

NATIONAL
COOPERATIVE
HIGHWAY
RESEARCH
PROGRAM

NCHRP

Web-Only Document 214:

Transportation Agency Self-Assessment of Data to Support Business Needs: Final Research Report

Spy Pond Partners, LLC
Arlington, MA

In association with

Iteris, Inc.
Santa Ana, CA

Contractor's Research Report for NCHRP Project 08-92
Submitted June 2015

ACKNOWLEDGMENT

This work was sponsored by the American Association of State Highway and Transportation Officials (AASHTO), in cooperation with the Federal Highway Administration, and was conducted in the National Cooperative Highway Research Program (NCHRP), which is administered by the Transportation Research Board (TRB) of the National Academies of Sciences, Engineering, and Medicine.

COPYRIGHT INFORMATION

Authors herein are responsible for the authenticity of their materials and for obtaining written permissions from publishers or persons who own the copyright to any previously published or copyrighted material used herein.

Cooperative Research Programs (CRP) grants permission to reproduce material in this publication for classroom and not-for-profit purposes. Permission is given with the understanding that none of the material will be used to imply TRB, AASHTO, FAA, FHWA, FMCSA, FRA, FTA, Office of the Assistant Secretary for Research and Technology, PHMSA, or TDC endorsement of a particular product, method, or practice. It is expected that those reproducing the material in this document for educational and not-for-profit uses will give appropriate acknowledgment of the source of any reprinted or reproduced material. For other uses of the material, request permission from CRP.

DISCLAIMER

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

The information contained in this document was taken directly from the submission of the author(s). This material has not been edited by TRB.

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, nongovernmental 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.

Table of Contents

Executive Summary	1
1.0 Introduction	3
1.1 Project Overview	3
1.2 Document Overview	5
2.0 Information Gathering	6
2.1 Literature Review Findings	6
2.2 Stakeholder Review Findings.....	8
2.3 Revisions to Assessment Methodology	16
3.0 Case Studies	19
3.1 Case Study State Selection Process	19
3.2 Case Study Preparation.....	19
3.3 Case Study Execution	20
3.4 Summary of Key Findings from the Case Studies	27
4.0 Assessment Guidance and Final Tools	33
4.1 Revisions to Assessment Methodology.....	33
4.2 Final Tools.....	35
4.3 Data Value Assessment Tool	36
4.4 Data Management Assessment Tool	48
4.5 Guide Development.....	60
5.0 Next Steps	61
References	64
Appendix A – Focus Group Results	68
Appendix B – Case Study Results	79
Michigan DOT	79
Utah DOT	99
Synthesis of Gaps and Improvements – Data Value.....	119
Synthesis of Gaps and Improvements – Data Management	120
Appendix C – Summary Presentation	124
Appendix D – Flyer	139

List of Figures

Figure 1. Data Value Tool – Menu Tab.....	37
Figure 2. Data Value Tool – Instructions Tab	38
Figure 3. Data Value Tool – Configuration Tab	39
Figure 4. Data Value Tool – Example Lists Tab	40
Figure 5. Data Value Tool – Importance Tab.....	42
Figure 6. Data Value Tool – Availability Tab.....	43
Figure 7. Data Value Tool – Quality Tab.....	44
Figure 8. Data Value Tool – Usability Tab	45
Figure 9. Data Value Tool – Results Tab.....	47
Figure 10: Data Management Tool – Menu Tab.....	49
Figure 11. Data Management Tool – Instructions Tab	50
Figure 12. Data Management Tool – Configuration Tab	52
Figure 13. Data Management Tool – Glossary Tab	53
Figure 14. Data Management Tool – Data Strategy & Governance Tab.....	54
Figure 15. Data Management Tool – Data Life Cycle Management Tab	55
Figure 16. Data Management Tool – Data Architecture and Integration Tab	56
Figure 17. Data Management Tool – Data Collaboration Tab.....	57
Figure 18. Data Management Tool – Data Quality Tab.....	58
Figure 19. Data Management Tool – Results Tab	59

List of Tables

Table 1: Key Assessment Categories and Topic Areas	3
Table 2. Executive Level Interviews – DOTs and MPOs	9
Table 3. Supplemental Interviews.....	10
Table 4. Focus Group Participation – Data Programs and Functions Represented.....	15
Table 5: Data Value Assessments.....	21
Table A.1. Focus Group Data Improvements	68
Table B.1. Michigan DOT Congestion/Mobility Data Value Activities and Data Sources	79
Table B.2. Michigan DOT Congestion/Mobility Data Value Team Consensus Ratings.....	80

Table B.3. Michigan DOT Congestion/Mobility Data Value Team Gaps and Consequences	81
Table B.4. Michigan DOT Facilities Management Data Value Activities and Data Sources.....	84
Table B.5. Michigan DOT Facilities Data Value Team Consensus Ratings.....	84
Table B.6. Michigan DOT Facilities Data Value Team Gaps and Consequences	85
Table B.7. Michigan DOT Real Estate Data Management Team Assessment Results.....	87
Table B.8. Michigan DOT Facilities Data Management Team Gaps and Consequences	88
Table B.9. Michigan DOT Aeronautics Data Management Team Assessment Results....	91
Table B.10. Michigan DOT Aeronautics Data Management Team Gaps and Consequences	92
Table B.11. Michigan DOT Agency-Wide Data Management Team Assessment Results.....	94
Table B.12. Michigan DOT Agency-Wide Data Management Team Gaps and Consequences.....	96
Table B.13. Utah DOT Maintenance Data Value Activities and Data Sources	99
Table B.14. Utah DOT Maintenance Data Value Team Assessment Results.....	100
Table B.15. Utah DOT Maintenance Data Value Team Gaps and Consequences	101
Table B.16. Utah DOT Project Scoping and Design Data Value Activities and Data Sources	103
Table B.17. Utah DOT Project Scoping and Design Data Value Team Assessment Results	104
Table B.18. Utah DOT Project Scoping and Design Data Value Team Gaps and Consequences.....	104
Table B.19. Utah DOT Asset Management Data Management Team Assessment Results	106
Table B.20. Utah DOT Asset Management Data Management Team Gaps and Consequences.....	108
Table B.21. Utah DOT As-Built Plans Data Management Team Assessment Results	110
Table B.22. Utah DOT As-Built Plan Data Management Team Gaps and Consequences.....	112
Table B.23. Utah DOT Agency-Wide Team Assessment Results.....	114
Table B.24. Utah DOT Agency-Wide Data Management Team Gaps and Consequences.....	116

EXECUTIVE SUMMARY

State Departments of Transportation (DOTs) and other transportation agencies collect and manage large quantities of data. These data are used to enhance internal decision making processes, provide information to the traveling public and meet external reporting obligations. Because data collection and management can be costly, it is important to ensure that dollars invested in data are well spent. This involves not only collecting the right data, but ensuring that the data collected are transformed into meaningful information that can be used for policy making, resource allocation and operational decisions across the organization. Increasingly, organizations are viewing data as an important asset that should be managed to maximize value – similar to pavement, bridges and other infrastructure assets. Asset management processes applied to data include creating a data inventory and establishing policies and processes for managing data throughout its life cycle to meet agency business needs and manage risk.

