availability of a data dictionary and metadata about the data set itself)
- Stress that data usability should be rated in the context of the identified business activities that involve use of data.
- Explain that ratings are to be assigned for each data type—considering the most demanding needs across the identified activities.
- Present the different usability ratings:
  - Poor: Data is available, but requires substantial effort to translate into usable form
  - Fair: Data is available, but requires moderate effort to translate into usable form
  - Good: Data is available in a usable form, but improvements to reporting capabilities would be helpful

- Excellent: Data is available in a usable form—no improvements are needed
- NA: Don't know—not enough information
- Ask each team member to individually complete the usability ratings, and provide a comment about why they assigned the ratings they did (i.e., what usability issues exist with each data source). When they are finished, ask if they had difficulty assigning ratings. Provide clarification as needed to address their difficulties.
- Select a member of the assessment team and ask them to state how they rated the first data type. Ask the group if anyone selected anything different. Discuss reasons for variation in each rating and enter the consensus rating for the group. If it is difficult to achieve consensus in a reasonable amount of time, use either a "majority rules" approach or appoint one member of the team to have the final say.
- Continue through the other activities using the same process. For ratings other than "Excellent," record a comment on the master assessment tool that documents the gaps in data usability.

Results Tab

- Discuss the results and describe how they are derived from the assessment exercise:
  - Results by Data Source—Importance: This is based on the highest importance rating assigned for the data source across all of the activities on the **Importance** tab.
  - Results by Data Source—Quality and Usability: These are taken directly from the responses on the **Quality** and **Usability** tabs.
  - Results by Business Activity—Availability: These are taken directly from the responses on the **Availability** tab.
  - Results by Business Activity—Quality: First, the overall quality rating is considered to be the limiting (lowest) quality rating among currency, accuracy, and completeness. Then, the overall quality ratings of each data source are weighted by activity-level importance ratings to determine activity-level quality ratings. This weighting is accomplished by first converting the quality and importance ratings from qualitative to quantitative values. Importance rating values are 0-NA, 1-Low, 2-Medium, 3-High. Quality ratings are 0-NA, 1-Poor, 2-Good, 3-Fair, 4-Excellent. Then, the quantitative quality ratings are multiplied by a weighting factor (the ratio of the data source importance score to the sum of the importance scores for all data sources). The resulting weighted numeric score is then converted back to a quality rating.
  - Results by Business Activity—Usability: the usability ratings from the Usability tab are weighted by importance using a method analogous to the one described for Quality.
  - Overall Business Area Results: The business activity results are aggregated to the overall business area level using two alternative methods. The "conservative" (or limiting activity) method takes the **lowest** rating across the different business activities. The "optimistic" method takes the **average** rating across the different business activities. Agencies can select one of these for inclusion in their summary presentation about the assessment results.
- Ask team members if the results appear to be an accurate reflection of data availability, quality, and usability. Provide an opportunity to go back and adjust the consensus ratings, making sure that any changes are adequately supported with reasons and documented in the comments blocks.
- Save the assessment results.

Wrap-up

- Thank the participants for their time and effort.
- Ask for feedback to be used for improving future assessment workshops.
- Remind participants of the timing and scope for the second workshop.

For Data Management Teams

The recommended approach for the data management assessment workshop is shown below.

Background: Why Are We Here?
- Describe why the agency is conducting the data self-assessment and how it plans to use the results. Identify the executive sponsor.
- Describe how this team was selected to participate.
- Provide an overview of the schedule of meetings—for both the Assessment and (if available) Implementation and Monitoring phases.
- Provide an opportunity for questions.

Assessment Content
- Describe each of the assessment elements—using the definitions in Table 2.
- Describe the maturity levels—using the definitions in Table 2.
- Talk about the result of the data management assessment—show the sample results in Figure 8. Indicate that this process will allow the agency to look at how different types of data are managed, and that this will lead to identifying what should be done to get more value from data investments.
- Describe the process that will be followed to complete the rating tool.

Data Strategy and Governance Tab
- Present an example of how to make a selection for each practice description for the first sub-element:
  - If the description does not reflect current practice in the agency, select 1-Totally Disagree
  - If the description partially reflects current practice in the agency, but is not the predominant way things are done, select 2-Somewhat Disagree
  - If the description is the predominant way things are done, but elements of it are not fully in place, select 3-Somewhat Agree
  - If the description accurately describes current agency practice, select 4-Totally Agree
  - In a case where there are substantially different practices in place within the agency or data program, the group can choose to qualify their rating in the comments section. For example "We have a rigorous data quality assurance program for data set X, but not for data sets Y and Z—our answers represent practice for data sets Y and Z only."
- Ask each member of the assessment team to complete the entries for each of the elements on the tab, and provide a comment about why they assigned the ratings they did. When they are finished, ask if they had difficulty assigning ratings. Provide clarification as needed to address their difficulties.
- Select a member of the assessment team and ask them to state how they rated the first sub-element type. Ask the group if anyone selected anything different. Discuss reasons for variation in each rating and enter the consensus rating for the group. If it is difficult to achieve consensus in a reasonable amount of time, use either a "majority rules" approach or appoint one member of the team to have the final say.
- Continue through the other activities using the same process. For maturity levels lower than "Functioning," record a comment on the master assessment tool that documents comments from the group about why the level was selected.

Data Life Cycle Management Tab
- Ask each member of the assessment team to complete the entries for each of the elements on the tab, and provide a comment about why they assigned the ratings they did. When they are finished, ask if they had difficulty assigning ratings. Provide clarification as needed to address their difficulties.

- Select a member of the assessment team and ask them to state how they rated the first sub-element type. Ask the group if anyone selected anything different. Discuss reasons for variation in each rating and enter the consensus rating for the group. If it is difficult to achieve consensus in a reasonable amount of time, use either a "majority rules" approach or appoint one member of the team to have the final say.
- Continue through the other activities using the same process. For maturity levels lower than "Functioning," record a comment on the master assessment tool that documents comments from the group about why the level was selected.

Data Architecture and Integration Tab

- Ask each member of the assessment team to complete the entries for each of the elements on the tab, and provide a comment about why they assigned the ratings they did. When they are finished, ask if they had difficulty assigning ratings. Provide clarification as needed to address their difficulties.
- Select a member of the assessment team and ask them to state how they rated the first sub-element type. Ask the group if anyone selected anything different. Discuss reasons for variation in each rating and enter the consensus rating for the group. If it is difficult to achieve consensus in a reasonable amount of time, use either a "majority rules" approach or appoint one member of the team to have the final say.
- Continue through the other activities using the same process. For maturity levels lower than "Functioning," record a comment on the master assessment tool that documents comments from the group about why the level was selected.

Data Collaboration Tab

- Ask each member of the assessment team to complete the entries for each of the elements on the tab, and provide a comment about why they assigned the ratings they did. When they are finished, ask if they had difficulty assigning ratings. Provide clarification as needed to address their difficulties.
- Select a member of the assessment team and ask them to state how they rated the first sub-element type. Ask the group if anyone selected anything different. Discuss reasons for variation in each rating and enter the consensus rating for the group. If it is difficult to achieve consensus in a reasonable amount of time, use either a "majority rules" approach or appoint one member of the team to have the final say.
- Continue through the other activities using the same process. For maturity levels lower than "Functioning," record a comment on the master assessment tool that documents comments from the group about why the level was selected.

Data Quality Tab

- Ask each member of the assessment team to complete the entries for each of the elements on the tab, and provide a comment about why they assigned the ratings they did. When they are finished, ask if they had difficulty assigning ratings. Provide clarification as needed to address their difficulties.
- Select a member of the assessment team and ask them to state how they rated the first sub-element type. Ask the group if anyone selected anything different. Discuss reasons for variation in each rating and enter the consensus rating for the group. If it is difficult to achieve consensus in a reasonable amount of time, use either a "majority rules" approach or appoint one member of the team to have the final say.
- Continue through the other activities using the same process. For maturity levels lower than "Functioning," record a comment on the master assessment tool that documents comments from the group about why the level was selected.

Results Tab

- Select the Results Tab and review the maturity levels assigned for each element and sub-element.
- Ask team members if the results appear to be an accurate reflection of agency practice. Provide an opportunity to go back and adjust the consensus ratings, making sure that any changes are adequately supported with reasons and documented in the comments blocks.
- Save the assessment results.

Wrap-up

- Thank the participants for their time and effort.
- Ask for feedback to be used for improving future assessment workshops.
- Remind participants of the timing and scope for the second workshop.

Step 3: Gaps and Candidate Actions Workshop Preparation Meeting

For Data Value Teams

In the Data Value Gaps and Candidate Actions Workshop Preparation Meeting, participants review the results of the assessment and prepare a list of gaps to be used as a starting point for the workshop. This ensures that there is continuity between the two workshops and enables the second workshop to proceed efficiently.

Worksheet 9 provides a format for listing the gaps. At the Preparation Meeting, use the comments from the Data Value Assessment Tool to partially complete this worksheet (leave the Business Impacts column blank). A sample is included to provide examples of the types of gaps that should be recorded.

Worksheet 9. Data value gaps.

ID	Type of Data	Gap Description	Business Impacts

Sample:

Type of Data	Gaps	Business Impacts
Maintenance Costs	Availability: We need historical cost information for maintenance budgeting, but we only have aggregate expenditures, not costs by activity.	Better data would improve our ability to link budget estimates with expected outputs. This would reduce the need for budget adjustments.
Freight	Availability: We would like to prioritize projects based on economic benefits, but lack good data on freight flows.	We are unable to understand and communicate the economic impacts of potential investments.
Right-of-Way	Availability: We need data on right-of-way limits, but this isn't systematically maintained.	Better data would reduce the time needed to research right-of-way limits where questions or problems arise.

Worksheet 9. (Continued).

Type of Data	Gaps	Business Impacts
Road Inventory	Availability: We need inventory data for local, non-state-maintained roads.	Better data is needed to meet federal reporting requirements and perform consistent safety analysis for both on- and off-system roads.
Underground Infrastructure	Availability: We only have underground infrastructure information where recent projects have been done.	Better data would reduce the need for special discovery efforts for project scoping.
Pavement Condition	Availability: We need trend data to estimate a pavement deterioration model, but only have 1 year of data.	Better data would allow us to produce more accurate and credible estimates of future pavement needs and set realistic performance targets.
Traffic	Availability: Traffic data for weekends and special events is very sparse.	There is a risk of over- or under-designing facilities based on faulty traffic assumptions.
Sign Inventory	Quality: Sign inventory is 3 years old and doesn't reflect recent work.	Districts won't use the inventory because they don't trust it is correct—they instead spend time re-collecting information in the field.
Roadside Assets	Quality: Roadside asset data has horizontal accuracy to the nearest 3 meters but sub-meter accuracy is needed.	Data can be used for planning but additional field collection will be required for project scoping and design.
Crashes	Quality: Reported crash locations don't match where crashes actually occurred.	We lack critical information needed to identify and correct safety issues. Considerable staff time is required to review each crash record and assign proper locations manually.
Traffic	Quality: We don't trust summarized traffic data because it doesn't adequately account for detector failures.	Our mobility performance measures lack credibility.
Pavement Condition	Quality: We don't trust that the pavement roughness data were measured accurately because equipment weren't calibrated properly.	We are not able to reliably track trends in pavement condition and understand how investments in pavement are affecting condition and performance.
Bridge Inspections	Quality: We think that that bridge condition data is biased because there wasn't sufficient independent verification of inspection results.	Priorities and needs for certain bridge projects may be overstated, resulting in suboptimal investment decisions.
As-Built Plans	Quality: We can't be sure that as-built plans are complete because there wasn't sufficient quality assurance.	We can't rely on as-builts as a data source for updating asset inventories and providing information on underground assets for project scoping. This means that we must pay to gather this information.
Incidents	Quality: Incident location information was entered in "free form text" so we can't use it for mapping.	Manual coding of locations is needed, which takes valuable staff time away from more productive activities.
Project Delivery	Quality: Project completion date wasn't defined in a consistent way—sometimes the physical completion date was used; other times the financial close-out date was used.	The data can't be used to compute statistics on project delivery performance and can't be used to help interpret historical crash, incident, and traffic data.
Traffic	Usability: We must submit a request to IT to get the traffic data we need.	Valuable IT resources are strained and business value of the data is diminished.
Traffic Signal Inventory	Usability: Our central traffic engineering unit maintains a signal inventory but districts weren't aware of that—so some started collecting their own data.	Effort and resources were wasted that could have been better used elsewhere.
Facility Inventory	Usability: We have a facility inspection database but no standard reports to summarize overall results.	Special staff effort is required to prepare one-off summaries of the data. The data is not used as much as they could be given the effort required for summarization and analysis.

Worksheet 10. Data management gaps.

ID	Assessment Element	Gap Description	Business Impacts

Sample:

Assessment Element	Gap Description	Business Impacts
Data Strategy and Governance	Lack of formal (role-based) definition of accountability and responsibility for data quality for each type of data	When existing key staff leave the agency or are re-assigned, there are risks that data quality will decline due to lack of formal accountability for specific job positions.
Data Life Cycle Management	No formal change notification process when data coding changes occur in one system that may affect another	Problem won't be noticed until managers request reports that rely on integration across systems. May result in delay in provision of information required for management and create need for emergency repairs to fix coding inconsistencies.
Data Architecture and Integration	Lack of standardization in data definitions — different districts maintain spreadsheets of information in various formats	Duplication of effort and lack of ability to aggregate data to produce information needed for management decisions.
Data Collaboration	Data is gathered from local jurisdictions ad hoc; formal data-sharing agreements do not exist	Current methods strain agency staff resources and do not reliably result in complete or current data.
Data Quality	No formal quality assurance process is in place	Lack of trust in the data; lack of ability for data managers to provide information on current level of quality—results in underused data and loss of potential value.

For Data Management Teams

In the Data Management Gaps and Candidate Actions Workshop Preparation Meeting, participants review the results of the assessment and prepare a list of gaps to be used as a starting point for the workshop. This ensures that there is continuity between the two workshops and enables the second workshop to proceed efficiently.

Worksheet 10 provides a format for listing the gaps. At the Preparation Meeting, use the comments from the Data Management Assessment tool to partially fill out this worksheet (leave the Business Impacts column blank.) A sample is included to provide examples of the types of gaps that should be recorded.

Step 4: Gaps and Candidate Actions Workshop

For both the data management and the data value assessments, a Gaps and Candidate Workshop is conducted—the second of two workshops held for each assessment team in Phase 2. Instructions for this workshop are the same for both assessment types. In this workshop, the

Gaps and Candidate Actions Workshop Agenda	
9:00 AM	**Background** • Review of Workshop Purpose and Agenda
9:15 AM	**Exercise 1: Validation of Data Gaps** • Review Process used to Develop Draft Gaps • Gap Screening and Validation • Business Impacts
10:15 AM	**Exercise 2: Candidate Actions** • Identify Current Initiatives that will Address Gaps • Recommend New Candidate Actions • Complete Action Evaluation Forms
12:00 PM	**Wrap-Up** • Feedback • Next Steps

Figure 15.   Gaps and Candidate Actions Workshop sample agenda.

assessment team is brought back together and asked to validate the draft list of data gaps and produce a list of candidate actions to address these gaps. The workshop consists of two exercises. A sample agenda for this workshop is provided in Figure 15.

Exercise 1: Gap Validation

The goal of the first exercise is to validate and complete the draft gaps produced at the Gaps and Candidate Actions Workshop Preparation Meeting—using Worksheet 10. The facilitator can use the following steps to review each gap:

- Read the draft gap aloud.
- Ask participants to confirm that this issue merits consideration in the data improvement action plan. This should screen out any gaps that the group thinks are not that important, as well as gaps for which realistic solutions aren't likely to be identified.
- If participants don't agree that the gap is significant, delete it from the list and move to the next gap.
- Ask participants if the gap can be described more precisely; re-word as needed.
- Ask participants to describe the business impacts of the gap: how is the gap creating risks, causing inefficiency, limiting value derived from available data, or affecting decision making? Why should agency management care about this gap?

After all draft gaps have been completed, the facilitator should ask the group if any gaps are missing and provide an opportunity to add entries to the list.

At the end of this first exercise, the group should have produced a complete set of gaps that the assessment team thinks are worth addressing in the data improvement action plan.

Exercise 2: Candidate Improvements

The second half of the workshop should be spent identifying candidate improvements that will be fed into the coordinated data improvement action planning process in Phase 3: Implement and Monitor. Worksheet 11 provides a format for recording the results of this exercise. Worksheet 12 provides a format for recording more detailed information about each candidate

Worksheet 11.   Candidate actions to address gaps.

ID	Action Description	Lead Responsibility	Current or New?

Worksheet 12.   Data improvement evaluation form.

Data Improvement Evaluation Form
Candidate Improvement:
Recommended By: Date:
Implementation
Proposed Lead for Development/Implementation: _____
Estimated Time for Development/Implementation:

❑ < 3 months
    ❑ 3-6 months
    ❑ 6-12 months
    ❑ > 12 months

Resource Requirements:

❑ Can be covered under existing project or initiative
    ❑ Requires new effort by In-house staff only
    ❑ Requires new effort involving external contractor support

Ongoing Maintenance & Support
Proposed Business Owner: _____
Other Staff Support Needs:_____
Resource Requirements:

❑ Can be maintained with existing staff
    ❑ Requires additional staff <= 1 FTE
    ❑ Requires additional staff > 1 FTE

Business Case (please describe for each applicable item)
External or Internal Agency Mandates:_____ _____
Staff Time Savings:_____
Other Agency Cost Reduction:_____
Risk Mitigation: _____
Improved Business Decisions:_____
Comments:

improvement that can be used to facilitate prioritization in Phase 3. A recommended process for completing these worksheets is as follows:

- Distribute copies of the data improvement idea checklist included in Appendix D. This is a master list of the types of data improvements that team members can consider to address the identified gaps.
- Identify Current Initiatives. Ask the assessment team members to list current agency initiatives—planned, funded, or in progress that are expected to address one or more of the identified gaps. Record these in Worksheet 11—identifying the lead business unit and manager of the effort and noting it as a "current" action.
- Ask each assessment team member to take 5 to 10 minutes and identify what they think would be the two most important actions to consider to close what they consider to be the remaining highest priority gaps. Participants should select actions that they think will provide the greatest business value, regardless of cost or level of effort. Ask each participant to describe their actions and identify the most likely lead business unit (and manager if possible). If a participant proposes an action similar to one already on the list, provide an opportunity to enhance the wording of the original action.
- When all participant actions have been recorded, ask the group if they think anything is missing from the list and add new entries as appropriate.
- Complete Improvement Evaluation Forms (see Worksheet 12) for each entry on Worksheet 11. Depending on the number of candidate improvements that the group has come up with, this can be done either as a group exercise, or through a "divide and conquer" strategy—with each team member completing forms for a portion of the candidate improvements.

Wrap-up

- Thank the participants for their time and effort.
- Ask for feedback to improve future workshops.
- Describe plans for the Implementation and Monitoring phase.

Step 5: Assessment Results Analysis and Summary

Following completion of all of the assessment workshops, the facilitators and staff supporting the assessment effort should get together and review Worksheets 9, 11, and 12 for data value teams, and Worksheets 10, 11, and 12 for data management teams. Staff should

- Edit the worksheets as needed to ensure that they are complete and consistent. This may require consultation with assessment team members to provide clarification or to fill in missing information.
- Prepare a presentation or briefing for the planning team to launch Phase 3 that includes
  - A description of each assessment team—including their topic and a list of their members
  - A data value assessment summary table (modeled after Figure 7) that summarizes ratings for data availability, quality, and usability for each of the business areas performing the data value assessment
  - A data management assessment summary table (modeled after Figure 8) that summarizes maturity levels for each of the groups performing the data management assessment
  - The consolidated list of gaps
  - Comments on themes common to the different groups
  - The consolidated list of current initiatives and candidate actions

CHAPTER 5

Phase 3: Improve and Monitor

Overview

The data self-assessment is envisioned as a cyclical process in which current practice is periodically evaluated, actions to improve are initiated, progress and results are tracked, and improvement plans are updated. Phase 3 of the data self-assessment involves monitoring the progress and results of actions previously recommended or implemented and initiating new improvement actions. This involves the following activities:

1. Consolidate Data Improvement Recommendations and Ongoing Initiatives
2. Prioritize New Improvements
3. Update Action Plan
4. Monitor Progress

Step 1: Consolidate Data Improvement Recommendations and Ongoing Initiatives

This step recognizes that agencies will likely not be conducting the data self-assessment process in a vacuum. Most transportation agencies are continually improving their data systems and programs (e.g., implementing system replacements and upgrades, new data collection programs, and data warehouses). In addition, various targeted assessment tools are available for transportation agencies that touch on data management and analysis practices within specific areas, including asset management, transportation operations, and safety. Agencies may have applied these tools and may be working through associated lists of recommendations.

Before considering a set of new candidate actions from the most recent set of data value and management assessments, it is helpful to compile a list of ongoing data initiatives, as well as data-related improvements that have been recommended through other assessments and/or agency business planning processes. This will ensure that the planning team comes into the process of prioritizing new actions with a good understanding of what is already underway.

Compiling a consolidated list of ongoing and recommended data improvements will likely require several days of staff effort. The level of effort can be managed by keeping the investigation at a relatively high level (i.e., focusing on major initiatives). Worksheet 13 can be used as a template for compiling this list. The following potential sources should be considered:

- The consolidated list of ongoing initiatives identified in Phase 2
- The consolidated list of recommended candidate improvements identified in Phase 2
- The agency's IT plan (where applicable)
- The agency's data business plan (where applicable)
- Recent business plans and budgets for units with data management-related responsibilities

Worksheet 13.   Data improvement tracking sheet.

Improvement Source	Improvement Description	Lead Responsibility	Current Status (Date Changed) - Recommended - Rejected - Planned - In Progress - Complete - Cancelled	Next Milestone and Date	Comments
Project Scoping Data Value Assessment	New QA Procedure for As-Builts	John Smith (Construction Division)	Recommended (05.25.2016)	Planning Team Consideration (07.01.2016)	See Assessment Team Summary Presentation

- Recent consultant studies or audits that have included reviews of data or information system capabilities
- Results of the following assessments (where applicable):
  - Transportation Asset Management Gap Analysis
  - NHTSA Traffic Records Assessment
  - FHWA Roadway Data Improvement Program (RDIP) and Crash Data Improvement (CDIP) Assessments
  - FHWA HPMS Assessment
  - AASHTO/FHWA Transportation Systems Management and Operations (TSMO) Assessment
  - FHWA Transportation Performance Management (TPM) Assessment

In subsequent cycles of the data self-assessment, the existing tracking sheet can serve as a starting point for this exercise. Staff can update this tracking sheet by checking with the designated lead person for each entry and asking key managers in the agency to identify new activities or sources of recommendations.

Step 2: Prioritize New Improvements and Assign Responsibilities

In this step, the planning team meets to review the tracking sheet and identify which new improvements should be advanced for implementation or more detailed investigation. The following process can be used for this meeting:

- **Key Gaps from the Assessment Phase.** The facilitator or champion reviews the most recent set of assessments and highlights gaps and business impacts identified by the assessment teams
- **Progress of Existing Initiatives.** Staff responsible for compiling the tracking sheet in Step 1 review the current initiatives underway—commenting on those that may address some of the concerns from the most recent assessments
- **Priorities.** The planning team identifies key priorities for action based on the following factors:
  - Relative areas of weakness (based on the assessment results)
  - Risks and opportunities (based on the "business impacts" columns of Worksheets 9 and 10)

- Current agency priorities and commitments
- Resource availability (e.g., staff, expertise, information technology, and funding)
- Likely effort to implement improvements relative to the likely benefits (based on Worksheet 12)

- **Targets.** As an option, the planning team establishes targets for data management maturity and data value ratings, based on the discussion of priorities. These targets would reflect where the agency wants to (or expects to) be in 2 years, based on an assessment of what is realistic to accomplish within a given timeframe. As noted under Priorities above, this assessment should be based on available budgets, personnel resources, technology capabilities, and the timing of related agency initiatives and priorities. Some agencies will find setting targets to be helpful to provide motivation and focus for advancement of data improvements. If desired, aspirational targets can be established to indicate the agency's desired eventual state. However, 2-year targets should be used to drive action—targets need to be realistically attainable to be of value.
- **Candidate Improvements.** The facilitator or champion reviews the candidate improvements identified by the assessment teams to address the gaps and highlights recommendations on the tracking sheet from related initiatives.
- **Synergies.** The planning team identifies opportunities for combining recommended improvements or folding them into already ongoing efforts to achieve efficiencies and synergies.
- **Triage.** Referring to the Data Improvement Evaluation forms (Worksheet 12), the planning team briefly discusses each candidate improvement and assigns it one of the following categories:
  - Definite Yes—compelling business case, addresses priorities and targets, resources in place to move forward
  - Definite No—lack of compelling business case, not feasible
  - Maybe—merits further consideration

  For each improvement in the "Definite Yes" category, the planning team assigns a lead person to move it forward. For each improvement in the "Definite No" category, the planning team provides a reason to record on the tracking sheet to provide feedback for members of the assessment teams. As with targets, triage ratings should take into account realistic expectations about budgets, personnel, technology capabilities, and the timing of other agency initiatives and priorities.
- **Next Steps.** For the improvements in the "Maybe" category, the planning team identifies the next steps to better assess the improvement. For example
  - Create more specific description of how this would be implemented
  - Gather information on other agency experience
  - Gather information on available products and services
  - Gather additional evidence to support the business case
  - Identify potential funding sources
  - Identify staff members who could lead this
  - Discuss with senior management to determine their level of support

  The planning team also assigns a person to take the lead on carrying these steps out.

The product of this meeting is an updated tracking sheet that does the following:

- Changes the status of the "Definite No" improvements to "Rejected"
- Changes the status of "Definite Yes" improvements to "Planned" and updates the Lead Responsibility for these
- Keeps the status of the "Maybe" improvements as "Recommended," adds the recommended next steps under Next Milestone, and updates the Lead Responsibility for those next steps.

Step 3: Create/Update Action Plan

Following the meeting of the planning team, an action plan for data improvements can be created (or updated, if one already exists). The action plan can be a formal document widely distributed across the agency, or it can be a simple document used primarily by the planning team to document the decisions made and track progress. The action plan should be used as a living document—regularly updated to accurately reflect the agency's current, ongoing, and planned set of activities related to data improvement. A "bare bones" action plan would compile descriptions of each data improvement initiative underway—using material from the Data Improvement Evaluation forms (see Worksheet 12) and the tracking sheet (see Worksheet 13).

Step 4: Monitor Progress

The purpose of this final step is to track the progress of improvement actions, act as necessary to keep them moving, and review results from completed improvements. This is best accomplished through quarterly meetings of the planning team. Prior to each meeting, the champion would work with staff to create an agenda that focuses on (1) improvements requiring decisions about next steps and (2) improvements that have been completed. Appropriate staff with lead responsibility would be invited to the quarterly planning team meeting to provide updates on these improvements. The agenda would include the following:

- Updates on recommended improvements with completed milestones
  - Briefing by lead staff
  - Planning team decision to move forward or request additional investigation
  - Planning team assignment of new lead (if appropriate)
- Updates on completed improvements
  - Briefing by lead staff on what was done
  - Discussion of whether or not the intended business benefits were achieved
  - Discussion of lessons learned for consideration in planning future improvements

Results of the first agenda item would be recorded on the tracking sheet. Completed improvements would be deleted from the tracking sheet. Entries for these would be created on the Data Improvement Log shown in Worksheet 14. This log is used to track accomplishments and lessons learned and can (1) be used to prepare periodic briefings for senior agency leadership and (2) serve as a valuable resource for the planning team when membership changes.

Worksheet 14.   Data Improvement Log.

Improvement Description	Lead Responsibility	Completion Date	Reported Benefits	Lessons Learned
New QA Procedure for As-Builts	John Smith (Construction Division)	12.15.2017	50% increase in usage of as-built plans	Involve district staff and contractors in defining the process; use shakeout period; provide feedback

Accommodating Decentralized Models of Improvement Planning

The process outlined in this portion of the Guide assumes that the planning team established in Phase 1 will be the focal point for moving forward with new improvements and tracking the progress of existing improvements. Some agencies may wish to pursue a more decentralized approach—in which improvements are prioritized, implemented, and tracked by existing management teams at the division or bureau levels. Agencies are encouraged to take advantage of existing structures and processes for reviewing, prioritizing, and tracking data-related improvements; however, having mechanisms to ensure coordination of and take advantage of synergies across improvement efforts is worthwhile. For example, each division- or bureau-level team could follow a similar approach for prioritizing improvements and delivering tracking sheets that are then consolidated and reviewed by the agency-wide planning team. The planning team could then meet to review the list and identify opportunities for collaboration or consolidation of efforts.

Glossary

This appendix provides a glossary of data management terminology used in this Guide and the accompanying tools. The following sources were consulted for definitions:

- AIIM—Association for Information and Image Management Glossary: http://www.aiim.org /community/wiki/view/glossary
- DAMA—the DAMA Dictionary of Data Management, 1st Edition, 2008
- IRMT—International Records Management Trust (IRMT) Glossary of Terms: http://www .irmt.org/documents/educ_training/term%20modules/IRMT%20TERM%20Glossary%20 of%20Terms.pdf
- ANSI/NISO Z39.19—Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies (2005) ISBN: 1-880124-65-3 is p. 157–167: http://www.niso .org/apps/group_public/download.php/12591/z39-19-2005r2010.pdf
- SAA—Society of American Archivists Glossary: http://www2.archivists.org/glossary
- OMB Circular A-130: http://www.whitehouse.gov/omb/circulars_a130_a130trans4/
- W3C—W3C Data Catalog Vocabulary—http://www.w3.org/TR/vocab-dcat/#class—dataset

Definitions derived from these sources are referenced accordingly. Where sources are not noted, definitions were developed from multiple sources. For many of these terms, definitions vary considerably across sources. These definitions are not intended to be authoritative beyond the scope of this Guide.

Business Rule. A formally stated constraint governing the characteristics or behavior of an object or the relationship between objects (entities) used to control the complexity of the activities of an enterprise. (Source: DAMA) *Example:* the width of an Interstate lane is 12 feet.

Change Management or Change Control. Processes in place to review, evaluate, and coordinate changes to data products, applications, and systems before they are implemented to minimize impacts to users and reduce any change–related errors. (Source: Adapted from DAMA)

Data. Representation of observations, concepts or instructions in a formalized manner suitable for communication, interpretation or processing by humans or computers. (Source: adapted from AIIM) *Examples:* a crash record; pavement roughness.

Data Accuracy. The degree to which data represents actual conditions as they existed at the time of measurement.

Data Architecture. A master set of data models and design approaches identifying the strategic data requirements and the components of data management solutions, usually at an enterprise level. (Source: DAMA)

Data Archiving. The process of moving data that is no longer actively used to a separate data storage device for long-term retention.

Data Business Plan.   A document that establishes data collection and management strategies that align with business objectives.

Data Catalog.   A listing of available data resources (e.g., data sets, query tools, maps, and reports) including descriptive information on what is included and how to access, compiled to facilitate discovery and understanding of available data.

Data Classifications.   A set of categories used to distinguish those key characteristics of a given data resource (e.g., level of sensitivity or degree of importance) used to determine appropriate governance policies.

Data Community of Interest.   The data owner, data steward, data users and other stakeholders with an active interest and role in the data program. (Source: Adapted from DAMA)

Data Completeness.   The degree to which the data provides sufficient coverage and includes values for all required data elements. For example, a data set may be considered incomplete because it is missing coverage of some portion of the road network, or some time periods, or some classes of travelers.

Data Currency.   The extent to which the data represents current conditions.

Data Customer.   A person or organization whose satisfaction with data products and services can determine the overall effectiveness and success of the programs. (Source: Adapted from DAMA)

Data Dictionary.   A place where a limited set of "data about the data" or meta data is stored. This may include technical meta data (e.g., column names and formats) and/or business meta data (e.g., data definitions, business rules, and code values). (Source: Adapted from DAMA)

Data Domain.   A category that can be used to group related data types in order to define stewardship roles. Examples of data domains include financial, human resources, and infrastructure.

Data Entities.   A classification of the types of objects found in the real world—persons, places, things, concepts, and events—of interest to the enterprise. (Source: DAMA)

Data Governance.   The accountability for the management of an organization's data assets to achieve its business purposes and compliance with any relevant legislation, regulation, and business practice.

Data Governance Body.   A high-level data governance structure in the organization that typically includes senior managers. Responsibilities may include identifying priorities for data governance policies, projects, or system enhancements, and the authorization, implementation, and enforcement of data governance policies and standards. (Source: Adapted from DAMA)

Data Inventory.   A compilation of information about an agency's data programs or major data categories that may include details on data types, storage locations, collection and update cycles, responsibilities, uses, and other information useful for data program management.

Data Timeliness.   The extent to which data is available within a useful time frame.

Data Management.   The processes and activities in place to develop, implement, and enforce policies and practices for protecting and enhancing the efficiency, value, and effectiveness of data and information. (Source: Adapted from DAMA)

Data Management Area.   This term can be used to provide an alternative, equivalent term for "Data Program" where use of the term "Data Program" would be perceived as a reference to a "computer program" as opposed to a programmatic function in an organization.

Data Mapping.   An activity to determine how data and associated information products (e.g., reports) are produced and used. Data mapping can be done to determine the value of the data for particular business processes or to identify data gaps or redundancies.

Data Owner(s).   People or groups with decision-making authority for initiating or discontinuing the data program and who determine the content of what data is collected.

Data Program.   An organizational function with significant data management responsibilities that can include scoping, collecting, managing, and/or delivering a particular category or form of data. Sometimes this function resides in a single organizational unit; at other times it is split across business units and IT units. Examples of DOT data programs include GIS, Road Inventory, HPMS, Traffic Monitoring, Crash Records, and Construction Project Data.

Data Quality.   The degree to which data is accurate, complete, timely, and consistent with requirements and business rules and relevant for a given use. (Source: Adapted from DAMA)

Data Quality Assurance.   Processes to ensure that data meets specified requirements.

Data Quality Control.   Processes to detect defects in collected data and take appropriate action.

Data Set.   A collection of data made available for access or download in one or more formats. (Source: adapted from W3C) Examples: a state's crash records for a single year; a database with roughness measures for pavement segments on the state highway system.

Data Steward(s).   People who are accountable for the quality, value, and appropriate use of the data.

Data Stewardship.   The formal, specifically assigned and entrusted accountability for business (as opposed to information technology) responsibilities ensuring effective control and use of data and information assets.