Building on a framework established by an earlier study (NCHRP Project 8-36, Task 100), NCHRP Project 08-92 developed a methodology and set of tools for assessing a transportation agency's data assets, and then developing a realistic action plan for improvement to both the data assets themselves and ongoing data management processes. The assessment is designed to help agencies address the following key questions:

- Do we have the right data to make good decisions and meet reporting requirements?
- Is our current data good enough?
- Are we getting full value from the data that we have?
- Are we making best use of our data collection and management resources?
- What do we need to improve?

The assessment process is intended to be cyclical and includes three major phases: Prepare, Assess and Improve:

- In the *Prepare* phase, agencies determine which components of the assessment they want to undertake, and which specific business and data subject areas they want to include. They also select and prepare an internal facilitator to conduct the assessments.
- In the *Assess* phase, agencies form teams and apply the assessment tools to understand areas of relative strength and weakness and identify gaps between current capabilities and a target level of capabilities. Two different types of assessments are available—one focusing on assessing data value in the context of a particular business function or policy area, and a second focusing on assessing data management capabilities at either the agency-wide level or for a particular type of data.
- In the *Improve* phase, agencies develop a prioritized plan of data improvement actions including “spot improvements” in response to the Data Value Assessments (e.g. more data, different data, address specific usability issues) and more systematic improvements in response to the Data Management Assessments (e.g. improved governance, technical analysis, processes, skill sets, automation).

In order to develop a meaningful data self-assessment approach for transportation agencies, Project 08-92 conducted interviews with 15 transportation agency executives, and focus groups at five state DOTs. Input from these consultation activities were used to develop an

initial set of assessment tools. These initial tools were tested at two state DOTs (Michigan DOT and Utah DOT). Based on the results of these tests, the tools were refined, and a detailed Guide was prepared that provides a step by step approach for agencies wishing to conduct the data self-assessment. Additional products of NCHRP 08-92 include a set of presentation slides that can be used to brief agency executives and managers on the purpose, benefits and content of the assessment, and a flyer providing an at-a-glance summary of the process.

1.0 INTRODUCTION

1.1 Project Overview

Background

NCHRP Project 08-92 builds upon the work completed through a prior study conducted under NCHRP Project 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 the future development of a tool and/or resource to help transportation agencies assess the adequacy, direction, and management of their data programs”. Table 1, reproduced from the final report for the study, shows the recommended topic areas to be included in the assessment.

NCHRP Project 8-36/Task 100 recommended using a maturity model approach, based on the Capability Maturity Model (CMM) originally developed by the Carnegie Mellon Software Engineering Institute (SEI) for software development. It developed an initial assessment framework, including a data typology, assessment criteria and examples of how a maturity scale might be constructed. It recommended that the assessment framework be further validated and refined, prototyped, and tested; and then potentially expanded into a tool that could be used to assess “a broad range of transportation data program elements.” It was envisioned that the resulting tool “would essentially need to be a ‘living resource’ and evolve as technology, the state of the practice, and agency needs change.”

Table 1: Key Assessment Categories and Topic Areas

Strategic Alignment <ul style="list-style-type: none">• Alignment with strategic goals• Clear and appropriate organizational roles• Alignment with user needs• Identification of data sources, uses, and users• Data utilization and visualization	Data Quality <ul style="list-style-type: none">• Accuracy• Consistency• Reliability• Timeliness• Completeness• Currency• Integrity• Confidentiality
Data Program Management Processes <ul style="list-style-type: none">• Clear definitions• Ability to segregate, aggregate, and analyze• Time and resources for conducting analysis and visualization• Regular audits and validation procedures• Consideration for program tradeoffs, costs and life cycles• Mechanisms for security, privacy, and ethical considerations• Data collaboration• Management continuity	

Source: NCHRP 8-36 Task 100 – “Transportation Data Self-Assessment Guide” Report

Research Objectives

The objectives of NCHRP Project 08-92 were to (1) test the feasibility of the Data Program Self-Assessment process and (2) produce a guide for transportation agencies to implement data self-assessment methods. The guide was to be aimed at helping decision makers and data practitioners at transportation agencies evaluate and improve their data programs to better support policy choices, infrastructure investments, and agency functions.

Research Scope and Tasks

The research was organized into three phases to review the assessment methodology, test and refine it, and develop a final set of tools and accompanying guidance.

Phase 1. Conduct Review of the State of the Practice

Task 1. Literature Review. The literature review focused on references that provided specific material on (1) data assessment criteria and methods and (2) data management practices suitable for incorporation or adaptation into the guide.

Task 2. Develop Interview Guide. In the task, materials were prepared to summarize the Project 08-36/Task 100 assessment approach and an interview guide was developed for vetting it with state DOTs and Metropolitan Planning Organizations (MPOs).

Task 3. Review Methodology with Stakeholders. Interviews were conducted with senior managers at state DOTs and large MPOs, and through focus group meetings with state DOT staff.

Task 4. Revise Assessment Methodology. Using data from Task 3, a revised methodology was developed for the data self-assessment.

Task 5. Update Phase 2 Work Plan. On completion of the revised assessment methodology, the Phase 2 work plan was updated.

Task 6. Phase 1 Interim Report. An interim report was prepared in March 2014 that provided: (a) a synopsis of the results of the review; (b) results of interviews; (c) a revised methodology for conducting data program self-assessment; (d) an initial draft outline of the guide and (e) an updated work plan for Phase 2, including identification of potential case studies.

Phase 2. Case Studies

Task 7. Finalize Case Study Agency Selection. Two case study agencies were selected and an information packet for the candidate agencies was prepared including: (1) a summary of the methodology to be tested, (2) description of who in the agency would need to be involved and how much time would be required, and (3) the project timeline.

Task 8. Conduct Case Studies. The case studies were conducted with working groups, within each agency. Results were summarized along with a series of lessons learned for refining the assessment tools and preparing the final guide.

Task 9. Phase 2 Interim Report. A second Interim Report was prepared in October 2014 documenting the results of Phase II including: (a) the process used to conduct the case studies; (b) the results of the case studies; and (c) lessons learned about the feasibility of the data self-assessment methods (e.g., DOT staff time used; meaningfulness of results; applicability of the results in terms of actionable items).

Phase 3. Guidance Development

Task 10. Draft Guide. Based on the results of Phase 2, a draft guide was prepared along with a set of accompanying stand-alone spreadsheet tools for the self-assessment process.

Task 11. Draft Final Report. A draft final report was prepared to document the entire project and summarizes the products of the research.

Task 12. Final Report. A final report was prepared based on comments from the panel, along with an accompanying Microsoft PowerPoint® presentation describing the project background, objective, research method, findings, and conclusions. A final Guide was produced as a separate deliverable of this task.