Data Visualization.   Techniques for graphical representation of trends, patterns, and other information. (Source: Adapted from DAMA)

Data Warehouse.   An integrated, centralized decision support data base and related software programs that can be used to collect, cleanse, transform, and store data from various sources to support business needs. (Source: Adapted from DAMA)

Enterprise Data Architecture.   An integrated collection of models and design approaches to align information, data, processes, projects, data systems/applications, and technology with the goals of the agency. (Source: Adapted from DAMA)

Findability.   The degree to which relevant information is easy to find when needed; findability is improved through application of meta data, taxonomies and other organizing tools, and search technologies. (Source: Adapted from AIIM)

Geospatial Data.   Data that includes location, specified with explicit geographic positioning information.

Information Management.   How an organization (e.g., a DT) efficiently plans, collects, creates, organizes, uses, controls, stores, disseminates, and disposes of data and information and ensures that the value of that data and information is understood and fully exploited.

Knowledge Management.   An umbrella term for various techniques for building, using and sustaining the knowledge and experience of an organization's employees.

Life Cycle.   The stages through which data or information passes, typically characterized as creation or collection, processing, dissemination, use, storage, and disposition. (Source: OMB Circular A-130)

Linear Referencing System (LRS).   A system for maintaining location information for events that occur along a linear network such as a road or rail line. It includes one or more methods for specifying the location of any point along the network based on distance from a known reference location (e.g., intersection-offset or county-relative milepoint).

Master Data.   Shared data about the core entities of an enterprise. In a private company, examples of core entities are customers, products, and vendors; in a DOT, examples of master data entities are routes, projects, funding sources, and district or regional offices.

Meta Data.   Data describing context, content, and structure of documents and records and the management of such documents and records through time. Literally, data about data. (Source: Adapted from AIIM/ISO 15489)

Records Management.   The systematic and administrative control of records throughout their life cycle to ensure efficiency and economy in their creation, use, handling, control, maintenance, and disposition. Similar to document management, but focused on documents that have been designated as official records with an emphasis on legal, regulatory, and risk management concerns. (Source: Adapted from SAA)

Sensitive Data.   Data that is confidential, privileged, or proprietary that should be protected from unauthorized disclosure, loss, misuse, or corruption to avoid serious consequences to the organization that owns it.

DOT Data Program Inventory

This appendix contains sample templates for compiling an inventory of DOT organizational functions that have data management responsibilities. Data management responsibilities may include data scoping, data collection or compilation, data validation, data quality improvement, data documentation, data loading, data integration, data analysis, data provisioning, data access management, and/or data user support.

The term "data program" (which can also be referred to as a "data management area" to avoid confusion with the term "computer program") is used here to refer to a logical unit of analysis for assessing capabilities and developing improvements. A data program can be a single organizational unit or team (e.g., a data management office), or an organizational function involving individuals from multiple organizational units to manage a particular category of data (e.g., financial data management).

Table B.1 provides an inventory listing that covers potential DOT data programs. Agencies can use this as a model for developing a master list of data programs for assessment and improvement. Figure B.1 is a sample template for compiling basic information about a specific data management area. These two templates are intended to be used together—once an inventory listing of data programs is developed, an inventory form can be completed for each data management area.

Table B.1.   DOT data program/data management area inventory list.

Data Category	Data Program/ Management Areas for Assessment	Sample Data Types Included	Responsible Unit(s)
General	IT Applications Development, Database Management and Administration	Multiple	
	Transportation Data Office	Multiple	
	Data Warehouse Group	Multiple	
	Business Intelligence/Dashboard /Reporting Group	Multiple	
	GIS Group	Geospatial Transportation Features (e.g., road centerlines, rail lines, and ferry routes), land and environmental features, multiple business data layers	
	Performance Management	Multiple performance measures—system condition, operations, agency efficiency	
Travel Data	Traffic Monitoring	AADT, Vehicle Classification, Turning Movements, Volume, Occupancy, Speed, Intersection Level of Service, Travel Time, WIM Data	
	Planning/Travel Modeling	Household Survey Data, Socioeconomic Data, Network Links and Nodes, Origin-Destination Matrices	
	Planning/Freight	Commodity flows, supply chain data, bottlenecks, infrastructure	
	Bicycle/Pedestrian Program	Bicycle Routes, Bicycle Paths, Bicycle and Pedestrian Counts	
System Inventory and Condition Data	Road Inventory	Mileage, Classification, Geometrics, etc.—including Model Minimum Inventory Elements (MIRE)	
	HPMS (typically combined with Road Inventory)	HPMS Data Elements—full extent and sample (e.g., road inventory, traffic, and pavement)	

Table B.1.   (Continued).

Data Category	Data Program/ Management Areas for Assessment	Sample Data Types Included	Responsible Unit(s)
	Pavement Management	Pavement inventory, IRI, cracking, summary condition, layer history	
	Bridge Management	Structure inventory and inspection	
	Traffic Engineering	Traffic signal inventory, guardrail inventory, sign inventory, railroad crossing inventory	
	ITS/Traffic Management Center	ITS device inventory, communications infrastructure inventory,	
Facilities Data	Property Management Fleet Management Maintenance Management	Plant and facilities inventory and condition, fleet inventory and utilization	
Financial/ Program Management Data	Capital Program/STIP	Federal Obligations, Construction Project Data, delivery performance (on-time, on-budget)	
	Financial Management	Funding and Allocations, Budgets and Expenditures	
	Contracts/ Procurement	Contracts, bid tabs	
	Human Resources	Personnel data	
Project Development Data	Design and Materials	Studies, surveys, non-destructive tests, core samples, design plans	
	Right-of-Way	Property inventory, transactions, appraisals, deeds	
	Environmental	Land use, water bodies, wetlands, groundwater, endangered species, historic sites, permits and commitments	
	Construction	Materials tests, inspections, payments, civil rights, claims, as-built plans	
System Operations Data	Incident Management	Incidents (real-time status, incident response time)	
	Traffic Management	Real-time traffic and travel time data	

Table B.1.   (Continued).

Data Category	Data Program/ Management Areas for Assessment	Sample Data Types Included	Responsible Unit(s)
	Equipment Management	Fleet/Equipment inventory, utilization, cost	
	Maintenance Management	Work requests, work orders, work accomplishments, resource utilization, cost	
	Road Weather Management	Weather/Road Condition (real time and historical)	
	Motor Carrier	Motor Carrier safety, operating statistics, IRP, IFTA, oversize/overweight permits	
	Modal Programs (e.g., transit and ferry)	Operations Statistics (e.g., vehicle miles, passenger miles, and revenues)	
Safety Data	Crash Records/FARS Reporting	FARS reports, police accident records, Crash location, Crash frequency	
	Safety Planning	Enforcement data (citations and convictions), injury surveillance, road safety audits, behavioral (e.g., seat belt and helmet compliance)	
Customer Relations	Public Affairs	Customer opinion surveys, website transactions	

Data Program/Management Area: Road Inventory

Primary Contact: Jane Doe, jane.doe@mystate.dot.us, 987-654-3210

Data Steward(s): Same as Primary Contact

Data Types Included: Functional Classification, NHS, Road Geometrics, Jurisdictional Responsibilities

Data Management Functions Performed:
- ✓ Data Requirements
- ❑ Data Architecture
- ✓ Data Collection/Acquisition
- ✓ Quality Assurance/Validation
- ✓ Data Transformation/Aggregation/Integration
- ❑ Data Analysis
- ✓ Data Distribution
- ✓ Data Storage and Backup
- ✓ User Access Privileges
- ✓ Data User Support
- ❑ Other_____

Federal Requirements Addressed by Program: HPMS, Certified Mileage Reporting

State Requirements Addressed by Program: NA

Internal Agency Users and Functions Served: Planning – HPMS Report, Long Range Planning, Corridor Planning; Pavement Management, Bridge Management, Maintenance Management, Traffic Monitoring, Project Development

External Users: FHWA, MPOs, Local Jurisdictions

Data Sources (by data type if multiple): Planning, FHWA, MPOs

Data Storage Location (by data type if multiple): (name and storage location of application or file(s)): Road Inventory Management System – Oracle database application on central servers managed in IT

Data Update Frequency (by data type if multiple): Annual

Figure B.1.   Sample data program/management area inventory form.

APPENDIX C

Data Management Assessment Elements

Element 1: Data Strategy and Governance

1.1	Strategy and Direction
	Leadership commitment and strategic planning to maximize value of data to meet agency goals
Level 1	*Agency-Wide:* Data collection and management is performed by individual business units with little or no agency-wide direction or coordination. Data improvements are not systematically or regularly identified—they are implemented reactively or opportunistically. *Program Specific:* Data improvements are not systematically or regularly identified—they are implemented on a reactive or opportunistic basis.
Level 2	*Agency-Wide:* Efforts to implement agency-wide data governance or assess agency-wide data needs are being discussed or planned. Data improvement needs are identified and communicated to management informally and efficiently. *Program Specific:* Data improvement needs are identified and communicated to management informally.
Level 3	*Agency-Wide:* Executive leadership has communicated the expectation that business units and IT functions should collaborate on identifying and implementing data improvements of agency-wide benefit. Data business plans or equivalent planning tools have been prepared to identify short and longer term data collection and management strategies that align with business objectives. Data improvement needs have been systematically reviewed, assessed, and documented. *Program Specific:* Data business plans or equivalent planning tools have been prepared to identify short and longer term data collection and management strategies that align with business objectives. Data improvement needs have been systematically reviewed, assessed and documented.
Level 4	*Agency-Wide:* Agency leadership regularly communicates and demonstrates active support for data improvements that will lead to improved agency effectiveness and efficiency. Agency leadership actively works to facilitate collaboration across business units on data improvements and maintain strong partnerships between IT and business unit managers. Data business plans or equivalent planning tools are regularly updated. A regular process of data needs assessment is in place and is used to drive budgeting decisions.

62

1.1	Strategy and Direction
	Program Specific: Data business plans or equivalent planning tools are regularly updated. A regular process of data needs assessment is in place and is used to drive budgeting decisions.
Level 5	*Agency-Wide and Program Specific* Data governance and planning activities are continually refined to focus on key risks and opportunities and eliminate activities without demonstrated pay off. Data governance and planning activities have a high probability of continuing through changes in executive leadership.

Element Description

This sub-element looks at the extent to which the agency leadership or data program manager has demonstrated a commitment to managing data as a strategic asset—through establishment of data governance structures, communications, and planning activities to ensure alignment between data investments and business needs.

Support for the AASHTO Data Principles

The following boxes show how the data management assessment elements tie to the AASHTO data principles. The first element is broad based and therefore covers all of the AASHTO data principles. Subsequent elements focus on specific aspects of data management and therefore are applicable to selected AASHTO data principles.

☑ Valuable ☑ Available ☑ Reliable ☑ Authorized ☑ Clear ☑ Efficient ☑ Accountable

Benefits of Moving Up the Maturity Scale

At lower levels of maturity, decisions about what data to collect and how to manage it are made in a highly decentralized fashion. As agencies mature, investments in data are made in a more deliberate and coordinated fashion. Agencies and data program managers can better answer questions such as "are we collecting the right data?" and "are we managing our data effectively?" Agencies can better identify where relatively unproductive or lower value data investments can be discontinued or diverted to higher value data investments.

Relevant Improvement Actions

❑ Data Governance Bodies

❑ Data Governance and Stewardship Policies

❑ Data Business Plans

Element 1: Data Strategy and Governance

1.2	Roles and Accountability
	Clear roles, accountability, and decision-making authority for data quality, value. and appropriate use
Level 1	*Agency-Wide and Program Specific:* Accountability for the quality, value, and appropriate use of data has not been clearly established.
Level 2	*Agency-Wide:* One or more individuals have been identified to lead agency-wide data governance activities. A business lead or point person has been designated for each major data set or application but the responsibilities of the role haven't been spelled out. *Program Specific:* A business lead or point person has been designated for each major data set or application but the responsibilities of their role haven't been spelled out.
Level 3	*Agency-Wide:* An agency-wide data governance body has been established with representation from IT and business functions and has defined its charter. Objectives and performance metrics for data governance and stewardship have been defined and documented. Role(s) have been designated to identify points of accountability for data quality, value, and appropriate use—for priority data programs or data subject categories. Decision-making authority has been defined for collection/acquisition of new data, discontinuation of current data collection, and significant changes to the content of existing data. Capabilities and skills for data management are included in staff position descriptions, agency recruiting, and staff development efforts. *Program Specific:* Role(s) have been designated to identify points of accountability for data quality, value, and appropriate use—for priority data programs or data subject categories. Decision-making authority has been defined for collection/acquisition of new data, discontinuation of current data collection, and significant changes to the content of existing data. Capabilities and skills for data management are included in staff position descriptions, agency recruiting, and staff development efforts.
Level 4	*Agency-Wide:* An agency-wide data governance body is active and achieving results recognized as valuable. The agency is successfully identifying and resolving situations where individual business unit interests are in conflict with agency-wide interests related to data collection and management Staff with responsibility for data stewardship and management have sufficient time and training to carry out these responsibilities. Staff with responsibility for data stewardship and management play an active role in defining data improvements and periodically produce reports of progress to their managers. *Program Specific:* Staff with responsibility for data stewardship and management have sufficient time and training to carry out these responsibilities. Staff with responsibility for data stewardship and management play an active role in defining data improvements and periodically produce reports of progress to their managers.

1.2	Roles and Accountability
Level 5	*Agency-Wide:* A charter for agency-wide data governance body is reviewed periodically and updated based on experience. Stewardship roles are periodically reviewed and refined to reflect new or changing data requirements and implementation of new data systems. Staff with responsibility for data stewardship and management are coordinating with their peers in the agency and with external data partners to deliver best value for resources invested. Data management-related metrics are routinely considered in employee performance reviews. *Program Specific:* Stewardship roles are periodically reviewed and refined to reflect new or changing data requirements and implementation of new data systems. Staff with responsibility for data stewardship and management are coordinating with their peers in the agency and with external data partners to deliver best value for resources invested. Data management-related metrics are routinely considered in employee performance reviews.

Element Description

This sub-element assesses the extent to which roles and accountability for data stewardship have been agreed-upon, defined, documented, and assigned to individuals.

Support for the AASHTO Data Principles

☑ Valuable ☑ Available ☑ Reliable ☑ Authorized ☑ Clear ☑ Efficient ☑ Accountable

Benefits of Moving Up the Maturity Scale

At lower levels of maturity, there is a lack of clarity about who "owns" data and who is accountable for making sure that data meets business needs. Responsibilities have not yet been defined for ensuring that different business units coordinate on data collection and data management activities to maximize efficiencies. As agencies move up the maturity scale, roles and responsibilities are more formalized. Managers ensure that staff are assigned to data stewardship and management roles and are sufficiently trained and provided with resources. Formalizing and documenting roles and accountability for data (1) creates a consistent and sustainable framework for proper data management, (2) reduces the agency's dependence on "heroic efforts" to take care of what needs to be done, and (3) helps to ensure that staff are proactive about providing the right data efficiently, with the right quality, and in the right form.

Relevant Improvement Actions

❑ Data Governance Bodies

❑ Data Governance and Stewardship Policies

❑ Data Business Plans

❑ Data Management Roles & Responsibilities

Element 1: Data Strategy and Governance

1.3	Policies and Procedures
	Adoption of principles, policies, and business processes for managing data as a strategic agency asset
Level 1	No formal policies and procedures have been defined.
Level 2	Executive leadership has endorsed basic data principles.
Level 3	The scope of agency-wide data governance has been established. Data classifications have been defined based on agency-wide importance or need for cross-business unit integration. A limited set of data management policies have been adopted for priority data categories. The agency has a documented procedure and decision-making process for requesting and evaluating new data collection or acquisition requests.
Level 4	A comprehensive set of data management policies has been adopted based on collaboration across the agency, including IT, business units, and records management. Processes are in place to monitor and enforce compliance with policies. The agency has a documented and implemented procedure for requesting and evaluating new data collection or acquisition requests (i.e., the documented procedure is routinely followed).
Level 5	Policies are regularly reviewed and updated based on factors such as awareness/reach within the agency, effectiveness, and cost burden.

Element Description

This sub-element looks at the extent to which the agency has established clear policies and procedures about how data is to be managed as a corporate asset.

Support for the AASHTO Data Principles

☑ Valuable ☑ Available ☑ Reliable ☑ Authorized ☑ Clear ☑ Efficient ☑ Accountable

Benefits of Moving Up the Maturity Scale

At lower levels of maturity, there are no written and adopted policies and procedures related to data governance and management. As agencies move up the maturity scale, policies and procedures are drafted, adopted, and implemented throughout the agency. The policies and procedures help standardize how an agency treats data. If implemented well, policies and procedures result in higher quality data, more effective use of data, and clear decision-making processes around data.

Relevant Improvement Actions

❑ Data Governance and Stewardship Policies

Element 1: Data Strategy and Governance

1.4	Data Asset Inventory and Value
	Tracking of agency data assets and their value added
Level 1	*Agency-Wide:* There are no inventories of available data sets. There is limited awareness of how data sets are used and what value is being provided. *Program Specific:* There is limited awareness of how data sets are used and what value is being provided.
Level 2	*Agency-Wide:* Some business units maintain lists of available data sets but there is no consistent, agency-wide data inventory. There is general awareness of how different data sets are used and what value is being provided, but no records are kept on this. *Program Specific:* There is general awareness of how different data sets are used and what value is being provided, but no records are kept on this.
Level 3	*Agency-Wide:* Data sets of agency-wide importance have been identified and documented with basic elements, including business contact, technical contact, location, and description. Primary users and uses of each data set have been identified and documented Data collection or acquisition costs are tracked. *Program Specific:* Primary users and uses of each data set have been identified and documented. Data collection or acquisition costs are tracked.
Level 4	*Agency-Wide:* A consistent agency-wide inventory of data sets is maintained and kept current as new data sets come on line. Data inventory information is used to identify duplicative data sets that can be eliminated or consolidated. Managers use information about data storage and management costs to evaluate opportunities for improved efficiencies. *Program Specific:* Managers use information about data storage and management costs to evaluate opportunities for improved efficiencies.
Level 5	*Agency-Wide and Program Specific:* There is a good understanding of the value provided by each data set with respect to agency efficiency and effectiveness. Data collection and management methods are regularly evaluated and improved.

Element Description

This sub-element looks at the extent to which the agency or data program manager has documented the data, its uses, and its value to the agency.

Support for the AASHTO Data Principles

☑ Valuable ☑ Available ☑ Reliable ☑ Authorized ☑ Clear ☑ Efficient ☑ Accountable

Benefits of Moving Up the Maturity Scale

At lower levels of maturity, information about data and its uses resides in the heads of a few staff members—nothing is written down. As agencies move up the maturity scale, they consistently document their data sets and track how data sets are used. This provides the basis for articulating the value of different types of data to the agency and weighing data collection and maintenance costs against value added. It also enables agencies to identify areas of duplication and opportunities for consolidation.

Relevant Improvement Actions

❑ Data Catalogs and Dictionaries

❑ Data Value Mapping

Element 1: Data Strategy and Governance

1.5	Relationships with Data Customers
	Connections between data producers and users
Level 1	There are no proactive outreach activities to understand data user needs.
Level 2	Informal, limited outreach to other business units has been conducted to identify how they might use available data sets.
Level 3	Meetings have been held with current or potential new users for our data to understand their needs. This information has been taken into account in developing plans for improvements.
Level 4	Input from data customers is routinely solicited, collected, and considered through various online and in person forums (e.g., Communities of Interest).
Level 5	There are formal, written agreements that document what data will be provided to customers, when, and how. A process is in place to periodically re-negotiate these agreements.

Element Description

This sub-element looks at the extent to which data program managers have established channels of communication with data users.

Support for the AASHTO Data Principles

☑ Valuable ☑ Available ☑ Reliable ☑ Authorized ☑ Clear ☑ Efficient ☑ Accountable

Benefits of Moving Up the Maturity Scale

At lower levels of maturity, data program managers do not actively communicate with data users to understand how they use data and obtain feedback on data quality. As agencies move up the maturity scale, data program managers reach out to data users and act on feedback received to make improvements. In some situations, service level agreements can be negotiated to formalize what data is provided, at what frequency, and in what form. Strengthening relationships between data providers and data customers helps agencies avoid situations in which data is being produced but not used as intended. A functioning feedback loop between data providers and customers helps data providers focus on data improvements that add value.

Relevant Improvement Actions

❑ Data Communities of Interest

Element 1: Data Strategy and Governance

1.6	Data Management Sustainability
	Continuity of data management knowledge and expertise through staff transitions
Level 1	*Agency-Wide and Program Specific:* Risks and needs associated with data management knowledge and core competencies are not well understood.
Level 2	*Agency-Wide and Program Specific:* There is some understanding of risks associated with retirement of key individuals with specialized knowledge of data systems—but these risks have not been systematically identified.
Level 3	*Agency-Wide and Program Specific:* Risks associated with potential loss of key individuals with specialized knowledge of data systems have been systematically identified. Strategies have been developed to mitigate these risks. Core competencies for data management are included in staff position descriptions, agency recruiting, and staff development efforts.
Level 4	*Agency-Wide and Program Specific:* There is a standard process in place to ensure continuity in data management practices through staff transitions. Staffing requirements for data management activities are understood and planned for. Processes are in place to ensure that work commitments are in line with available staff resources.
Level 5	*Agency-Wide and Program Specific:* People with specialized knowledge about agency data sets have been identified and there are succession plans and mentoring strategies in place to pass on this knowledge to others. There is a functioning process to bring on new skills and capabilities as needed to meet changing technologies and data management methods.

Element Description

This sub-element assesses the extent to which the agency can sustain data management functions through staff transitions.

Support for the AASHTO Data Principles

☑ Valuable ☑ Available ☑ Reliable ☑ Authorized ☑ Clear ☑ Efficient ☑ Accountable

Benefits of Moving Up the Maturity Scale

At lower levels of maturity, the agency is not aware of risks associated with departures of staff with specialized knowledge and skills related to particular data sets or data management practices. As agencies move up the maturity scale, these risks are systematically identified and mitigation actions are in place—including succession plans and mentoring strategies. A proactive approach to ensuring data management sustainability reduces risks of disruption to data access or reporting activities. It also provides for an orderly and efficient transition of responsibilities.

Relevant Improvement Actions

❑ Succession Planning and Management

❑ Core Competency Definition

Element 2: Data Life Cycle Management

2.1	Data Updating
	Well-defined and coordinated data update cycles
Level 1	*Agency-Wide and Program Specific:* Data updating cycles and business rules for data updates have not been defined.
Level 2	*Agency-Wide and Program Specific:* Updating cycles have been established but have not been documented.
Level 3	*Agency-Wide and Program Specific:* Updating cycles have been documented. Business rules have been defined for how key data entities are added, updated, and deleted.
Level 4	*Agency-Wide and Program Specific:* Updating cycles are being consistently followed. Business rules for data updating are embedded in and enforced by applications (where applicable).
Level 5	*Agency-Wide and Program Specific:* Data updating methods are periodically reviewed to identify opportunities for improved efficiencies.

Element Description

This sub-element assesses the extent to which update methods and cycles have been defined and documented for key data sets.

Support for the AASHTO Data Principles

❑ Valuable ❑ Available ☑ Reliable ❑ Authorized ❑ Clear ❑ Efficient ❑ Accountable

Benefits of Moving Up the Maturity Scale

At lower levels of maturity, data updates are made ad hoc and users are not aware of data updating frequencies or methods. In addition, rules for adding and deleting key data entities (e.g., routes, projects, and employees) have not been developed. As agencies move up the maturity scale, they create and maintain business rules for how each major data set is to be updated. Where applicable, business rules are embedded into applications. For example, an HR system may include a wizard for adding a new employee that makes sure that all required data elements are entered. Defining rules for data updates is a critical step that affects the cost of data maintenance and also the level of quality that will be provided. Formalizing rules for data updates provides clarity for both data users and data managers.

Relevant Improvement Actions

❑ Standard Operating Procedures

Element 2: Data Life Cycle Management

2.2	Data Access Control
	Well-defined policies and guidelines for managing access to data sets
Level 1	*Agency-Wide and Program Specific:* There are no established policies for determining if access to data sets should be limited.
Level 2	*Agency-Wide and Program Specific:* A process of defining what data is sensitive and needs to be protected is underway. A process of defining what data can be shared outside the agency is underway.
Level 3	*Agency-Wide and Program Specific:* Standard guidelines are available for identifying and protecting sensitive data. Criteria and processes have been defined for making data available to the public.
Level 4	*Agency-Wide and Program Specific:* All core data sets have been classified based on guidelines for data protection and access. Data owners/providers are complying with guidelines for data protection and access.
Level 5	*Agency-Wide and Program Specific:* Data access guidelines and procedures are well established and periodically reviewed and updated.

Element Description

This sub-element assesses the extent to which the agency manages access to data sets to protect sensitive information and maintain data integrity.

Support for the AASHTO Data Principles

❑ Valuable ❑ Available ☑ Reliable ☑ Authorized ❑ Clear ❑ Efficient ☑ Accountable

Benefits of Moving Up the Maturity Scale

At lower levels of maturity, the agency's approach to managing access to data is ad hoc. As agencies move up the maturity scale, there is a standard method for classifying sensitive information and a formal process for defining access privileges as new data sets are brought on line. Standardizing and formalizing data access control supports compliance with applicable information security regulations and prevents data corruption due to unauthorized or unmanaged changes. It also enables agencies to define and apply consistent criteria for what data is to be shared openly versus kept internal to the agency.

Relevant Improvement Actions

❑ Data Access Policies

Element 2: Data Life Cycle Management

2.3	Data Findability and Documentation
	Availability of data catalogs and dictionaries that enable discovery and understanding of available agency data assets
Level 1	*Agency-Wide and Program Specific:* Users rely on "word of mouth" to discover what data is available. There are variations across data sets in terms of the level and type of available documentation.
Level 2	*Agency-Wide and Program Specific:* Efforts are under way to improve data findability and documentation through adoption of common meta data standards, development of data set catalogs, or creation of web pages with links to commonly requested data sets.
Level 3	*Agency-Wide:* An agency-wide data catalog or meta data repository has been established to improve data findability for business users and is being populated. Standards and policies have been defined to ensure that a data dictionary is available for each data set. Templates for describing data collection, updating, and reporting processes have been developed and are starting to be used. *Program Specific:* Standards and policies have been defined to ensure that a data dictionary is available for each data set. Templates for describing data collection, updating, and reporting processes have been developed and are starting to be used.
Level 4	*Agency-Wide:* Business users can access a list of available agency data sets to discover data of potential value to meet their needs. Consistent documentation is available describing data collection, updating, and reporting cycles for most of the agency's core data sets. Data dictionary information is available and current. Quality assurance processes are in place to ensure that data dictionary information is complete and useful. Processes are in place to keep the data set listing (or catalog) current when data sets are added or discontinued. *Program Specific:* Data dictionary information is available and current. Quality assurance processes are in place to ensure that data dictionary information is complete and useful. Processes are in place to keep the data set listing (or catalog) current when data sets are added or discontinued.
Level 5	*Agency-Wide:* Business users can search for availability of agency data on a particular subject or entity type. The agency periodically evaluates opportunities to refine its approach to data documentation based on user needs and new technologies. Documentation of data sets is periodically improved based on feedback from users and research into best practices. *Program Specific:* Documentation of data sets is periodically improved based on feedback from users and research into best practices.

Element Description

This sub-element assesses the extent to which the agency ensures that potential data users can discover what data is available and understand the potential applicability of a data set for a given business need.

Support for the AASHTO Data Principles

☑ Valuable ☑ Available ❑ Reliable ❑ Authorized ☑ Clear ☑ Efficient ❑ Accountable

Benefits of Moving Up the Maturity Scale

At lower levels of maturity, data sets are discovered primarily by word of mouth. As agencies move up the maturity scale, standard information is maintained and made available about what each data set contains, including the meaning of each data element. Providing an easily accessible catalog of data sets (or sources) adds value to existing data by promoting its re-use and minimizes the chances that duplicate data will be collected. Documenting the source and derivation of data elements also reduces risks associated with data misuse.

Relevant Improvement Actions

❑ Data Catalogs and Dictionaries

❑ Data Management Plans

❑ Data Curation Profiles

Element 2: Data Life Cycle Management

2.4	Data Backups and Archiving
	Guidelines and procedures for protection and long-term preservation of data assets
Level 1	*Agency-Wide:* There may be important data sets managed using desktop applications within individual business units, but these have not been systematically identified. Each business unit is responsible for ensuring that its data sets are backed up and periodically archived to enable future retrieval and use. *Program Specific:* Backups of data sets are made ad hoc.
Level 2	*Agency-Wide:* Several of the agency's important data sets are managed using desktop applications (e.g., spreadsheets) but plans are in process to bring these into enterprise databases. Data owners receive informal (unwritten) guidance regarding frequency and storage locations for backups and archive copies. *Program Specific:* Backups of data sets are made regularly, but there are no written procedures on backup frequency or storage locations. Archive copies of data sets exist, but there are no written procedures on how to create these and how to retrieve them.

2.4	Data Backups and Archiving
Level 3	*Agency-Wide:* Most of the agency's important data sets are managed within enterprise databases (e.g., Oracle, SQLServer) and regular backups are made. Written procedures on backup frequency and storage locations are available. Written procedures on data archiving and retrieval are available. *Program Specific:* Written procedures on backup frequency and storage locations are available. Written procedures on data archiving and retrieval are available.
Level 4	*Agency-Wide:* All of the agency's important data sets are managed within enterprise databases (e.g., Oracle, SQLServer) and regular backups are made. Backup procedures are consistently followed. Archiving procedures are consistently followed. Backup procedures have been fully tested. Archiving procedures have been fully tested. *Program Specific:* Backup procedures are consistently followed. Archiving procedures are consistently followed. Backup procedures have been fully tested. Archiving procedures have been fully tested.
Level 5	*Agency-Wide and Program Specific:* Data managers and stewards periodically review existing data backup and archiving procedures and update them as appropriate to reflect user feedback or changing needs.

Element Description

This sub-element assesses the extent to which active data sets are backed up and inactive data sets are archived for future use as needed.

Support for the AASHTO Data Principles

☑ Valuable ☑ Available ☑ Reliable ❑ Authorized ❑ Clear ❑ Efficient ❑ Accountable

Benefits of Moving Up the Maturity Scale

At lower levels of maturity, backups and archiving are performed ad hoc. As agencies move up the maturity scale, the agency has developed and reliably follows guidance and procedures that specify what types of data will be centrally managed (e.g., stored in enterprise databases), how frequently backups will occur, where backups will be stored, and who is responsible for making and testing backups. In addition, there will be a well-defined process for identifying which inactive or historical data sets should be archived and what type of business user access to the archived information should be provided to meet business needs. Formalizing backup processes and verifying that they are working reduces the risk of data loss due to hardware failures and other sources of data corruption. Formalizing archiving processes (1) enables agencies to identify where data sets can be retired to reduce data maintenance costs and (2) ensures that business user needs are considered when determining appropriate archive methods.

Relevant Improvement Actions

❑ Data Governance and Stewardship Policies

❑ Data Retention Schedules and Archiving

Element 2: Data Life Cycle Management

2.5	Data Change Management
	Processes to minimize unanticipated downstream effects of data changes.
Level 1	*Agency-Wide and Program Specific:* There are no defined processes for analyzing how changes to the data structure in one system may affect reports or other dependent systems.
Level 2	*Agency-Wide and Program Specific:* A standard data change management process is being developed.
Level 3	*Agency-Wide and Program Specific:* A standard change management process has been defined for changes to data elements that may affect multiple systems. This involves consultation and communication with affected data owners and users and propagation of the changes across databases as needed. Change analysis and propagation processes are mostly manual.
Level 4	*Agency-Wide:* A meta data repository or other tool is available to conduct change impact analysis (i.e., identification of which systems and database tables contain a particular data element). Automated processes are in place to manage changes to code lists and propagate these changes to various business systems. Automated processes are in place to manage additions, deletions, and changes to master data entities. A change management process is in place and functioning as intended. *Program Specific:* A change management process is in place and functioning as intended.
Level 5	*Agency-Wide and Program Specific:* A periodic review is conducted of the nature and extent of data changes to improve future data architecture and change management practices.

Element Description

This sub-element assesses the extent to which procedures are in place to manage the process of making changes to data structures in one data set or system that may affect other systems or reports.

Support for the AASHTO Data Principles

❑ Valuable ❑ Available ☑ Reliable ❑ Authorized ❑ Clear ☑ Efficient ❑ Accountable

Benefits of Moving Up the Maturity Scale

At lower levels of maturity, changes to data structures, definitions, or unique identifiers are made as needed—without awareness of potential unintended consequences. Effects may be

discovered only when downstream applications or reports stop working as a result of the changes made. As agencies move up the maturity scale, an active approach is in place for anticipating downstream effects of changes, communicating with the data stewards of these downstream systems, and implementing changes in a controlled, automated, and coordinated manner. Such an approach focuses on management of "master data" that exists across multiple agency systems. Putting active and robust change management processes in place helps to prevent (1) business disruptions from broken reports or queries and (2) introduction of inconsistencies in data structures and definitions across systems that hinder creating an integrated view of data.

Relevant Improvement Actions

❑ Data Change Management

❑ Reference Data Management

Element 2: Data Life Cycle Management

2.6	Data Delivery
	Delivery of data to users in various convenient, useful, and usable forms
Level 1	*Agency-Wide and Program Specific:* Data reporting is accomplished in a decentralized fashion—individual data or application owners separately plan and implement reporting capabilities. *Program Specific:* Limited standard reports are available, but there are no ad hoc query capabilities. New reports require software development resources to implement.
Level 2	*Agency-Wide:* The agency is exploring agency-wide needs and opportunities for improving access to integrated agency data in usable forms. Pilot initiatives may be underway. Ad hoc query tools are available but geared to a few individuals with specialized training. *Program Specific:* Ad hoc query tools are available but geared to a few individuals with specialized training.
Level 3	*Agency-Wide:* The agency has implemented enterprise solutions for data access, reporting, visualization, and analysis (e.g., data warehouse, data marts, and dashboards). Agency employees have access to a common map-based interface that allows them to view and analyze various information (e.g., pavement condition, bridge condition, crashes, traffic counts, programmed projects, and completed projects). The agency provides access to both "live" data and "frozen" or "snapshot" data, depending on an assessment of business needs. The agency has developed a standard approach to accessing agency data from mobile devices. *Program Specific:* Reporting and query tools are available for general use within the agency and do not require specialized training. Business needs for access to both live data and frozen/snapshot data have been identified.