1.2 Document Overview

This is the Draft Final Report for NCHRP Project 08-92 -- “Implementing a Transportation Agency Data Self-Assessment”. It summarizes the objectives, scope, methodology and deliverables of the project. This first chapter has provided an overview of the project. Chapter 2 documents the information gathering phases of the project including the literature review, focus groups and interviews with state Department of Transportation (DOT) and Metropolitan Planning Organization (MPO) executives. Chapter 3 covers the work that was done to pilot the research in two state DOTs. Chapter 4 provides a brief summary of the products from the research and an overview of the Self-Assessment Guide that was produced to assist transportation agencies in conducting transportation data self-assessments. Chapter 5 outlines next steps for implementation of the self-assessment process in transportation agencies.

Appendix A provides detailed findings from the focus groups.

Appendix B provides detailed findings from the case studies.

Appendix C includes presentation slides that can be used for communicating the purpose and value of the self-assessment process.

Appendix D includes a flyer that can be used for outreach and implementation activities.

Volume 2 of this report includes the Data Self-Assessment Guide.

2.0 INFORMATION GATHERING

2.1 Literature Review Findings

Methodology

The first project task included an in-depth review of several data assessment models, and covered a balanced set of resources from each of the major dimensions identified in the earlier NCHRP 8-36, Task 100 research framework related to:

- Strategic Data Alignment
- Data Program Administration
- Data Quality

The research team also conducted a review of selected references providing guidance on several specific data programs related to travel, safety, system inventory, system conditions and financial data.

Summary of Findings

Several data assessment models (and related resources) were identified in the state of the practice review:

- NCHRP 8-36 Task 100 (transportation data) [1]
- USDOT Federal Highway Administration (FHWA) Roadway Safety Data Capabilities Assessment [2]
- CMMI Data Management Maturity [3] (general)
- IBM Data Governance Council (general)[4]
- Data Flux Data Governance Maturity Model (general) [5]
- Gartner Enterprise Information Management Maturity (general) [5]
- Aikin – Community Self- Assessment (general) [6]
- Transportation Asset Management Guide Volume I Self-Assessment [7]
- Transportation Asset Management Guide Volume 2 Gap Analysis (for asset inventory and performance, planned and completed work) [8]
- NCHRP 666 Data Management Maturity Model (transportation performance management) [9]
- Information Quality Benchmarks (health care)[10]
- Transportation Systems Management and Operations Capability Assessment [11]
- Assessment and Accountability Comprehensive Center (AACC) Data Needs Assessment (for education) [12]

Of those reviewed, the following models were particularly helpful in illustrating examples of a robust assessment methodology:

- The FHWA *Roadway Safety Data Capabilities Assessment* examines data needed for different analysis and decision making steps, and uses a set of elements that easily lend themselves to development of an action plan (people, policies, technology, etc.). Summary displays of results are very effective for communicating “at a glance” maturity levels for a single agency, as well as maturity levels across multiple agencies. Also, there is a strong linkage between assessment findings and development of an action plan for improvement.
- The *Transportation Asset Management Guide* self-assessment provided information on the strategic alignment of data programs and included tools for helping agencies assess whether they have the information needed for data-driven investment and resource allocation decisions.
- An early draft of a joint Carnegie Mellon Software SEI and the Enterprise Data Management Council (EDM) data management maturity model provided a useful list of practices and maturity level definitions.

Collectively, the examples included in the literature review provided a rich resource base for project 08-92, with respect to how to structure assessment questionnaires; key substantive elements of data alignment, management and quality; methods and formats for summarizing and visualizing results; and approaches for linking results to strategies and actions.

Other Resources

Several other sources were reviewed that provided useful content for development of the data self-assessment methods.

- **Strategic Alignment.** In addition to the *Roadway Safety Data Partnership* work and the *Asset Management Guide* materials discussed above, there have been several FHWA and NCHRP studies and peer exchanges on the topics of data for decision making, and how to integrate consideration of data needs and data management practice into performance-based planning and resource allocation processes [13, 14, 15, 16, 17]. In addition, several examples of state DOT data business planning were reviewed that included processes for identifying and prioritizing data needs [18, 19, 20, 21]. The Transportation Research Board’s Special Report 304, completed in 2011 presents a comprehensive look at sustainable approaches to providing passenger and freight travel data to support public and private sector transportation planning and decision making [22]. A 2010 American Association of State Highway and Transportation Officials (AASHTO) transportation finance briefing paper [23] presents examples of how different agencies have approached data compilation and management to support performance management. Finally, a 2012 FHWA-sponsored report presented a useful benefit-cost analysis methodology for evaluating potential safety data investments [24].
- **Data Program Administration.** The Data Management Association’s (DAMA) Data Management Body of Knowledge (DMBOK)[25] presents a comprehensive treatment of data management functions, terminology and best practices. A white paper presenting alternative models of data stewardship [26] provided insights into different ways to approach roles and responsibilities for data management. Within the transportation

domain, several references provide data definitions and collection standards – including The Highway Performance Monitoring System (HPMS)[27], the Model Minimum Uniform Crash Criteria (MMUCC) [28], and the Model Inventory of Roadway Elements (MIRE)[29, 30]. A 2012 FHWA study examined data management strategies in the context of open data programs [31].

- **Data Quality.** Several useful references were identified for specific data areas (crash data [32, 33], pavement [34], bridge [35], traffic [36, 37, 38], operations [39]) that cover key elements of data quality, and how to assess data quality in the context of specific applications.

2.2 Stakeholder Review Findings

Following completion of the literature review, the research team reviewed the methodology developed in the Task 100 Report with key stakeholders. The objective was to identify any needed changes to the conceptual design, and incorporate that input into the revised assessment tools and guidance. The review process included interviews with 12 executive level individuals at state DOTs and MPOs, and five state DOT focus groups. Supplemental interviews were also conducted with additional individuals at AASHTO and FHWA. Table 2 lists the executive interviews conducted. Table 3 lists the supplemental agency interviews conducted. The focus group agencies were Oregon DOT, Minnesota DOT, Kentucky Transportation Cabinet, Maryland State Highway Administration and Colorado DOT.

Table 2. Executive Level Interviews – DOTs and MPOs

Agency	Name	Title
Arizona DOT	John Halikowski	Director
Colorado DOT	Heidi Humphreys (with William Johnson)	Director of Human Resources and Administration
Kentucky Transportation Commission	Mike Hancock	Secretary
Maryland State Highway Administration	Gregory Slater Gregory Welker	Director of Planning and Preliminary Engineering Deputy Administrator/Chief Engineer for Operations
Michigan DOT	Kirk Steudle	Transportation Director
Minnesota DOT	Tim Henkel	Division Director, Modal Planning and Program Management
NYSDOT	Roderich Seachrist	Assistant Commissioner, Operations and Asset Management

Agency	Name	Title
Oregon DOT	Jerri Bohard	Transportation Development Division Administrator
Utah DOT	Carlos Braceros	Executive Director
Utah DOT	Cory Pope	Program Development Director
Virginia DOT	Greg Whirley	Commissioner
Minneapolis-St. Paul Metropolitan Council	Arlene McCarthy	Transportation Services Director
Puget Sound Regional Council	Charlie Howard	Director of Integrated Planning
Washington DC Council of Governments	Ronald Kirby	Director of Transportation Planning

Table 3. Supplemental Interviews

Agency	Name	Title
AASHTO	Matt Hardy	Program Director for Planning and Policy
FHWA – Office of Highway Policy Information	David Winter	Director, Office of Highway Policy Information
	Ralph Davis	Chief, Motor Fuel and Driver License, and Vehicle Registration Division
	Chris Allen	Division Leader, Recovery Act and Highway Finance Division
	Tinjia Tang	Chief, Travel Monitoring and Travel Surveys
FHWA – Office of Safety	Heather Rothenberg	Transportation Specialist
FHWA – Office of Safety	Bob Pollack	Safety Data Manager
FHWA – Office of Transportation Performance Management	Pete Stephanos	Director, Office of Pavement Technology
FHWA – Office of Transportation Policy Studies	Mary Lynn Tischer	Director, Office of Transportation Policy Studies

Agency	Name	Title
VHB – Contractor for the Safety Data Capabilities Assessment	Mike Sawyer	State Safety Data Capabilities Assessment Project PI

Synthesis of Interview Findings

Executive Interviews

Findings from the executive interviews are divided into two categories: (1) aspects of data program maturity that an assessment should measure and (2) approach to the assessment process itself.

Aspects of Data Program Maturity of Interest to Executives

The executives who were interviewed consistently identified the need for better translation of data into meaningful information that could be applied for policy making, resource allocation and operational decisions across the organization. Multiple barriers to this process were cited, including silo-based information systems that were not integrated, as well as availability of staff with the time and expertise needed to integrate and analyze data and present it effectively. A second common concern was related to the overall efficiency of data programs, i.e. the extent to which available data from internal and external sources were being fully leveraged to avoid duplication and minimize collection and management costs.

Specific aspects of data program maturity of interest to executives are synthesized below.

Data Value for Decision Making

- Extent to which the organization can *expeditiously integrate and leverage* their available data sources in order to answer *key questions of interest* from a decision making standpoint – For example:
 - Providing a bird’s eye view of the agency’s performance, activities and resources;
 - Understanding the interrelationships across asset inventory, maintenance, design, capital program, project management and financial information;
 - Understanding how system performance is being impacted by projects and programs;
 - Understanding impacts of policy changes and resource allocation decisions;
 - Understanding costs and impacts of operations activities in order to make future improvements (e.g. determining when to use new materials for pavement markings); and
 - Understanding mobility and safety from a multi-modal and intermodal perspective.
- Extent to which there is a consistent, repeatable and transparent process for making *data-driven decisions*.

- Extent to which the organization is able to *capture full value from all data programs*, for example, by building predictive capabilities and using data to drive decisions.
- Extent to which data programs *inform policy development and evaluation* – which may involve highly specialized data sources or case studies.
- Extent to which the organization and its partners can *adapt existing data systems to respond to new policy questions* that emerge; agility to obtain new summaries or reports or add new variables of interest. For example, recent interest in the topic of distracted driving has created the need to re-tool existing accident reports and the systems that process, store and report the data.
- Extent to which the organization has *aligned their data collection activities with the intended uses* for the data and the value that the data will provide – and defined minimum/standard data elements for different areas (e.g. construction projects, asset inventory) that are sustainable given available resources.

Data Packaging, Presentation and Delivery

- Extent to which the agency is able to effectively develop and present information with the context and interpretation needed for effective application (e.g. presentation of trend information highlighted with explanatory factors).
- Extent to which *casual users* (at both executive and operational levels) can *discover and use* information of interest from data systems – e.g. via Geographic Information System (GIS) interfaces, dashboards or other business intelligence environments that offer “pre-digested” results.
- Extent to which information users are able to get a good *understanding of what is “behind the numbers”* – particularly when there are multiple handoffs in development of analysis results.

Efficiency and Integration

- Extent to which the organization is *using its resources for data collection and management efficiently* – e.g. taking advantage of GIS, web, and mobile technologies, utilizing sampling methods, tapping in to available external data sources, developing partnerships, ensuring that data collected to meet one purpose (e.g. federal reporting) can be used for other purposes, aligning data sources used to address executive, technical and customer needs for a given topic; considering potential regional/MPO geographic aggregation requirements for state DOT-collected data.
- Extent to which the organization has *upgraded its technology infrastructure* to enable its data programs to operate smoothly – e.g. reduce reliance on hard-copy plans and documents, retire inflexible mainframe-based systems, replace spreadsheets with enterprise management systems, build automated interfaces, etc.
- Extent to which data can be *integrated or exchanged across source systems* – to enable analysis and eliminate duplication. There may be business, technology, architecture, and cross-vendor software compatibility issues to address.

Data Quality

- Extent to which specific *gaps in data quality* (coverage, timeliness, accuracy, etc.) have been *identified and assessed* based on risk or opportunity cost. These gaps are often nuanced and can include, for example: variations in quality by source, sub-network, jurisdiction, age of infrastructure. Solutions to these quality issues also vary – some may be relatively straightforward (e.g. more quality assurance); others may be very difficult (e.g. achieving consistency across a large number of data providers.) Where substantial investment/effort is required for data quality improvement, agencies must decide if this effort is worth it.
- Extent to which *people responsible for collecting data understand the value and uses* of the data – this can have a dramatic positive impact on data quality.

Assessment Process Characteristics

Executives interviewed felt that there would be value to having a high level summary of how well the agency's data programs were working, and an ability to track the status of improvements. There was a high degree of consensus that in order to be meaningful, the assessment should be organized around major agency policy or program areas rather than around specific types of data. This organization better allows for consideration of how well data are supporting decision making.

Executives also talked about distinguishing questions about whether the agency was *collecting the right data* from questions about whether the agency was *making effective use of existing data*. Another common theme related to the need for guidance to help agencies focus on data improvement initiatives that are realistic and have the highest payoff. An assessment process that simply provides a laundry list of the areas where they might improve is not helpful. A number of specific comments were also offered to strengthen the assessment process.

Supplemental Interviews

AASHTO and FHWA

Supplemental interviews conducted with representatives of AASHTO and FHWA echoed several of the themes that were identified in the executive interviews. They felt that a data self-assessment would be of value to raise awareness of tradeoffs between data cost, quality, timeliness and access. In addition, the process could help improve data literacy – for example, understanding and appropriate application of sampling and confidence levels. They also suggested that the assessment be kept simple in order to encourage its application.

Additional specific suggestions were:

- With respect to assessment of data quality, it is important to recognize that (1) perceptions of data quality may vary across different business units, (2) data quality improves with the level and frequency of data use and (3) establishing accountability for data quality is essential.

- With respect to assessment of data management and administration, a mature approach is one where there is a reasonable level of staffing and resources devoted to data programs, and the establishment of well-defined roles.