2.6	Data Delivery
Level 4	*Agency-Wide:* The agency has the expertise and tools in house to develop data marts that allow employees to "slice and dice" data sets, perform ad hoc queries, and produce reports at the desired level of summarization. The agency has the expertise and tools in house to combine data sets based on different road sections (e.g., 10th mile sections for pavement information and 2-mile sections for AADT). Agency employees can easily visualize trend information for asset condition, capital and maintenance expenditures, traffic, crash rates, and other important agency performance indicators. Agency field staff can access information about assets, projects, or work orders on mobile devices. The agency has sufficient network connectivity and bandwidth to enable remote data access from field offices. *Program Specific:* Data is made available through various formats and platforms (e.g., GIS portal, mobile devices, and dashboards) to meet identified business requirements.
Level 5	*Agency-Wide:* The agency has implemented a flexible architecture for reporting and mapping that enables easy addition of new data sources and enhanced analysis capabilities in response to newly identified requirements. The agency routinely improves data access and usability based on feedback from users and monitoring of the latest technology developments. *Program Specific:* Data is shared outside of the agency via a statewide or national GIS portal or clearinghouse Access to data is provided through a service or application programming interface (API).

Element Description

This sub-element assesses the extent to which data is delivered to end users in convenient forms suited to best meet business needs.

Support for the AASHTO Data Principles

☑ Valuable ☑ Available ❑ Reliable ❑ Authorized ❑ Clear ☑ Efficient ☑ Accountable

Benefits of Moving Up the Maturity Scale

At lower levels of maturity, data is collected or acquired without careful consideration of the wide range of potential uses and the types of delivery formats that would best serve these uses. As agencies move up the maturity scale, the agency implements tools and processes to ensure that data is delivered in usable forms. This may involve various techniques including data integration and transformation (e.g., to combine traffic and pavement data or to aggregate financial transactions into meaningful categories), development of exception reports, use of GIS portals and business intelligence platforms, and creation of open data feeds. Emphasizing data delivery promotes data use and re-use, thereby producing more value from data investments, and increases agency efficiency by reducing the need for time-consuming data manipulation and custom report development.

Relevant Improvement Actions

❑ Data Delivery Platforms

Element 3: Data Architecture and Integration

3.1	Location Referencing
	Common location referencing methods across agency data sets
Level 1	*Agency-Wide:* The agency does not have a single common LRS. Data sets, including location elements, cannot be spatially integrated with other agency data sets. *Program Specific:* Data sets, including location elements, cannot be spatially integrated with other agency data sets.
Level 2	*Agency-Wide:* The agency is working toward establishing a single common LRS. Representation of location information is being standardized. *Program Specific:* Representation of location information is being standardized.
Level 3	*Agency-Wide:* The agency has developed a single common LRS. Quality standards for the LRS have been established with input from various business units. The agency has defined a process for propagating changes in the LRS to various agency data sets. New data sets that include location elements are collected using the agency's LRS. *Program Specific:* New data sets that include location elements are collected using the agency's common LRS.
Level 4	*Agency-Wide:* The agency's LRS is used for all agency data sets that include location. The agency's LRS meets established quality standards. Methods are in place and functioning to propagate changes in location referencing resulting from road network changes to business data sets. Methods are in place and functioning to translate between coordinate-based location referencing (e.g., latitude/longitude) and linear referencing (e.g., route-milepoint). *Program Specific:* Methods are in place and functioning to propagate changes in location referencing resulting from road network changes to business data sets. Methods are in place and functioning to translate between coordinate-based location referencing (e.g., latitude/longitude) and linear referencing (e.g., route-milepoint).
Level 5	*Agency-Wide:* The agency has a standard architecture for linking agency GIS and LRS data to business data systems. Methods for propagating changes in location referencing resulting from road network changes are automated. Data owners/managers work closely with agency GIS staff and proactively work to improve their data sets' consistency with agency-wide standards. *Program Specific:* Methods for propagating changes in location referencing resulting from road network changes are automated. Data owners/managers work closely with agency GIS staff and actively work to improve data sets' consistency with agency-wide standards.

Element Description

This sub-element assesses the extent to which the agency has standardized methods for location referencing, including linear referencing for its road-related data sets.

Support for the AASHTO Data Principles

❑ Valuable ☑ Available ❑ Reliable ❑ Authorized ☑ Clear ☑ Efficient ❑ Accountable

Benefits of Moving Up the Maturity Scale

At lower levels of maturity, different data sets use different methods for location referencing and standards for location referencing have not been established. This results in an inability to map information reliably and integrate different data that have a spatial component. As agencies move up the maturity scale, location referencing standards are developed and adopted, existing data sets are transformed as needed to use the standard referencing methods, and the standards applied for new data sets are collected or acquired. In addition, a process is in place to propagate changes in linear referencing to various data sets as road changes occur or as errors are corrected. Standardization and management of location referencing enables agencies to visualize and integrate data efficiently.

Relevant Improvement Actions

❑ Common Geospatial Referencing
❑ Data Change Management

Element 3: Data Architecture and Integration

3.2	Geospatial Data Management
	Standardized approach to collection and management of geospatial data
Level 1	The agency does not provide enterprise-wide planning and support for management and integration of geospatial data. Management of geospatial data is not integrated with other agency data management and IT functions.
Level 2	The agency has designated responsibilities for enterprise-wide planning and support for managing geospatial data. The agency manages a collection of spatial data sets and makes them available for internal use.
Level 3	The agency has written policies and standards defining how geospatial data is to be collected, stored, managed, shared, and integrated with non-spatial data attributes. The agency considers spatial data in their IT strategic plan (or equivalent) that identifies investment needs and priorities for hardware, software, and data. The agency has identified data entities that should have standard location referencing.
Level 4	The agency has a well-understood and functioning process for collecting, adding, and updating geospatial data sets. The agency has a standard approach to assigning spatial location to key data entities (e.g., construction projects and assets.) Training and support is provided to ensure adherence to adopted policies and standards for geospatial data collection and management and to build skills in spatial data analysis.
Level 5	Spatial data collection, management, and visualization requirements are fully integrated within the agency's IT and data management planning and operational functions. The agency periodically reevaluates and updates its approach to geospatial data management to reflect changes in technology, data availability and cost, and user requirements.

Element Description

This sub-element assesses the extent to which the agency has a standard approach to collecting, managing, and integrating spatial data.

Support for the AASHTO Data Principles

❏ Valuable ☑ Available ❏ Reliable ❏ Authorized ☑ Clear ☑ Efficient ❏ Accountable

Benefits of Moving Up the Maturity Scale

At lower levels of maturity, various methods may be used across the agency for collecting and managing spatial data. Hardware, software and services related to GIS are not standardized and are not well coordinated with "mainstream" agency functions for data management, reporting, integration, or application development. As agencies move up the maturity scale, the agency views spatial data management and reporting/mapping as integral to its overall data management and delivery function. Standard methods, processes, and tools are provided to ensure that GIS data is integrated with other agency business data. Training and support are made available to agency staff to ensure that they can make effective use of available data. Building a consistent agency approach for managing spatial data (1) promotes efficiency in use of hardware, software, and staff expertise; (2) standardizes and streamlines data integration processes, thereby reducing the need for time-consuming, repetitive tasks; and (3) ensures that various data is spatially enabled to provide business value.

Relevant Improvement Actions

❏ Data Stewardship and Governance Policies

❏ Data Delivery Platforms

❏ Common Geospatial Referencing

Element 3: Data Architecture and Integration

3.3	Data Consistency and Integration
	Standards and practices to ensure use of consistent coding and common links so that different data sets can be combined to meet business information needs
Level 1	*Agency-Wide:* There have been no formal efforts to plan for data integration/linkage across business applications outside of the context of individual application development projects. Lists of values for coded fields are defined for each application and there are no policies requiring consistency of code lists across applications. *Program Specific:* Data sets have not been reviewed to determine consistency with applicable agency or industry standards.

3.3	Data Consistency and Integration
Level 2	*Agency-Wide:* Efforts are underway to identify key integration points across data sets. Efforts are underway to identify data duplication and inconsistencies across sources. Some code lists are standardized across applications (e.g., city/county codes and organizational unit codes). *Program Specific:* Cross-reference lists have been developed to allow for data to be used in conjunction with other data sets (e.g., state versus federal project ID).
Level 3	*Agency-Wide:* The agency has identified and defined fundamental master data entities (e.g., projects, roadway segments, bridges, and employees) present in multiple business applications and has mapped which physical systems contain information related to these data entities. The agency has established a list of key link fields (e.g., route ID and project ID) and have standardized these across systems to integrate different data sets to provide answers to business questions of interest. The agency has identified single authoritative source systems for key data elements of agency-wide interest. The agency has identified common code lists and maintains these lists in a central location. *Program Specific:* Data sets/applications adhere to agency standard link fields that have been established to facilitate cross-system integration. Standard code lists are used within data sets/applications if they are available (e.g., city/county codes and organizational unit codes).
Level 4	*Agency-Wide:* The agency has procedures in place to ensure that externally procured data sets and applications adhere to established data standards and can be linked to existing data. The agency has one or more skilled individuals with responsibility for data architecture and integration across systems. *Program Specific:* Procedures are in place to ensure that externally procured data sets and applications adhere to established data standards.
Level 5	*Agency-Wide:* The agency has a process in place to assess opportunities for re-use of existing data sources as new applications come on line or new data acquisition efforts are considered. The agency maintains an agency-wide data model and uses this model to minimize data duplication and inconsistencies as new data and systems come on line. The agency has developed a "to be" data and system architecture and uses this architecture to guide system addition, replacement, consolidations, and updates. *Program Specific:* Opportunities to improve data integration and consistency with other agency data sets are reviewed annually.

Element Description

This sub-element assesses the extent to which the agency manages database creation and application development processes to minimize duplication and ensure integration.

Support for the AASHTO Data Principles

❑ Valuable ☑ Available ❑ Reliable ❑ Authorized ☑ Clear ☑ Efficient ❑ Accountable

Benefits of Moving Up the Maturity Scale

At lower levels of maturity, each new database and application development effort is implemented in isolation. Any efforts to ensure linkage with existing data is the result of individual development team initiatives—as opposed to a standard agency process. As agencies move up the maturity scale, they seek to ensure that different data sets can be linked. They manage the database and application development process to include an architectural review function that enforces standards and uses common code lists and services. This approach minimizes data duplication and facilitates data integration. It also increases efficiency of data maintenance requirements by consolidating code lists and other data tables.

Relevant Improvement Actions

❑ Data Architecture Practices and Roles

❑ Reference Data Management

❑ Master Data Management

Element 3: Data Architecture and Integration

3.4	Temporal Data Management
	Standardization of date-time data elements to enable trend analysis and integration across data sets collected and updated on varying cycles
Level 1	*Agency-Wide and Program Specific:* No standards or guidelines are in place regarding date- and time-related data elements (e.g., naming of fiscal versus calendar year, distinguishing data collection dates, data loading or update dates, and data effective dates). There is no defined strategy for integrating data sets to provide a consistent "point-in-time" view of integrated information.
Level 2	*Agency-Wide and Program Specific:* Naming conventions and common practices are in use regarding date- and time-related data elements, but no written guidelines exist. There is some understanding of user needs for trend analysis and creating snapshot views of data for analysis and reporting, but these needs have not been explored systematically or comprehensively. There is experience with integrating data to create a snapshot-in-time view, but no repeatable procedures for this have been defined.

3.4	Temporal Data Management
Level 3	*Agency-Wide and Program Specific:* There are documented guidelines for ensuring consistency in use of date- and time-related data elements across data sets and applications. Data user requirements for trend analysis, snapshots, and other uses of temporal information have been documented. There are documented procedures or models for integrating across data sets to create a snapshot-in-time view.
Level 4	*Agency-Wide:* There is consistency across the agency's major data sets in use of date- and time-related data elements across data sets and applications. Data user requirements for trend analysis, snapshots, and other uses of temporal information can be met without major changes to data structures or substantial new development effort. *Program Specific:* Data user requirements for trend analysis, snapshots, and other uses of temporal information can be met without major changes to data structures or substantial new development effort.
Level 5	*Agency-Wide and Program Specific:* Data user requirements for trend analysis, snapshots, and other uses of temporal information can be met through largely automated processes.

Element Description

This sub-element assesses the extent to which requirements for standardizing temporal data elements are considered so as to ensure that data representing different periods can be combined as needed to support analysis.

Support for the AASHTO Data Principles

❏ Valuable ☑ Available ❏ Reliable ❏ Authorized ☑ Clear ☑ Efficient ❏ Accountable

Benefits of Moving Up the Maturity Scale

At lower levels of maturity, treatment of temporal data elements is not standardized—each new database and application development effort determines its own formats and requirements. As agencies move up the maturity scale, they consider business requirements for time-based queries and trend analysis. Based on these business requirements, they establish standards for temporal data elements (e.g., always use month and year to convert between calendar and fiscal year; always distinguish between data update date and the effective date of an observation.) In addition they establish processes to create snapshots of data sets to represent point-in-time conditions as needed for specific business purposes (e.g., safety analysis). Analogous to standardization of spatial referencing, a standard approach to temporal referencing ensures that different data sets can be integrated to provide business value. For example, both "when" and "where" are key questions for understanding cause-and-effect relationships among system performance, crashes and fatalities, asset condition, construction project completion, weather events, and land development.

Relevant Improvement Actions

❏ Data Architecture Practices and Roles

❏ Standardized Approach to Temporal Data

Element 4: Data Collaboration

4.1	Internal Agency Collaboration
	Collaboration across agency business units to leverage opportunities for efficiencies in data collection and management
Level 1	*Agency-Wide:* Most data collection efforts in the agency are independent—there has been little or no effort to coordinate across business units. The agency does not have information about the extent of data duplication. *Program Specific:* There have been no efforts to coordinate data collection or management activities with other business units.
Level 2	*Agency-Wide:* The agency has assessed the extent to which there is duplication across data sets within the agency. Opportunities for coordinating data collection and management across business units (e.g., safety and asset management) are periodically discussed, but limited progress has been made. *Program Specific:* Opportunities for coordinating data collection and/or management activities with other business units have been discussed, but no action has been taken.
Level 3	*Agency-Wide:* The agency has implemented a data collection effort involving coordination of more than one business unit (e.g., use of video imagery from pavement data collection to extract data on other assets). The agency has defined metrics to track improvements in data collection and storage efficiency. *Program Specific:* A specific opportunity for coordinated data collection has been identified and is being pursued.
Level 4	*Agency-Wide:* Agency business data owners are encouraged and incentivized to share their data with a broader audience within the agency (where appropriate). Agency business data owners are encouraged and incentivized to plan new data collection initiatives in partnerships with other business units where information needs of multiple units can be simultaneously addressed. The agency monitors progress of efforts to reduce data duplication. *Program Specific:* Data collection is routinely coordinated with one or more other business units.
Level 5	*Agency-Wide:* The agency periodically reviews its data collection programs to identify opportunities to leverage new technologies and externally available data sources. The agency regularly seeks opportunities to minimize or reduce redundancy in data collection, storage, and processing. *Program Specific:* New internal agency partnerships on data collection and management are actively sought to achieve economies of scale and make best use of limited staff and budget.

Element Description

This sub-element assesses the extent to which there is collaboration and coordination across different organizational units on data collection and management.

Support for the AASHTO Data Principles

❑ Valuable ☑ Available ❑ Reliable ❑ Authorized ☑ Clear ☑ Efficient ❑ Accountable

Benefits of Moving Up the Maturity Scale

At lower levels of maturity, data collection and acquisition efforts are planned and executed independently to meet the needs of different business units. Each business unit views the data they collect as "their own" and doesn't consider the possible value of sharing the data with others in the agency. As agencies move up the maturity scale, data collection efforts are coordinated across business units and data is shared. Data partnerships are encouraged and incentivized. New data collection technologies are pursued that can provide multiple types of data at once (e.g., videologs and LiDAR). In addition, business units work closely with the agency's IT group to take advantage of enterprise reporting and other data access platforms. A collaborative approach to data collection and management reduces duplicative efforts and prevents proliferation of multiple overlapping and inconsistent data sets.

Relevant Improvement Actions

❑ Multi-Purpose Data Collection

❑ Data Outsourcing

❑ Data Business Plans

❑ Data Governance Bodies

Element 4: Data Collaboration

4.2	External Agency Collaboration
	Partnerships with external entities to share data and avoid duplication
Level 1	*Agency-Wide:* Individual business units obtain and use publicly available data from external entities as needs and opportunities arise. The agency has acquired single "point-in-time" data sets from external entities. *Program Specific:* Publicly available data from external entities is obtained and used as needs and opportunities arise.
Level 2	*Agency-Wide:* The agency is exploring partnerships with other public- and private-sector organizations to share data on an ongoing basis. *Program Specific:* Partnerships with other public- and private-sector organizations are being explored to share data on an ongoing basis.
Level 3	*Agency-Wide:* The agency has data-sharing agreements with external entities. The agency provides "self-serve" access to data sets of value to external users. *Program Specific:* Data-sharing agreements are in place with external entities. "Self-serve" access is provided to data sets of value to external users.

4.2	External Agency Collaboration
Level 4	*Agency-Wide:* The agency has sustained partnerships with external entities involving regular update cycles. *Program Specific:* Data-sharing agreements with external entities have been sustained over time (2+ years) and through multiple data update cycles.
Level 5	*Agency-Wide:* The agency routinely seeks new opportunities for data partnerships with external entities. They have designated staff liaison responsibilities for managing these external partnerships. *Program Specific:* New opportunities for data partnerships with external entities are actively sought. Staff liaison responsibilities for managing these external partnerships have been designated.

Element Description

This sub-element assesses the extent to which the agency seeks out externally available data and develops data-sharing arrangements and partnerships with external public- and private-sector entities.