AASHTO and FHWA staff emphasized the need for this project to be cognizant of ongoing national transportation data-related efforts, including:

- The data principles adopted by AASHTO Standing Committee on Planning
- FHWA data programs that depend on consistent data to be reported from states – HPMS, Traffic Monitoring, Motor Fuel, Licensed Drivers, Vehicle Registration, Financial.
- MAP-21 requirements for performance target setting and reporting related to pavement and bridge condition, system performance, freight, safety and Congestion Mitigation and Air Quality (CMAQ).

State Safety Data Capabilities Assessment Representatives

Interviews with FHWA and contractor representatives of the State Safety Data Capabilities Assessment initiative identified successful elements of that effort that are relevant for the development of the data self-assessment approach in NCHRP 08-92. These included:

- identifying a champion and liaison person in each state to facilitate the effort;
- using a structured survey instrument to gather assessment information;
- using branching logic to skip elements that are not relevant based on initial answers and
- using follow-up roundtable discussions to review and validate assessment ratings.

They recommended use of a flexible approach that allows states to provide information in a way that was easiest for them; development of a glossary of terms to facilitate communication across different stakeholders; and a process that assures anonymity and confidentiality. They urged the research team to avoid the assumption that “one size fits all” and be conscious of the level of effort required to complete the assessment. They suggested that the process incorporate regular follow-up on the action plan in order to track progress and avoid loss of momentum.

These interviews also explored the relationship between the detailed safety data assessments that have been completed and the self-assessment being designed for NCHRP 08-92. The tools designed for NCHRP 08-92 are sufficiently different from those used for the safety efforts. The research team envisions that states could use the NCHRP 08-92 tools for a higher-level assessment of their data programs and rely on the existing safety assessment tools for more detailed assessments and diagnostics. Alternatively, they could make use of the completed safety assessments to inform the higher-level assessment developed under NCHRP 08-92.

Focus Groups

The five state DOT focus groups provided an opportunity for data program managers and staff, as well as mid-level managers to provide input on the assessment process.

Table 4 presents the data programs represented in the focus groups. Overall, there was broad representation across traffic, safety/crash, road inventory, GIS, capital programs/financial, ITS, planning, maintenance, pavement and bridge, and asset management/performance management functions.

Table 4. Focus Group Participation – Data Programs and Functions Represented

Data Programs/Functions	CO	MD	KY	MN	OR
GIS	✓✓	✓✓	✓✓		✓✓
Roadway Inventory/ HPMS		✓✓		✓✓	✓✓
Traffic Monitoring	✓✓	✓✓		✓✓	✓✓
Traffic Operations/ ITS	✓✓	✓✓	✓✓	✓✓	✓✓
Safety/ Crash Data	✓✓		✓✓	✓✓	
Performance/ Asset Management	✓✓			✓✓	✓✓
Bridge	✓✓			✓✓	
Pavement	✓✓		✓✓		
Maintenance			✓✓	✓✓	
Planning/ Travel Modeling/ Freight	✓✓	✓✓	✓✓	✓✓	✓✓
Bike/Pedestrian				✓✓	
Financial				✓✓	✓✓
Capital Program Management		✓✓			
District Project Development			✓✓		
Motor Carrier (IFTA/IRP)			✓✓		
Data Services/ Data Management	✓✓	✓✓	✓✓	✓✓	✓✓

Data Program Costs and Desired Data Improvements

Focus group participants were asked to share information on their roles, estimated data program costs and their personal “wish lists” for improvement. The following definitions were described for various data-related roles:

- Data Owner(s) – people with decision making authority for initiation (or discontinuance) of the data program and who determine the content of what data are collected
- Data Steward(s) – people who are accountable for the quality, value and appropriate use of the data
- Data Manager(s) or Custodian(s) – people with hands-on responsibility for data loading, validation, extraction, transformation and integration tasks.

- Data Analyst(s) – people with hands-on responsibility for reviewing, reporting, visualizing, charting, summarizing, synthesizing and highlighting findings using one or more data sets produced by the program.
- Data Users – both hands-on users of the data and indirect users that receive information that draws on data provided by the program.

Discussion revealed that many were not generally familiar with distinctions among these different roles. Many of the participants stated that they played multiple roles – including steward, manager and user. This reflects the largely siloed nature of data programs in DOTs.

Many of the participants were able to estimate the cost for certain elements of their data programs – for example, data collection contracts, salaries for staff playing significant data management roles, and equipment costs. However, only Oregon DOT appeared to have a relatively formal approach to tracking data program costs.

A majority of the items on focus group participants' wish lists related to either getting new data, improving decision support using existing data, or obtaining a better understanding of how data are being used. There was not much discussion of data quality improvements, though a few participants identified needs for improved timeliness or enhanced quality assurance activities. Several participants commented on the need for upgraded technology in order to improve access and integration. Participants with data management responsibilities also cited the need for improved, automated data integration and formalized governance processes for updates.

Sample Questionnaires

Focus group participants were also asked to complete sample questionnaires for each of the three assessment dimensions, and then comment on the questions that were included. Participants offered a number of specific comments that were helpful in refining the strategic alignment, data quality and data program administration and management dimensions of the assessment approach. These are incorporated in the summary of general observations and revisions to the assessment methodology that are outlined below. Additional detail on the focus group results can be found in Appendix A.

2.3 Revisions to Assessment Methodology

Summary of General Observations

In general, the interviews and focus groups validated the conceptual framework for data self-assessment presented in the final report for Project 08-36 Task 100. Transportation agency executives felt that a high level data assessment would be of value, and the types of issues and assessment factors discussed in the Task 100 report resonated with the DOT focus group participants. The overall three step process of *Prepare, Assess and Improve* was readily understood and accepted as logical.

However, the interviews and focus groups did suggest the need for some refinements to the methodology. Key findings and associated conclusions for the assessment methodology are summarized below:

- **Assessment Categories.** Distinctions across the major assessment categories (Alignment, Data Quality and Program Administration/Management) were not clear-cut due to the inherent interrelationships across these categories – e.g. a poor quality data set may reflect inadequate data management and stewardship processes and results in a lack of alignment with user needs. One person might therefore categorize a data currency problem under Data Quality whereas another might categorize this same problem under Administration/Management. In general, focus group participants asked for greater clarity about whether the assessment was dealing with the data itself, or the processes used to manage the data. *Conclusion: The assessment needs to treat availability and usability of data from an end-user perspective separately from data administration, integration and quality assurance processes from an organizational or data program management perspective.*
- **Assessment Frame of Reference.** DOT executives and many focus group participants expressed a strong preference for assessment results based on substantive policy or business areas rather than at the data program level. This approach makes sense for questions about alignment between data and business needs. However, in considering specific aspects of data management and data quality, an enterprise-wide or data program-based approach is more appropriate. *Conclusion: Assessment elements related to alignment should be evaluated using a policy or functional area frame of reference; others should be evaluated at the data program or the enterprise level.*
- **Business Unit or Agency-Wide Focus.** Much of the discussion about data maturity levels emphasized that the real challenge being faced by transportation agencies is managing data *across* rather than *within* organizational units. *Conclusions: (1) Maturity levels for at least some assessment elements should be defined based on the degree to which an agency-wide approach is taken. (2) The guidance for the Prepare step should encourage agencies to scope the assessment to allow for consideration of cross-unit data coordination and sharing, even when the scope is limited to a particular policy area or set of data programs.*
- **Maturity Levels.** Finally, there were questions about whether maturity levels should be based on the *processes and practices* in place, or whether they should also consider the specific *type, extent and quality of data* available. Typically maturity models emphasize processes and practices – and maturity level definitions reflect the level of standardization and management controls in place. However, the data assessment would be more meaningful if it included substantive criteria that went beyond processes. *Conclusions: (1) The methodology should use different maturity scales for assessment of data management processes versus assessment of data availability, extent and quality. (2) Assessment of data availability, extent and quality should build on the safety and asset management assessment tools, as well as other available standards and models, and provide flexibility for future expansion.*