Support for the AASHTO Data Principles

❑ Valuable ☑ Available ❑ Reliable ❑ Authorized ☑ Clear ☑ Efficient ❑ Accountable

Benefits of Moving Up the Maturity Scale

At lower levels of maturity, staff in different business units may seek out and acquire data sets from external entities on a one-time basis as needs arise. External requests for agency data sets are considered on a case-by-case basis. As agencies move up the maturity scale, data-sharing agreements are developed as appropriate to make best use of both internal and external data resources. Rather than making or fulfilling a series of one-time, special data requests, regular processes are set up to share data on an ongoing basis. The agency provides self-serve access to key data sets for which there are frequent requests. An active approach to external data collaboration saves the agency staff time in fulfilling data requests and provides opportunities for the agency to gain access to a richer pool of data at a lower cost than would be required if it were to collect and manage the data in house.

Relevant Improvement Actions

❑ Data Clearinghouses/Open Data Platforms

❑ Data-Sharing Agreements

❑ Data Partnerships

Element 5: Data Quality

5.1	Data Quality Measurement and Reporting
	Metrics and reporting to ensure user understanding of current data quality
Level 1	*Agency-Wide:* There are no agency-wide activities related to data quality measurement and reporting. *Program Specific:* Data quality metrics have not been identified.
Level 2	*Agency-Wide:* The agency is exploring establishment of common data quality metrics for shared data elements. *Program Specific:* Metrics for data quality are being defined.
Level 3	*Agency-Wide:* The agency has defined common data quality metrics across data programs to integrate data (e.g., locational accuracy). *Program Specific:* Metrics and standards for accuracy, including location accuracy, are defined and documented. Metrics and standards for timeliness and currency are defined and documented. Metrics and standards for completeness, including coverage or required entities/areas and inclusion of required attributes, have been defined and documented.
Level 4	*Agency-Wide:* The agency has implemented data quality standards, verification techniques, and reports for common data elements. *Program Specific:* Processes are in place to measure and document the level of accuracy, currency, and completeness of data sets. Information about data accuracy, currency, and completeness is routinely shared with users. Where data is based on sampling, information about confidence levels is made available to data users.
Level 5	*Agency-Wide:* The agency identifies new areas where common data quality metrics across data programs would be beneficial. *Program Specific:* Data quality measurement processes, metrics, and measurement techniques are reviewed periodically and refined as needed.

Element Description

This sub-element assesses the extent to which data quality metrics have been defined and used to inform users about the level of currency, accuracy, coverage, and completeness of a given data set. Data reliability is considered to be related to accuracy and is not distinguished here as a separate characteristic. Data integrity, consistency, and confidentiality are other important aspects of data quality considered as part of Assessment Elements 2 and 3.

Support for the AASHTO Data Principles

❑ Valuable ❑ Available ☑ Reliable ❑ Authorized ☑ Clear ❑ Efficient ☑ Accountable

Benefits of Moving Up the Maturity Scale

At lower levels of maturity, there is a lack of awareness about the quality of different agency data sets beyond anecdotal information. As agencies move up the maturity scale, they measure and report on data quality using metrics reflecting key characteristics of concern to potential users. The agency provides standard definitions of different data quality metrics and models for how to measure data quality to facilitate application within different data program areas and enable data users to become familiar with a consistent set of measures. Providing data users with data quality metrics can help users to determine whether a data set is sufficiently accurate to meet their needs and help to address lack of trust in data that users may have as a result of seeing a single example of an erroneous data value. Finally, it can provide a basis for initiating data quality improvement efforts and tracking their progress. Data quality measurement can be costly, so it is important to focus on a few essential measures and take advantage of quality metrics that can be automatically generated (e.g., adherence to validation rules).

Relevant Improvement Actions

❑ Data Quality Measurement and Improvement

Element 5: Data Quality

5.2	Data Quality Assurance and Improvement
	Practices for improving quality of existing data and ensuring quality of newly acquired data
Level 1	*Agency-Wide:* Data quality is assessed and improved in the context of individual data programs. No agency-wide support is provided. *Program Specific:* Data quality is addressed ad hoc in response to reported issues. There is no standard approach in place for quality assurance for new data collection and acquisition.
Level 2	*Agency-Wide:* Limited technical assistance is available for data program managers or business units on fundamental data quality concepts and practices. *Program Specific:* There have been efforts to work with data users to discuss and define data quality requirements. Standard practices are being defined to ensure the quality of data collected or acquired.
Level 3	*Agency-Wide:* The agency has established guidelines for determining spatial accuracy requirements and appropriate collection methods for new data collection efforts. The agency incorporates practices supporting data quality within the standard software development process, including definition and documentation of business rules for data validation and use of standard lists of values. *Program Specific:* Standard, documented data quality assurance and improvement processes are defined. Business rules for assessing data validity have been defined. Specific guidance and procedures for data collection and processing is routinely provided to ensure consistency. A formal process for data certification and acceptance has been defined—including provision for correcting or re-collecting data when it does not meet minimum requirements for accuracy.

5.2	Data Quality Assurance and Improvement
Level 4	*Agency-Wide:* The agency provides standard tools for gathering and tracking response to user feedback on data quality issues. The agency has deployed data profiling and cleansing tools and uses these tools to identify (and, where possible) correct data quality issues. *Program Specific:* Standard, documented data quality assurance processes are routinely followed. Data collection personnel are trained and certified based on demonstrated understanding of standard practices. Business rules for data validity are built into data entry and collection applications.
Level 5	*Agency-Wide:* The agency provides tools for specification, maintenance, and management of business rules. *Program Specific:* Data quality assurance processes are regularly assessed and improved. Data collection and acquisition practices are regularly reviewed to identify lessons learned and areas for improvement. Automated error reporting tools are available for data users. Data validation and cleansing tools are used to identify and address missing or invalid values.

Element Description

This sub-element assesses the extent to which the agency pursues a systematic and proactive approach to data quality assurance and improvement.

Support for the AASHTO Data Principles

☑ Valuable ☑ Available ☑ Reliable ☑ Authorized ☑ Clear ☑ Efficient ☑ Accountable

Benefits of Moving Up the Maturity Scale

At lower levels of maturity, data quality is addressed as issues are reported. Staff responsible for initiating new data collection efforts do not have any standard agency guidance to follow for inclusion of data quality assurance practices in the effort. As agencies move up the maturity scale, data quality is addressed actively, using multiple techniques. These include use of standard quality control and quality assurance processes for new data collection, development and application of data validation business rules, use of automated data cleansing processes to identify potentially erroneous data values, and establishment of efficient error reporting and correction processes. Data quality improvement efforts need to be tailored to specific data types and collection methods. Appropriate application of data quality improvement techniques is important to ensure that data can be used as intended and can be used to produce reliable information that is valuable for decision making.

Relevant Improvement Actions

❑ Data Quality Measurement and Improvement

APPENDIX D

Data Improvement Catalog

This appendix provides two resources for identifying potential data improvements. The first resource is a checklist of data improvement ideas, organized by type of improvement (e.g., information technology, data collection, and data governance). This checklist includes ideas that can address gaps from the data value assessment (availability, quality, or usability) or the data management assessment. The second resource is a longer list of improvements to data management practices, organized by the four data management assessment elements. This resource provides example applications of different improvement types, as well as selected references that can provide additional information.

1. Checklist of Data Improvement Ideas by Category

Information Technology
- ❑ Implement new or upgraded source system for data
- ❑ Sunset existing data source system
- ❑ Implement new or improved data access solution (e.g., GIS portal, business intelligence tool)
- ❑ Deploy mobile solution for data collection and/or data access
- ❑ Implement new or upgraded data integration solution (e.g., data warehouse, Extract-Transform-Load (ETL), enterprise service bus, and master data management)
- ❑ Implement an enterprise meta data repository
- ❑ Implement data profiling/data cleansing tools

Data Presentation and Analysis
- ❑ Create new or improved data summaries and reports
- ❑ Create new or improved data visualizations
- ❑ Make data available via GIS
- ❑ Develop new or improved data mining or analysis methods
- ❑ Integrate data sources for improved insight

Data Investigation and Documentation
- ❑ Conduct a data needs study for one or more business areas
- ❑ Conduct a data risk assessment (e.g., identify data in non-enterprise systems and data sets with personally identifiable information)
- ❑ Conduct data value mapping (map how data is used within business processes)
- ❑ Conduct data production mapping (map how data is produced and transformed)
- ❑ Rationalize/prune existing data (drop data elements and tables that are redundant or not actively used or needed)
- ❑ Develop or improve data collection manuals and associated training materials
- ❑ Establish a Data Community of Interest
- ❑ Create or update data catalogs and data dictionaries—establish processes to keep these updated as data sets change
- ❑ Develop data quality metrics
- ❑ Formalize procedures for data quality assessment and reporting

❑   Develop and document data validation business rules
❑   Identify business requirements for historical data
❑   Enhance communication strategies to increase awareness of agency data products and services
❑   Create standard new staff orientation briefing on available data and how to use it

Data Collection/Processing/Quality Improvement

❑   Initiate new data collection or acquisition effort
❑   Discontinue current data collection effort
❑   Change spatial coverage and/or frequency of existing data collection effort
❑   Change existing data collection method (e.g., outsourcing, new technology, different equipment, and new sampling approach)
❑   Implement new/improved data quality assurance/quality control procedure
❑   Coordinate or combine existing data collection efforts across business areas or across the enterprise
❑   Harvest data from existing processes or sources (e.g., asset extraction from as-builts)
❑   Change data elements, level of granularity, or data structure for existing data set
❑   Add/improve spatial referencing for existing data sets
❑   Negotiate data-sharing agreement
❑   Establish Service Level Agreement between data provider and user
❑   Convert or migrate existing data
❑   Clean up existing data

Data Governance/Policy/Procedure/Standards

❑   Establish data governance body or modify/strengthen charter for existing body
❑   Develop and adopt data principles and supporting policies
❑   Develop a strategic data business plan for the organization or a particular business function
❑   Define and designate data stewardship roles and responsibilities
❑   Identify categories of data to be managed as a corporate asset
❑   Classify data based on sensitivity—including designation of data that can be shared externally
❑   Designate a source system of record for shared agency data elements
❑   Implement a reference (code list) data management strategy
❑   Implement a master data management strategy
❑   Develop a data change management process
❑   Develop agency data standards (e.g., for location data and project identifiers)
❑   Develop an agency business data glossary
❑   Document standard "design patterns" for managing access to historical data (e.g., snapshot creation)
❑   Review and revise data retention policies
❑   Create standard methods for managing data access
❑   Develop "standard operating procedures" for data collection, updating, and quality assurance

Data Management Staffing and Responsibilities

❑   Integrate data management core competencies into position descriptions
❑   Conduct staff training for specialized data management tasks
❑   Create new staff positions (e.g., data architect, business analyst, and data manager)
❑   Identify staff with critical data knowledge and begin knowledge transfer and succession planning

2. Data Management Improvement Strategies, Examples and Resources (organized by assessment element)

ELEMENT 1:	Data Strategy and Governance
Improvements	Agencies can draw on various techniques for strengthening data governance and data strategic planning: • **Data Governance Bodies**—formal organizational structures to oversee data management policies, projects, initiatives, and investments • **Data Governance and Stewardship Policies**—adopting principles, policies, and business processes for managing data as a strategic agency asset • **Data Business Plans**—developing plans that identify the data required to meet agency business needs and develop strategies for ongoing management of data • **Data Management Roles and Responsibilities**—defining and documenting data management responsibilities for data management; building these into employee position descriptions • **Data Value Mapping**—diagramming how data is used within agency business processes to understand and document the value such data add • **Data Communities of Interest**—bringing together data users and producers to provide an ongoing forum for identifying and prioritizing data improvement needs and strategies based on multiple perspectives
Examples	**Alaska Department of Transportation and Public Facilities (DTPF)** In Alaska, a comprehensive data business planning effort produced an inventory of data programs, applications, and systems; an enhanced understanding of the relationships between data, applications, and business needs; and a framework for implementing data governance. As part of Alaska DTPF's data business planning activities, they have made extensive use of Unified Modeling Language (UML) diagrams to define use cases and illustrate how data is produced and consumed. For more information: Jack Stickel, or Jill Sullivan, Alaska DTPF, Jack.Stickel@alaska.gov or Jill.Sullivan@alaska.gov **California Department of Transportation (Caltrans)** Caltrans has established a Transportation System Data Governance Board made up of eleven division and three district office representatives. The Board's mission is to ensure that Caltrans creates and maintains reliable transportation system data accessible to the department and its partners. The Board is responsible for all aspects of data ownership, standards setting, collaboration, reporting, and other activities. Caltrans has also developed a Transportation System Data Business Plan. The plan identified key issues facing the department, a data governance approach for resolving issues, and an organizational framework to initiate and manage the plan. The plan includes

- A governance approach to use data as an enterprise asset, including mission goals and success measures
- Definitions of governance roles and responsibilities
- The development of core business processes for successful governance implementation and detailed flow charts outlining steps, roles, and responsibilities for governance activities (e.g., prioritizing governance issues, monitoring progress on tasks and projects, and establishing and implementing new policies)
- Descriptions of data products and 12 corporate datasets
- Processes for doing data quality assessments and the results of applying the processes to several core datasets
- Assessment criteria for evaluating data processes and corporate data and the results of applying the criteria to three high-priority datasets
- Recommendations for an enterprise data architecture, including attributes and benefits of meta data and an outline for data catalogs
- An implementation plan and schedule

For more information: Coco Briseno, California DOT, coco.briseno@dot.ca.gov

http://dot.ca.gov/hq/tsip/data_library/data_governance/CTS_DataBusinessPlan_8_29_11.pdf

State of Colorado Data Management Program

The State of Colorado has established a strong data management program in the Governor's Office of Information Technology. The program uses data and information as enterprise assets and establishes standards and processes to support more flexible government services. In 2009, Colorado created the nation's first state chief data officer position and began developing enterprise data models, enterprise architecture meta data management, and data quality management. The state's data management program requires a strong program of information sharing to serve citizens more effectively, improve efficiency and effectiveness, and inform policy making.

For more information: http://www.colorado.gov/cs/Satellite/OIT-Main/CBON/1251575408707

Colorado Department of Transportation (CDOT)

Colorado DOT data management activities build on what was initiated at the state level in the Colorado State Office of Information Technology. Consistent with state policy, CDOT established a Knowledge Management Governance Oversight Committee to support a knowledge management governance framework within the department. The vision of the Committee is to implement policies, procedures, and standards to be used in managing information, data, and content within CDOT to support the department's mission and goals. The scope of the Committee is to provide recommendations to the Information Technology Management Team for implementation of standards, procedures, and work products for the enterprise as defined above, in coordination with implementation teams. Each year the Committee publishes a report of accomplishments

Responsibilities of the Committee are to

- Develop a strategy and process for implementing knowledge management governance throughout the organization, to include the following:
    - Guidance on priorities for implementing governance on enterprise information, data, and content
    - Prioritization of governance tasks
    - Creation and tasking of implementation teams
    - Guidance on developing a governance manual documenting the framework
    - Developing a plan for communicating the initiative to the organization
    - Identifying a process for change management and training to support the governance framework and initiatives.
- Develop and recommend a detailed governance framework to
    - Define roles and responsibilities in governance (e.g., board/council, steward, custodian, and stakeholder)
    - Define goals/objectives pertaining to the creation, retention, distribution, and use of information, data, and content
    - Identify the value, business use, and priority of information, data, and content
    - Define the requirements for developing a knowledge catalog for the department

In 2011, CDOT completed a Performance Data Business Plan to support enhanced data management, performance reporting, and decision making within the agency. The report recommended performance measures for fatalities, bridge conditions, pavement conditions, roadside conditions, snow and ice control, roadway congestion, on-time construction, on-budget construction, and strategic action item implementation. In addition, the plan addressed data management methods, best practices, and recommendations for data governance. The work produced a data inventory, a data catalog, and a sample data governance work team charter. It also recommended measures for assessing data quality.

For more information: William Johnson, Colorado DOT, will.johnson@state.co.us

Michigan Department of Transportation (MDOT)

MDOT has established a Department Data Governance Council. The Council's charter includes a data management vision, mission, and core principles. The Council meets at least monthly, and IT resources have been assigned to assist the Council. Council responsibilities include

- Supporting the creation of data life cycle documentation
- Establishing and maintaining data management principles
- Developing, maintaining, and ensuring adherence to data management best practices, standards, functions, and data use and re-use guidelines
- Advocating project-level standards
- Providing direction to IT teams

- Coordinating information sharing
- Moderating issues from data stewards
- Sponsoring the knowledge base
- Verifying adherence to standard data concepts and definitions defined by the department and the statewide DTMB
- Empowering functional areas to audit and enforce data integrity in compliance with the Data Management Policy
- Recommending data stewardship resources and resource levels
- Reporting progress to executives and monitoring industry and corporate trends

The Council has adopted a three-tier model for implementing data governance in the department. MDOT has defined a data manager role with the current focus on key data categories—including capital program data, GIS, and asset data. The data manager works with the agency Chief Data Steward to implement data improvements.

For more information: Ron Vibbert, Michigan DOT, VIBBERTR@michigan.gov

State of Minnesota

In Minnesota, executives from eight cabinet-level agencies meet quarterly in a State Data Governance Advisory Council to discuss coordination opportunities and strategies for

- Increasing efficiencies in data management
- Using data sets across state agencies
- Minimizing and managing breaches in data security
- Establishing an overall statewide data architecture

At the time of this writing, a survey of state agencies was planned for 2015 to assess data management maturity levels.

For more information: Minnesota IT, Jon Eichten, Jon.eichten@state.mn.us

Minnesota Department of Transportation (MnDOT)

Data business planning was undertaken to recommend strategies and actions to address priority data gaps and needs in the areas of safety, infrastructure preservation, and mobility; strengthen data governance; and validate and provide strategic direction for GIS.

In 2011, MnDOT established a Governance Council to implement actions to strengthen data governance. Among these actions were the establishment of business data domains and the definition of stewardship roles and responsibilities. Nine data domains have been identified that cover all basic data uses of the department:

- Human resources data
- Financial data
- Planning, programming, and project data
- Infrastructure data
- Business stakeholder and customer data

- Spatial data
- Regulatory data
- Recorded events data
- Supporting assets data

Within the domains, 120 subject area stewards have been identified and are undergoing training to clarify stewardship expectations, role, and responsibilities. Data domain stewards meet monthly. A representative from the statewide IT group also attends. To date, MnDOT domain stewards have focused on

- Scoping IT projects in the context of identified data principles to minimize redundancies and foster discussion of how a project in one area may have broader effects on other areas or data systems in the department
- Identifying enterprise and authoritative sources of data and clarifying ownership responsibilities
- Discussing data retention needs and policies
- Reviewing data access policies
- Identifying data-sharing opportunities within and external to the department and developing service level agreements to establish expectations.

For more information: Mark Nelson, Minnesota DOT, mark.b.nelson@state.mn.us

US DOT FHWA

The FHWA Office of Operations Management developed Roadway Transport Data Business Plan Phase I and Phase II reports to improve coordination and communications and strengthen data governance across USDOT offices involved with roadway travel mobility data. Recommendations to improve coordination and communications focused on identifying gaps and redundancies in roadway travel mobility data programs and devising "rules of engagement" regarding collaboration of the data functions for roadway travel mobility data.

The Data Business Plan recommends the establishment of a Mobility Data Coordination Group to (1) coordinate data issues affecting roadway mobility data and (2) work on cross-cutting data management issues related to data quality, privacy, and security. The USDOT Mobility Data Coordination Group would serve as an umbrella structure for smaller working groups who would meet to coordinate on specific data issues (e.g., travel data, modal data, or climate data). The plan also recommended the creation of an internal Community of Interest that would

- Coordinate on cross-cutting issues that affect data from all of the working groups
- Represent those who use, access, integrate, or benefit from improved coordination and data management.

The plan includes specific information on the roles and responsibilities of the Mobility Data Coordination Group, the working groups, and the internal and external communities of interest.

For additional information: http://ntl.bts.gov/lib/48000/48500/48531/6E33210B.pdf -http://ntl.bts.gov/lib/48000/48500/48531/6E33210B.pdf

Virginia Department of Transportation (VDOT)

VDOT's System Operations Directorate created a data business plan that provided a framework for meeting the Department's data needs related to maintenance, traffic engineering, and traffic operations. The plan development process involved vision development sessions with key stakeholder groups, analysis of existing data sets, and identification of key gaps in data. The plan recommended a stewardship model to ensure ongoing management and improvement of the agency's data resources. One of the strategies recommended in the data business plan was formation of Data Communities of Interest (COIs). A data COI includes staff from different units with the department to collaborate on developing recommendations and guidelines about data needs. Five different COIs were established—for work planning and tracking data, financial data, bridge data, traffic and safety asset data, and ITS asset data.

For further information: Bob Boothe, VDOT—Bob.Boothe@vdot.virginia.gov

Virginia Transportation Research Council (VTRC)/Virginia DOT (VDOT)

VTRC conducted a business process modeling project for VDOT to describe how several planning and programming activities could be integrated. Underlying reasons for doing the process modeling were to ensure that

- Resources for construction projects are used effectively
- Employees know where projects are in their construction life cycles and how projects may have been changed
- The time of agency employees is used effectively
- The employees are working together to complete transportation projects in a reasonable time

The process modeling effort included a step for documenting who generates what information, products, and services; for whom; how; and for what reasons. The process encouraged the development of integrated systems across functional areas and business activities.

For more information: http://www.virginiadot.org/vtrc/main/online_reports/pdf/05-cr15.pdf

Washington State Department of Transportation (WSDOT)

Data Governance. In January 2015, WSDOT issued a Secretary's Executive Order creating a high-level Enterprise Information Governance Group (EIGG). The EIGG serves as the policy-setting body for the department on data and information management issues and is responsible for establishing direction and setting policy that facilitates management of data and information in alignment with eight identified data and information principles. The Executive Order directs the EIGG to

- Review existing data and information policies and periodically prepare reports summarizing the effectiveness of current practices while implementing work plans to address gaps, inconsistencies, and any conflicting or unclear direction
- Develop policies that promote efficient and strategic use of data and information resources for all aspects of data collection, storage, management, findability, and access
- Identify roles and responsibilities for enforcement, accountability, and authority that support conformance with the data and information principles
- Provide executives with annual reports on accomplishments, improvements resulting from policy changes, and policy issues under consideration.

The Executive Order further directs all employees to make efficient and strategic use of data and information and directs WSDOT regions, executives, directors, and employees to align their practices with the data and information principles and policies.

WSDOT has adopted the following data and information principles:

1. Data and information are critical to effective business decision making at WSDOT and shall be maintained in a manner appropriate to meet business needs.
2. Data and information are strategic, long-term assets owned by WSDOT, not by individual business units. They are findable, retrievable, and shared.
3. Data and information shall be collected once, stored once, and used multiple times.
4. Data and information that is not used shall not be collected or stored.
5. Data and information that is used by multiple applications or shared across business units shall be defined and managed from an enterprise perspective and fit for various applications.
6. Data and information investments will consider business priorities, program impacts, and tradeoffs.
7. Data and information shall be managed to provide availability, security, and integrity—they shall be safe from harm and accessible by those who need them.
8. Data and information governance, costs, and stewardship processes will be transparent

For more information: Leni Oman, WSDOT, Omanl@wsdot.wa.gov

Data Value Mapping. WSDOT conducted a data value mapping exercise for their Highway Safety Project Programming Process. A diagram was produced that shows the data collection, supporting data, data processing, and information needed to support all of the safety project activities associated with the planning process, preliminary programming, design, construction, maintenance and traffic operations, and monitoring, reporting performance, and asset management.

For more information: Ida van Schalkwyk, WSDOT, VanSchI@wsdot.wa.gov

Data Business Planning. WSDOT conducted research and developed a state Freight Data System to address user needs for data on the economic impact of freight, system bottlenecks, and supply chains. As part of the research, the department

- Completed an inventory of current freight data sources and compiled a database
- Surveyed other state DOTs to learn about how freight data is being used, needs for freight data, approaches for addressing data needs, and attempted solutions

- Conducted workshops around the state to determine freight data user needs
- Identified data gaps, redundancies, inaccuracies, and weaknesses in current data collection.

Research results revealed the absence of links between different data sources and gaps in the availability of needed data. Combining data sources additionally raised concerns about the quality and consistency of fused data. The research recommended development of a maintainable, systematic, and coordinated data collection framework. The framework will have

- A new Freight Data Librarian/Educator to lead the effort and interact with state freight clientele, develop data partnerships, and serve as the freight data source for the state and
- A Freight Database Manager who will develop a freight data warehouse and provide technical support.

Other recommendations called for additional ongoing origin and destination surveys and studies of carriers at the statewide, urban area, and county road levels.

For more information: http://www.wsdot.wa.gov/research/reports/fullreports/690.1pdf

Resources	The Data Management Association Data Management Body of Knowledge http://www.dama.org/content/body-knowledge
	National Association of State Chief Information Officers Data Governance Articles: http://www.nascio.org/publications/documents/NASCIO-DataGovernance-Part1.pdf
	http://www.nascio.org/publications/documents/nascio-datagovernanceptii.pdf
	http://www.nascio.org/publications/documents/nascio-datagovernanceptiii.pdf
	Data Governance Institute www.DataGovernance.com
	Object Management Group—Business Process Modeling http://www.bpmn.org/
	Object Management Group—Unified Modeling Language http://www.uml.org/
	The Open Group Architecture Framework https://www.opengroup.org/togaf/

ELEMENT 2:	Data Life Cycle Management
Improvements	Data life cycle management practices include • **Standard Operating Procedures**—for data collection, updating, loading, backups, and archiving • **Data Change Management**—data change impact analysis and governance processes to minimize unintended consequences of changes to data structures or codes • **Data Catalogs and Dictionaries**—documenting data tables and columns in a standard manner; providing catalogs of agency data sets that facilitate understanding of and access to available data • **Data Curation Profiles**—a standard method for documenting "the story" of a research data set—describing its origin and role in a research project • **Data Management Plans**—plans that describe how data sets are to be managed throughout their life cycle, covering formats, documentation, storage, access, and re-use • **Data Retention Schedules and Archiving**—processes for determining how long different data sets will be kept, and strategies for archiving data that need to be retained, but are not in active use • **Data Access Policies**—classification of data sets for controlling access to sensitive or confidential data; establishment of policies for data access • **Data Delivery Platforms**—implementation of data query and reporting tools to facilitate delivery of data to users in various convenient, useful, and usable forms
Examples	**Cornell University—Data Curation** Scientists at Cornell University researched methods for increasing University and public access to demographic data. The research was designed to investigate the idea of using an external data repository that could offer web APIs similar to those being used at the University and that the public could use to access data. The scientists acquire demographic data from the US Census and various other sources and add value by processing, analyzing, and distributing the data on their project website so as to make the data more accessible and easier to use. The research included an inventory of data sources (including how they are acquired, what they contain, and their size, format, and meta data), an analysis of aggregation and analysis needs, an assessment of mapping requirements, and recommendations for data access tools. The research was designed to improve current methods and strengthen accessibility for internal and external demographic data users. For more information: http://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1026&context=dcp

Michigan DOT (MDOT)

Change Management. MDOT is using a commercial information management platform to build a meta data repository. When fully populated, this repository will enable change impact analysis by identifying data tables that contain a specific data element— or an element derived from another element—that is being changed.

For more information: Ron Vibbert, Michigan DOT, VIBBERTR@michigan.gov

Data Delivery. MDOT publishes the Michigan Traffic Crash Report as an interactive website that summarizes historical and annual crash trends and characteristics. The website was the winner of the 2014 "Best Traffic Records Web Page" award presented by the Association of Transportation Safety Information Professionals (ATSIP).

For more information: www.michigantrafficcrashfacts.org

Minnesota Department of Transportation (MnDOT)

Data Catalog. MnDOT implemented a Data Business Catalog application. Designated data stewards throughout the department identity and document data items within their designated domain areas. Data terms, along with pertinent information or meta data, are published in the Business Data Catalog and made available to staff. The meta data elements for each data item include the approved term name, term definition, and source of record, data classification, and responsible data steward. The Business Data Catalog helps prevent data redundancy and identifies opportunities to leverage investment in information technology.

For more information: John Solberg, Minnesota DOT, john.solberg@state.mn.us

Data Access Policies. MnDOT has established an online guide and process for requesting information consistent with the Minnesota Government Data Practices Act. The Guide outlines who has the right to access public data, how to make a request, how the department will respond, and how long it will take to get requested information. The policy includes a copy request form, along with data practices contacts. The Guide also addresses requests for creating new data or providing data in a specific form as well as copy costs.

For more information:
http://www.dot.state.mn.us/information/datapractices/index.html

Retention Schedules. MnDOT has established a records retention schedule organized by data domain and subject area. The retention schedule has been added to the Business Data Catalog so that users can search for records and achieve more reporting flexibility. Staff can get the complete MnDOT records retention schedule or generate reports only for records assigned to a particular data domain or subject area or filtered by other criteria. Additional enhancements to boost usability include

- Creation of pages for publishing separate information and updates for data terms and records

- Replacement of static reports with more interactive, flexible, user-driven reports in the agency's crystal reports web portal

For more information: Charles Stech, Minnesota DOT, Charles.stech@state.mn.us

Oregon Department of Administrative Services

The Oregon Department of Administration Services (DAS) has adopted a formal policy to ensure that the state's information assets are identified, properly classified, and protected throughout their life cycles. The policy provides that all state agency information will be classified and managed based on its confidentiality, sensitivity, value, and availability requirements, consistent with the Oregon Public Records Law. The four sensitivity levels are

- Level 1—"Published." This includes low-sensitivity information that is not protected from disclosure and will not jeopardize the privacy or security of agency employees, clients, and partners. This includes information regularly made available to the public.
- Level 2—"Limited." This includes sensitive information that may not be protected from public disclosure but, if made easily or readily available, could jeopardize the security or privacy of employees, clients, and partners. Examples might include audit reports and risk management planning documents.
- Level 3 –"Restricted." This includes sensitive information intended for limited business use. The information in this category typically may only be accessed and used by authorized internal parties in the performance of their duties. External parties must be under contractual obligation of confidentiality. Security threats at this level include changes to or destruction of data, unauthorized disclosure, and violation of privacy practices. Unauthorized access and use could result in financial loss or identify theft.
- Level 4—"Critical." This includes information that is extremely sensitive and intended for use only by "named" individuals. This information is generally exempt from public disclosure because it may cause major damage or injury to named individual(s), employees, clients, or partners or cause damage to the agency.

The policy includes labeling and handling conventions for limited or restricted critical information and outlines data disposal guidelines.

For more information: http://www.oregon.gov/DAS/OP/docs/policy/state/107-004-050.pdf

Texas Department of Transportation (TxDOT)

TxDOT has established a formal policy that assigns responsibilities for maintaining roadway information. The policy defines how the Transportation Planning and Programming (TPP) Division shares the responsibilities of roadway data maintenance with all district offices and the Construction Division. The manual includes electronic links to definitions for all of the terms included in the policy and links to district personnel responsible for maintaining data.

For more information:
http://onlinemanuals.txdot.gov/txdotmanuals/trm/data_maintenance_responsibility.htm

Utah Department of Transportation (UDOT)

UDOT has gained national attention with the implementation of the UDOT GIS Access to the Transportation Enterprise (UGATE) and UPlan projects. UGATE is a spatially enabled data warehouse; UPlan provides access to a wide variety of geographic information in the department—in map, tabular, and straight-line diagram forms. The system was designed to provide a flexible, scalable platform for data sharing to promote effective decision making throughout the department.

For additional information:
http://www.gis.fhwa.dot.gov/documents/Cloud_Technologies_for_GIS.htm#utah

http://environment.fhwa.dot.gov/integ/case_utah.asp

http://gisticinc.com/category/case-studies/

Virginia Department of Transportation (VDOT)

VDOT's Pavement Management Team within the agency's Maintenance Division has developed a Standard Operating Procedures (SOP) document that describes the standard process to be followed for collection, processing, loading, analyzing, and reporting of pavement condition data. This SOP defines specific responsibilities for the Pavement Management Team and Information Technology Division staff to ensure a clear understanding of roles and dependencies.

For more information: Tanveer Chowdhury, VDOT: Tanveer.Chowdhury@vdot.virginia.gov

Washington State Department of Transportation (WSDOT)

Data Catalog. WSDOT created an online "DOTS" (Data Or Term Search) application designed to connect knowledge workers with data and information and promote a common data vocabulary within the agency. It provides information on what data is available, what the data means (meta data), where the data is housed, and who is responsible for managing the data. The application integrates the work of business data stewards, subject matter experts, knowledge workers, application developers, librarians, and technical stewards. The application allows users to search for data resources. In addition, they can subscribe to information on changes in business concepts and attach support documentation or URL references.

For more information: Andy Everett, EveretA@wsdot.wa.gov

Data Delivery. The WSDOT GeoPortal is an application that allows users to view the agency's spatial data via a web browser. Types of data in the GeoPortal include

functional classification, interchange drawings, city limits, and state routes. Users can check boxes to choose from base maps and data layers. The GeoPortal allows users to

- Measure distance or areas
- Share maps via a URL link
- View various city, district, and legislative boundaries
- Select from various imagery, topographical, aerial, and other base maps
- Locate mileposts, geographic coordinates, and street addresses

For more information: http://www.wsdot.wa.gov/mapsdata/tools/geoportal_ext.htm

Resources	The Data Management Association Data Management Body of Knowledge http://www.dama.org/content/body-knowledge
	FHWA Asset Management Data Collection for Supporting Decision Processes http://www.fhwa.dot.gov/asset/dataintegration/if08018/assetmgmt_web.pdf
	Council on Library and Information Resources—Data Curation http://www.clir.org/initiatives-partnerships/data-curation
	Data Curation Profiles Toolkit http://datacurationprofiles.org/
	Data Management Plan Guidance http://www.dcc.ac.uk/resources/how-guides/develop-data-plan and https://purr.purdue.edu/dmp

ELEMENT 3:	Data Architecture and Integration
Improvements	The following strategies can be pursued to establish standard data structures and management approaches to enable improved integration across different data sources: • **Common Geospatial Referencing**—development, adoption, and ongoing maintenance of standard methods for measuring and referencing spatial locations, including locations along linear networks (e.g., based on highway route and distance from known reference point). • **Standardized Approach to Temporal Data—**establishment of standard attributes and common definitions to describe temporal aspects of data sets and allow disparate data sets with a temporal dimension to be integrated. • **Reference Data Management**—ensuring of consistency of standard code lists across applications. • **Master Data Management—**ensuring that the organization maintains a "single version of the truth" with respect to core data entities (e.g., projects and roadways) through centralized management of master data and use of synchronization or replication services. • **Data Architecture Practices and Roles**—maintenance of an integrated agency-wide view of business data entities and their relationships; establishment of review processes for new databases to ensure consistency and appropriate linkages with existing data entities. • **Business Glossaries—**development of agreed-upon definitions of data elements to facilitate integration. • **Data Integration Tools—**use of Extract-Transform-Load (ETL) and other data integration tools to formalize data mappings and automate transformations.
Examples	**Idaho Transportation Department (ITD)** ITD has established procedures for maintaining data in their LRS. The department implemented a commercial product for synchronizing location information across separate systems in place for managing bridge, safety, and traffic data. The system allowed the agency to reduce high mainframe maintenance costs, automate event location stability, and eliminate "data integration by memo." Implementation involved resolving issues related to the management and maintenance of both LRS and event temporal data. To maintain integrity, the department needed to create new standards and data maintenance rules to resolve issues related to temporal queries and temporal event conversion and provide capabilities to correct temporal mistakes. For more information: Brian Emmen, GIS Manager, brian.emmen@itd.idaho.gov **Michigan DOT (MDOT)** Michigan's statewide Geographic Framework is a collaborative, integrated topological model made up of points, lines, and polygraphs all spatially related to one another. There are 21 feature classes (e.g., transportation, boundaries, hydrography, and points of interest) that are topologically maintained with role-based stewardship. There are 160 plus data elements and 1.2 million line segments. The state's entire roadway network is fully linearly referenced, with additional layers for cities, villages, townships, school

districts, legislative boundaries, unincorporated places, census tracks and block groups, and adjusted census urban areas.