Revised Assessment Structure

The findings of the interviews and focus groups led the research team to conclude that while the basic elements of the process outlined in the Task 100 report were sound, a restructuring of the assessment portion was warranted. The main change to the structure is to split the assessment into two distinct parts: a first focusing on data value to the organization, and a second focusing on data management processes:

- The *Data Value Assessment* is conducted for one or more policy or functional areas and reflects the data user and data analyst perspectives. This first assessment looks at the availability, quality and usability of data to meet business needs. It does not consider how data are managed but rather how well data are working for the business.
- The *Data Management Assessment* is conducted for one or more data program areas – or at the agency-wide level. This assessment is concerned with process – i.e. how data are managed – and reflects data owner, steward and manager perspectives.

Splitting the assessment was not an easy decision since it inevitably would lead to greater complexity. It implies that two different types of assessment tools are needed. However, the interviews and focus groups clearly pointed the research team in this direction. On balance, the research team concluded that the additional complexity was justified and could be mitigated by designing a flexible approach in which agencies could choose to select one or both types of assessments to conduct. Therefore, the research team recommended moving forward with developing and testing two different assessment tools. The results of the testing process are summarized in Chapter 3.

3.0 CASE STUDIES

Case studies of tool application in two states were designed to test the data self-assessment approach and obtain feedback on the methodology and tools. This section summarizes activities to select and carry out the two case studies.

3.1 Case Study State Selection Process

The research team identified the following criteria for case study agency selection:

- Leadership support for participation in this effort
- Ability to designate a liaison for the effort who can make available the time required to work with the research team and orchestrate the assessment activities
- Ability to involve a range of staff members and ensure that they have the time to complete assessment questionnaires and provide feedback to the research team
- Willingness to have the case study results published
- Past or current involvement in data program assessment and/or implementation of data governance programs
- Variation across agency data programs in level of maturity
- Experience with both rural and urban data programs, and with both traditional and real-time data programs
- Experience with outsourcing data program elements and/or using data from surrogate sources

Based on discussion with the research panel, the candidate list was narrowed down to five states, with a preference for selecting two states in different geographic regions. The research team made initial contacts with the states of Michigan and Utah to determine their level of interest and willingness to participate. These states were selected from the list of five because of strong executive-level support for data management initiatives. Both states indicated a willingness to participate.

3.2 Case Study Preparation

Develop Draft Assessment Tools

Three separate draft data self- assessment tools were developed based on the results of the literature review, focus groups and interviews with stakeholders. These included: (1) a spreadsheet tool for the Data Value Assessment, to be conducted for business or policy areas rather than for specific data programs; (2) an online survey for assessment of a specific data program, and (3) an online survey for assessment of agency-wide data management functions. While the original intent was to have a single tool for the Data Management Assessment, a decision was made to split this into two different tools because many of the questions about agency-wide or enterprise data management wouldn't be relevant for managers and staff within specific data programs.

Assessment Planning and Set-Up

In order to provide a good set of test cases, a plan for testing the Data Management Assessment called for conducting the Data Value Assessment for two different business areas in each agency, and the Data Management Assessment for two different data programs as well as for the agency-wide data management function.

The three phase assessment process was tested as follows:

- **Prepare:** the research team worked with each agency by phone to scope the assessment, and identify a planning team and participants for each of the data value and data program management assessments.
- **Assess:** assessment questionnaires were sent out for each participant to complete, results were compiled, and on-site workshops were held with each group to discuss results and develop consensus results and discuss gaps.
- **Improve:** a second set of workshops was held with each group to review gaps and identify priority actions for improvement.

Briefing materials were prepared for the two participating agencies that summarized the purpose of the assessment, the steps required by the agency to participate, the planned timeline of activities, and an overview of the assessment methodology. Each state scheduled a kickoff meeting with the planning team, and two sets of workshops for each assessment group.

3.3 Case Study Execution

3.3.1. Data Value Assessment Workshops

The research team conducted two Data Value Assessment workshops in each state. The two areas selected for the assessment by the Michigan were Mobility/Congestion and Facilities Management. At Utah DOT, the two areas selected for assessment were Maintenance and Project Scoping & Design.

Questionnaire Configuration

Prior to the first workshop, conference calls were held to configure the data value tools for the selected business areas. Configuration of the data value tools involved (1) identifying key data sources used for the business area and (2) reviewing the generic categories of data-consuming processes and substituting more specific activities that would be recognized and understood by the participating staff. The business areas, activities and data sets selected for each of the Data Value Assessments in the pilot are shown in Table 5.

Table 5. Data Value Assessments

Agency	Business Activities	Data Sets
<p>Michigan DOT – Mobility/ Congestion</p>	<p>Transportation system performance monitoring</p> <p>Scoping and design of candidate projects</p> <p>Corridor and long-range planning, multi-modal planning</p> <p>Real-time traffic and incident management</p> <p>Improvement program development/prioritization</p>	<p>Road Inventory – LRS and road characteristics</p> <p>Traffic Monitoring – counts, AADT, classification</p> <p>Bike/Pedestrian – bike routes, paths, counts</p> <p>Real Time Travel – Current volume/occupancy, travel time</p> <p>Capital Program – Construction project scope and status</p>
<p>Michigan DOT – Facilities Management</p>	<p>Track facility inventory and condition</p> <p>Track facility capital and maintenance expenditures and work accomplishment</p> <p>Identify candidate projects for rehabilitation, replacement and expansion/ addition</p> <p>Diagnose causes of high maintenance costs or inefficiencies</p> <p>Prioritize candidate projects and develop resource-constrained improvement program</p>	<p>Facility Inventory</p> <p>Facility Inspection/Condition</p> <p>Maintenance Records</p> <p>Facility Improvement Program</p> <p>Budgets and Expenditures</p>

Agency	Business Activities	Data Sets
Utah DOT – Maintenance	Track maintenance level of service (outcomes) 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 - Highway Referencing System Maintenance Feature Inventory Maintenance Work Orders Funding allocations Snow Plans Plan for Every Section (pavement management system recommendations)
Utah DOT – Project Scoping and Design	Project Management/Project Control (tracking with Executive Dashboard) Prepare Design Plans Environmental Review Review Existing Conditions/ Identify Needs (Using As-Builts) Create Concept Reports	Roadway Inventory (LRS and Road Geometry) Traffic Monitoring (Counts, AADT, Classification) Asset Inventory, Concept Reports, Environmental Data Crash and Fatality Data Contract Management, Project Management Right of Way (ROW) Executive Dashboard ProjectWise (As-Builts)

Stakeholders for each group were identified by the agency lead, and a three hour workshop was scheduled. Questionnaires were sent to each identified stakeholder the week prior to the workshop. The questionnaires included a series of question designed to assess data availability, data quality and data usability. A 4 point scale (Excellent, Good, Fair, Poor) was used to rate individual assessment elements.