The "master" LRS data is made possible by collaborative contrib utions under a shared services model where all agencies contribute to the overall budget and can benefit from results. The system is web-based in an Oracle Spatial Topology Model. The system has an established migration process and there are built-in work flow and audit/approval processes.

For more information: Ron Vibbert, Michigan DOT, VIBBERTR@michigan.gov

New York State Department of Transportation (NYSDOT)

NYSDOT created a data warehouse and implemented new enterprise-wide business intelligence tools to comply with the reporting requirements associated with the American Recovery and Reinvestment Act (ARRA). The agency focused on identifying and aggregating data from 13 existing systems into a data warehouse to meet reporting needs. Efforts were also initiated to develop tools for gatheri ng information from local governments and other partners who do not have access to agency source systems so that they could enter data directly into the warehouse. New enterprise business intelligence visualization tools were developed to use warehouse data for creating new data views, dashboards, and ways of consuming the information.

For more information:
http://www.ctg.albany.edu/publications/journals/dgo_2010_recoveryact/dgo_2010_recoveryact.pdf

Ohio Department of Transportation (ODOT)

ODOT initiated a major research effort to develop a customized, executable, strategic enterprise architecture design for the departments. The research recommended the design of an enterprise architecture, consisting of

- Business architecture—which defines the functional structure of ODOT in terms of its business processes and organization and its associated business information needs
- Applications architecture, which delineates the capabilities of specific applications used to support ODOT's business functions and how these various applications work together or integrate to support ODOT's enterprise-wide information requirements
- Data architecture, which establishes data standards for all of ODOT's systems to support integration and information sharing between these systems
- Technical architecture, which describes the technical infrastructure and software technologies, which are shared services and not application specific, and other specific hardware and operating system-level software technologies required to support the various business applications.

For more information:
www.dot.state.oh.us/Divisions/Planning/SPR/Research/reportsandplans/Reports/2014/Administration/134756_FR.pdf

Oregon DOT

Oregon DOT has drafted a statewide Road Centerline Data Standard to

- Ensure the compatibility of datasets within the same framework feature set and between other framework feature sets and themes
- Help agencies responsible for creating, maintaining, and distributing road centerline data sets by reducing the costs of data sharing, data development, and data maintenance
- Ensure that road centerline attribution (including geometry) is as current, accurate, and complete as possible by relying on local road authorities

The standard describes the essential elements and data structure necessary to adequately describe, produce, and use road centerline data produced in Oregon. This includes a core set of geospatial information and geometry for an accurate and current representation of the state's roadway network.

The data environment for the standard is a vector model, composed of points and linking logical relationships between those points. The exchange medium for road centerline data files is the ESRI shapefile, which is a public domain data structure relating points, lines, and feature attribution (including shape geometry). This exchange medium is supported by all known GIS software suites in use in Oregon. The standard includes specifications for data characteristics, graphic data elements, attributes, resolution and accuracy, and other geospatial data requirements. The standard also includes a data dictionary and glossary of terms.

For more information: http://www.oregon.gov/ODOT/TD/TDATA/gis/docs/TFIT/T-FIT_20061117_TransStandard_Draft_5_0.pdf

Resources	US Federal Highway Administration—ARNOLD Reference Manual: www.fhwa.dot.gov/policyinformation/hpms/documents/arnold_reference_manual_2014.pdf
	FHWA Data Integration Primer: http://www.fhwa.dot.gov/asset/dataintegration/if10019/dip00.cfm
	The Data Warehousing Institute: www.infor@tdwi.org
	Open Methodology MDM Page: http://mike2.openmethodology.org/wiki/Master_Data_Management_Solution_Offering

ELEMENT 4:	**Data Collaboration**
Improvements	Strategies for establishing and supporting data partnerships include • **Multi-Purpose Data Collection**—adopting principles, policies, and business processes for managing data as a strategic agency asset • **Data Clearinghouses/Open Data Platforms**—platforms that enable multiple parties to post data sets that others can discover and use; may include open data API access that enables data to be integrated into apps • **Data Partnerships**—interagency initiatives to collaboratively acquire, build, and maintain data sets of common interest • **Data-Sharing Agreements**– agreements between organizations that establish ground rules for data sharing, including restrictions on use • **Data Outsourcing—**Leveraging available private-sector data sources
Examples	**Massachusetts Department of Transportation (MassDOT)** MassDOT began an Open Data Initiative in 2009, making available some data feeds to the public, and encouraging app developers to use these data feeds to provide the traveling public with useful information. Current data feeds include transit schedules, real-time travel time on selected highways, planned highway construction and maintenance projects, bicycle facilities, and current Registry of Motor Vehicles branch wait times. Real-time travel data is provided through anonymous tracking of Bluetooth-enabled devices carried by motorists and their vehicles. For more information: https://www.massdot.state.ma.us/DevelopersData.aspx **Minneapolis-St. Paul (Twin Cities) Metropolitan Council** The Metro Regional Centerline Collaborative (MRCC) is a joint collaborative project involving managers and GIS staff from the 7-county Minneapolis-St. Paul metropolitan area, the Metropolitan Emergency Services Board, and the Metropolitan Council regional planning agency. The goal of the project is to facilitate the creation of an authoritative roadway centerline database that is locally sourced and maintained and that can meet the business needs of participating agencies. Thus far business needs have been documented and a draft data model has been developed. The MRCC can be used for • Vehicle routing • Address geocoding • Next generation 911 call routing and location validation • Emergency services dispatching • LRS use • Cartographic representation of road features For more information: http://www.metrogis.org/projects/centerlines-initiative.aspx

State of Utah

Spatial data is being shared across agencies and with the public through a Utah state managed Automated Geographic Reference Center (AGRC) clearinghouse. The site includes various address, aerial photography, bioscience, demographic, economy, elevation/terrain, energy, environmental, farming, health, history, planning, recreation, transportation, utilities, and water data layers.

For more information: http://gis.utah.gov/

Utah DOT (UDOT)

UDOT initiated a comprehensive LiDAR data collection effort in 2011 to capture information needed for asset management and other related business needs. The effort was intended to lower overall data collection costs by gathering multiple types of information at the same time. Several different departments participated in funding the data collection, which included pavement condition (roughness, distress, rutting), roadway geometrics, and inventory for several different roadway assets including bridges, walls, signs, signals, barriers, power poles, striping, curb cuts, intersections, drainage, shoulders, and ATMS devices.

Wisconsin Department of Transportation (WisDOT)

WisDOT has developed a Wisconsin Information System for Local Roads (WISLR), an internet-accessible system that local government road authorities can use to report local roadway information on lane/shoulder widths, surface type, surface year, road category and functional classification, and pavement condition ratings to WisDOT. The tool uses GIS technology to combine local road data with interactive mapping functionality. Users can display data in tabular formats, spreadsheets, or maps. Local governments are using the WISLR query, analysis, and spreadsheet tools to analyze, update, and edit their data. WISLR is improving the accuracy of roadway inventory and pavement condition data.

For more information: http://www.dot.state.wi.us/localgov/wislr/

Resources	USDOT Planning for Operations Data Collection and Sharing Resources: http://www.plan4operations.dot.gov/data_coll.htm
	FHWA GIS-T Operating Agreements Page: http://gis.fhwa.dot.gov/gdc_agreements.asp
	FHWA Office of Safety—Noteworthy Practices web page: http://safety.fhwa.dot.gov/rsdp/noteworthy_practices.aspx
	FHWA Research Data Exchange: https://www.its-rde.net/home
	New York State GIS Data Clearinghouse: (https://gis.ny.gov/gisdata/inventories/member.cfm?OrganizationID=539)
	Outsourced Probe Data Symposium Proceedings (January 2015) http://www.ntc.umd.edu/sites/default/files/documents/Publications/Proceedings_1st_Probe_Data_Symposium.pdf
	TRB Special Report—How We Travel: A Sustainable National Program for Travel Data (suggests a collaborative approach to building a National Data Program): http://onlinepubs.trb.org/onlinepubs/sr/sr304.pdf

ELEMENT 5:	Data Quality
Improvements	Data quality improvement strategies include • **Metrics**—establishing and reporting metrics for assessing and describing data quality. • **Data Validation Rules—**establishing business rules for data validity (e.g., acceptable ranges or variations from a prior observation). • **Data Cleansing—**Identification and correction (or exclusion) of data records that do not meet established validity criteria. • **Data Collection Quality Management Processes**—establishing roles, responsibilities, and processes to ensure quality data from field data collection including training, equipment calibration, personnel and equipment certification, comparison against control sections, and independent verification and validation.
Examples	**FHWA Report: Quality Management for Pavement Condition Data** The February, 2013, FHWA Report, *A Practical Guide for Quality Management of Pavement Condition Data Collection*, provides guidance on the principles and practices of data quality management as applied to pavement condition data collection. The report includes suggestions for • Specifying the data collection rating protocols to be used, including those related to tracking linear referencing locations • Establishing quality standards and acceptance criteria • Identifying responsibilities, including training and succession planning of personnel • Performing quality control activities • Monitoring and testing for acceptance • Taking timely and appropriate corrective action • Performing quality management reporting • Developing a data quality management plan The Guide also presents case studies of practices in use by several transportation agencies including • Oklahoma DOT data quality standards • Louisiana DOTD standards for the quality of video image • Pennsylvania DOT data acceptance process and criteria • Nebraska's Department of Roads data collection quality control process that includes calibration of the profiler's laser sensors, accelerometers and DMI, control site testing, real-time system checks, and time-series comparisons Additional information can be found at http://www.fhwa.dot.gov/pavement/management/qm/data_qm_guide.pdf **FHWA Traffic Monitoring Guide** An updated FHWA Traffic Monitoring Guide released in 2013 outlines practices for • Traffic volume monitoring • Vehicle classification data collection

- Truck weighing and data collection at truck weigh sites
- Traffic monitoring data formats

The guide includes integrated sample designs for traffic monitoring and discusses sampling techniques as well as methods and measures for managing variability, enhancing data quality, and developing estimation procedures.

For more information: http://www.fhwa.dot.gov/policyinformation/tmguide/

FHWA Crash Data Improvement Program (CDIP) Guide

The purpose of the CDIP Guide is to "assist state crash database managers and other traffic safety professionals in identifying, defining and measuring the characteristics of the data quality within the state crash database." The foci of the Guide are the quality-related attributes of timeliness, accuracy, completeness, consistency, integration, and accessibility of crash data. The CDIP Guide provides a mechanism by which States can establish baseline measures and subsequent assessments related to the crash data quality characteristics.

For further information:
http://safety.fhwa.dot.gov/cdip/finalrpt04122010/finalrpt04122010.pdf

NHTSA

In 2011, NHTSA issued a report containing model measures for state traffic records systems. The measures covered six key data quality attributes: timeliness, accuracy, completeness, uniformity, integration, and accessibility—across the six core state traffic record data systems—crash, vehicle, driver, roadway, citation/adjudication, and emergency medical services (EMS)/injury surveillance.

Subsequently, NHTSA initiated a series of state Traffic Records Program Assessments where teams of experts would comprehensively review state data, application, organizational, planning, coordination, and investment practices related to traffic data collection, management, and reporting. Recommendations from these NHTSA team assessments addressed data quality-related topics including

- Custodial responsibilities for crash data
- Data timeliness issues
- Data accuracy issues, including a data quality control program with the following components:
    - Data quality metrics for timeliness, accuracy, completeness, consistency, integration, and accessibility
    - Data quality monitoring and reporting
    - Procedures for returning erroneous data and reports
    - Continuous auditing of data quality
    - Periodic reviews by independent auditors
    - Training procedures
    - Feedback mechanisms for reporting performance
- Final acceptance criteria for data submissions

For more information: www.nrd.nhtsa.dot.gov/Pubs/811441.pdf

http://www.nhtsa.gov/Data/Traffic+Records

The Inter-American Development Bank

The Inter-American Development Bank, Department of Infrastructure and Environment, completed an Assessment of Transport Data Availability and Quality in Latin America to identify transport data availability, coverage, and quality within selected developing countries to determine where there are gaps in data needed to estimate greenhouse gas emissions. The assessment established a scale of 1-5 to rate quality and also evaluated

- The availability of time-series data
- Whether the data was subject to quality assurance protocols
- The accessibility of the data
- The entities responsible for data collection
- The cycle for collecting data

For more information: www.iadb.org

MDOT

The Michigan DOT Intermodal Management System (IMS) business processes define data needs and accuracy, completeness, and timeliness requirements. The system includes 54 categories of data assessed quarterly for quality and completeness. Quarterly data quality reports include information on data currency (update due versus actual), known flaws (e.g., missing data), and importance (e.g., use to meet reporting requirements). Data quality categories are assigned as follows on the reports:

- Green—data is complete, correct and capable of supporting business processes
- Yellow—Data is incomplete or incorrect and could pose problems supporting business processes
- Red—Data is incomplete or incorrect and incapable of supporting business processes

For more information: http://mdotcf.state.mi.us/public/tms/idm.cfm

Resources	The Data Management Association: www.dama.org
	International Association for Information and Data Quality: http://iaidq.org/
	FHWA Data Quality White Paper (http://ops.fhwa.dot.gov/publications/fhwahop08038/pdf/dataqual_whitepaper.pdf)
	FHWA Traffic Data Quality Measurement: http://ntl.bts.gov/lib/jpodocs/repts_te/14058.htm
	FHWA Traffic Monitoring Guide—Compendium of Quality Control Criteria: http://www.fhwa.dot.gov/policyinformation/tmguide/tmg_2013/compendium-data-quality.cfm
	FHWA Recommended Framework for a Bridge QA/QC Program: http://www.fhwa.dot.gov/bridge/nbis/nbisframework.cfm

Abbreviations and acronyms used without definitions in TRB publications:

A4A	Airlines for America
AAAE	American Association of Airport Executives
AASHO	American Association of State Highway Officials
AASHTO	American Association of State Highway and Transportation Officials
ACI–NA	Airports Council International–North America
ACRP	Airport Cooperative Research Program
ADA	Americans with Disabilities Act
APTA	American Public Transportation Association
ASCE	American Society of Civil Engineers
ASME	American Society of Mechanical Engineers
ASTM	American Society for Testing and Materials
ATA	American Trucking Associations
CTAA	Community Transportation Association of America
CTBSSP	Commercial Truck and Bus Safety Synthesis Program
DHS	Department of Homeland Security
DOE	Department of Energy
EPA	Environmental Protection Agency
FAA	Federal Aviation Administration
FHWA	Federal Highway Administration
FMCSA	Federal Motor Carrier Safety Administration
FRA	Federal Railroad Administration
FTA	Federal Transit Administration
HMCRP	Hazardous Materials Cooperative Research Program
IEEE	Institute of Electrical and Electronics Engineers
ISTEA	Intermodal Surface Transportation Efficiency Act of 1991
ITE	Institute of Transportation Engineers
MAP-21	Moving Ahead for Progress in the 21st Century Act (2012)
NASA	National Aeronautics and Space Administration
NASAO	National Association of State Aviation Officials
NCFRP	National Cooperative Freight Research Program
NCHRP	National Cooperative Highway Research Program
NHTSA	National Highway Traffic Safety Administration
NTSB	National Transportation Safety Board
PHMSA	Pipeline and Hazardous Materials Safety Administration
RITA	Research and Innovative Technology Administration
SAE	Society of Automotive Engineers
SAFETEA-LU	Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (2005)
TCRP	Transit Cooperative Research Program
TDC	Transit Development Corporation
TEA-21	Transportation Equity Act for the 21st Century (1998)
TRB	Transportation Research Board
TSA	Transportation Security Administration
U.S.DOT	United States Department of Transportation

NON-PROFIT ORG.
U.S. POSTAGE
PAID
COLUMBIA, MD
PERMIT NO. 88

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

The nation turns to the National Academies
of Sciences, Engineering, and Medicine for
independent, objective advice on issues that
affect people's lives worldwide.
www.national-academies.org

ISBN 978-0-309-37485-9

90000

9 780309 374859