The questionnaire also provided opportunities for team members to rate the importance of the data to meeting business needs. This helped ensure that subsequent improvement planning focused on those gaps that are most important. Questionnaire results were compiled in advance of the workshop and emailed to each participant.

Data Value Assessment Workshop

Questionnaire results and the range of ratings received from participants were reviewed and discussed at the first workshop. The goal at the workshop was to develop consensus ratings for each of the assessment elements and begin thinking about gaps and issues. Where ratings were less than desired, team members were asked to identify specific gaps related to the following three types of gaps:

- **Data availability gaps** – data are not available at the right level of detail, with sufficient coverage.
- **Data quality gaps** – data are not sufficiently accurate, credible and current to support decision making.
- **Decision support gaps** – data are not integrated, analyzed and presented as needed to support decision making or too much manual effort is required to package and deliver meaningful information.

The consequences and impacts of gaps were discussed and team members shared information on current initiatives underway to strengthen their data programs.

Improvement Planning Workshop

A second workshop was held to review and validate gaps identified at the assessment workshop, and to identify priority actions for closing the gaps. The workshop focused on identifying a set of priority actions to move forward. Each participant was asked to suggest 1-2 actions that they felt were most important to pursue in order to address one or more of the gaps.

Actions identified had to be considered feasible to implement and had to have a clear business case. For each action, the team was asked to identify any related gap(s), whether it was a current initiative or something new, the estimated timeframe for completion, and the person or unit who would be responsible for moving it forward. They were also asked to characterize the level of resources that would be needed for implementation. Teams were then asked to rank or sequence the different actions.

3.3.2 Data Management Assessment Workshops

A similar process was used to conduct the Data Management Assessments in each state. At Utah DOT, the two data programs selected for assessment were Asset Management Data and As-Built Plans. Michigan DOT selected Aeronautics Data and Real Estate Data. Both agencies also convened an agency-wide Data Management Assessment group. A link to an online questionnaire was sent to each identified stakeholder in advance, and then a workshop was held with each group to develop consensus ratings and identify gaps.

Questionnaire Configuration

The data management questionnaire included a series of questions designed to assess the maturity of the following three components:

- **Data Governance**, including roles and responsibilities, ensuring alignment with agency priorities, processes for managing user expectations, ensuring that data sets are documented and discoverable, data archiving, and processes to ensure data security and integrity.
- **Data Architecture and Integration**, including establishment of linkages across key data sets, setting and enforcing data standards to ensure consistency and enable integration, managing data in a manner that allows it to be presented and accessed in multiple forms, and managing changes to data structures and code lists to minimize impacts and ensure that multiple existing systems are kept in synch.
- **Data Quality Management**, including general quality assurance processes in place, as well as specific processes for establishing standards, measuring, monitoring, communicating and improving levels of accuracy, reliability, completeness, timeliness and currency.

A 5 level maturity scale was utilized, with definitions as follows:

- **1-Initial:** Processes, strategies and tools are generally ad-hoc rather than proactive or enterprise-wide; successes are due to individual efforts.
- **2-Aware:** Widespread awareness of more mature data management practices; recognition of the need to improve processes, strategies and tools, planning for improvements is underway.
- **3-Defined:** Processes, strategies and tools have been developed, agreed-upon and documented. Initial pilot efforts may be underway to implement improvements.
- **4-Proficient:** Processes, strategies and tools are generally being implemented as defined.
- **5-Optimizing:** Processes, strategies and tools are routinely evaluated and improved.

Questions were tailored based on the perspective of the respondent and whether the assessment element is the agency-wide data management function or a specific data program. A summary explanation of the purpose of the assessment was included.

In addition to completing assessment questionnaires, respondents were asked to provide comments on their answers, and provide insights into why the organization has not moved further along the maturity scale, and what steps would be required to build a more mature data management function.

Data Management Assessment Workshop

Questionnaire results and the range of ratings that were received from participants were reviewed and discussed at the workshop. Each question was mapped to an assessment element and maturity level. Responses were discussed, and consensus maturity levels were assigned for the different assessment elements based on the discussion. In some instances multiple levels were assigned to a given element due to variations in current practices for different data subsets, or variations in levels implied by individual questions related to an element. The rating process involved discussion of current activities, opportunities and challenges.

Where ratings were less than desired, team members were asked to identify specific gaps in data governance, architecture and integration or data quality and briefly describe the consequences and implications of not addressing gaps were discussed.

Improvement Planning Workshop

A second workshop was held with each of the assessment teams to identify and validate data management gaps and further discuss the business consequences and impacts associated with the gaps. Each participant was asked to suggest 1-2 actions that they felt were most important to pursue in order to address one or more of the gaps. Actions identified had to be considered feasible to implement and had to have a clear business case.

For each action, teams were asked to identify any related gap(s); indicate whether it was a current initiative or something new; estimated timeframes for completion, and the person or unit who would be responsible for moving actions forward. Workshop participants were also asked to characterize the level of resources that would be needed for implementation based on the following categories:

- A1- can be done with existing staff or under the umbrella of existing/planned initiatives
- A2- requires modest level of additional resources –consultant contract (<\$100K), summer interns, temporary staff assignment
- A3- requires significant level of additional resources – consultant contract > 100K, new system acquisition/upgrade

3.3.3 Feedback on the Assessment Process

Participants in each assessment team were asked for their feedback on the value of the assessment process and suggestions for improvement. In general, participants felt that the process was valuable in that it provided an opportunity to take a step back and systematically look at the value added by data and the current data management practices. The process can help agencies go through a formal approach to identify what data are currently available and how they are being used. It can assist agencies to define and designate roles for collecting and managing core data.

Participants appreciated the opportunity to have a constructive dialog about data across staff working in different roles and areas of the agency. One individual noted that more collaboration and communication about data can be really valuable in helping the agency become more productive. Others commented that the assessment exercise reaffirmed the need for the projects and initiatives that are already underway to improve data and tools. Several suggestions were made to improve the assessment process and are synthesized below:

Preparing for the Assessment

- Involve a variety of staff with *varying roles and perspectives* – particularly in the Data Value Assessment. In particular, consider participants from district or regional offices, since they are often the beneficiaries of data improvements.
- For the Data Value Assessment, limit the number of different activities and data sets in order to streamline the assessment process. In addition, configure the Data Value Assessment to list *business activities in a logical order* (e.g. don't discuss program development prior to project scoping).
- Some agencies may want to consider developing business process maps and identifying how data are currently being used and by whom as one of the preparatory steps prior to the assessments. This would provide a *common foundation of information* for assessment participants.

Conducting the Assessment

- Ensure that participants have a *common understanding of the purpose* of the assessment and its anticipated outcomes. A communication from the executive sponsor prior to the workshop that articulates this can be helpful.
- Ensure that participants *understand terminology and meanings of the rating scales*.
- For both data value and Data Management Assessments, participants may have difficulty assigning ratings where the answer is “it depends”. For example, the quality of road inventory data might be very good for Interstates but poor for local roads. In these situations, *ratings will need to be assigned reflecting the predominant situation* – but participants should be given the opportunity to note exceptions or describe variations in practice.
- For the Data Value Assessment, it is important to *be very clear about what each business activity involves, and the differences between data availability, quality and usability*. Taking some time at the start of the initial workshop to walk through fundamental definitions is important to achieve meaningful and consistent assessment results.
- For the Data Value Assessment, participants may need *some initial education about each data set*. In particular, it is important that participants understand how certain foundational data sets are used to create other data sets (e.g. use of linear referencing data to establish location definitions for asset and project data).
- In the Data Value Assessment, participants are asked to assign ratings of the importance of each data set to each of the business activities. Participants need to understand that these ratings should consider the *potential value* of each data set for decision making – not just the degree to which it is currently used.
- Keep the assessments to a *manageable length* to avoid participant fatigue. The Data Value Assessment required participants to assign ratings to each combination of activity and data set; participants suggested streamlining this approach.
- For the Data Value Assessment, there is the potential for one or two low ratings for individual data sources to have a larger than expected influence on the overall

summary rating. It is important to *review and validate the overall results* to make sure that the summary scores sufficiently account for data value or importance.

- For the Data Management Assessment, it is important to remind participants and senior managers reviewing assessment results that the purpose of the maturity levels is to provide guidance for potential improvement. A lower maturity level shouldn't be viewed as a poor grade, and it is not realistic for agencies to reach the highest maturity levels across all assessment elements. Each agency needs to assess the cost and benefit of moving to a higher level of maturity – and recognize that there may be a point of diminishing returns – where additional efforts to improve data management capabilities will not be justified.
- There were varying views about setting targets. On the one hand, several participants felt that identifying gaps based on the difference between the current assessment level and a target level would be a useful way to share results with senior managers. On the other hand, participants stated that they would have difficulty identifying a target level without an understanding of what specifically would be involved (and how much it would cost) to achieve a particular level. Therefore, in the case studies, specific targets were not set – rather, gaps were identified based on discussion of the assessment ratings and the associated areas for improvement.

Developing Improvement Plans

- In considering candidate actions, make sure participants understand that a broad *range of improvements* – to data sets, information systems and data management processes should be considered.
- Recognize that *there isn't a one-to-one correspondence between improvements and gaps* identified in the assessment phase. A given improvement could address multiple types of gaps.
- Rather than trying to attack all types of data at the agency, *focusing on low hanging fruit* might be the way to start. Develop a standard checklist of activities for data governance that can be tested and rolled out incrementally.
- Emphasize the value of *repeating assessments periodically* to check over time how things are going.
- One result of the assessment could be to *reduce the amount of data that are being collected*. Having a data governance framework in place can make it easier to identify areas where data creates real value and where it does not. Agencies need to make sure that there are real business needs for their data – is anybody using it; how are they using it and are there other alternatives?

3.4 Summary of Key Findings from the Case Studies

The case studies proved to be a very valuable way to test the assessment methodology and obtain concrete feedback that can be used to enhance the methodology and produce the final guidance and tools. They demonstrated that the methodology is generally sound, feasible to implement and valuable to DOTs. They also resulted in identification of a

substantial number of refinements that can be made to improve the usefulness and likely impact of the assessment process. Specific findings are presented below.

Assessment Process and Results

Value. Participants saw value in the data self-assessment goals and the process. Groups agreed the process was valuable in building awareness of data customers, data uses, challenges, issues and opportunities. Participants were engaged and there was good discussion. Participating in the process can help DOTs understand the impacts and risks of paying insufficient attention to data management, especially with high turnover and a maturing workforce.

Results. The self-assessment process worked. Using a combination of survey tools and facilitated meeting discussions, participants were able to identify data program successes, challenges, gaps and consequences as well as actions and priorities for improvement.

Process. Simpler is better. The piloted approach was thorough and comprehensive and perhaps tried to cover too many different data quality and data management dimensions. Additional streamlining of assessment elements, along with the flexibility for agencies to exclude certain assessment elements would be helpful.

Applicability. The process seemed to be adaptable and flexible enough to address specific data program areas (mobility/travel data); functional areas (real estate and aeronautics); business processes (project scoping and design); and agency-wide initiatives (asset management). However, where there are multiple business processes and multiple data sets to consider, the assessment process can become complex and nuanced. For example, long-range planning involves several different activities and a wide range of data types.

Concepts and Terms. Data programs are not a well-defined entity within DOTs and there is not universal understanding of data management terms and concepts such as governance, stewardship, change management, data mapping or master data.

Time Requirements. Based on the case studies, the research team got a good feel for time requirements for the different assessment elements. The case studies included 2 2-3 hour workshops. Time requirements for the different groups varied based on group size and specific personalities involved, but in general, additional time was needed, particularly for the action planning workshop. We would also include time in the initial workshop to clarify goals of the assessment, review the scope of the process, discuss expectations and reach consensus on the definition of terms.

Participation. Engaging a cross section of data providers, data users, Information Technology (IT) professionals, GIS and application/system owners in the self-assessment process enhances overall success. Including participants from multiple business units ensures that a variety of perspectives are brought to the table.

Facilitation. Many of the groups agreed that having an outside or “unbiased” facilitator would be desirable. There was some discussion about whether the process could be conducted without outside assistance. It will be important to ensure that facilitators are

sufficiently prepared so that they have the necessary background to be successful. One idea discussed at the case study states was to have different states work together and provide each other with facilitators in order to have an unbiased perspective.

Assessment Preparation

Inventory. While the case studies didn't allow for completion of a detailed data program inventory, work with key contacts at each agency did confirm that some kind of data inventory would be a valuable, though not necessary essential element of the assessment preparation process. Both states thought it would be beneficial to define a relatively small number of data categories (e.g. less than 20) as a way to review and prioritize potential areas for data assessments.

Data Program Definition and Identification. The inventory exercise conducted by the two states flagged the need for additional work to define "data programs" for the purpose of the assessment. Challenges in identifying data programs that met that met the definition that was provided by the research team were related to three factors:

- The data management functions included in the definition are often split across business and IT units. The scope of responsibilities of the IT units may extend beyond the data sets or systems of concern to the business units.
- IT systems rather than "clusters of data" are the natural "unit" for consideration of data management practices. (In addition, several participants interpreted the word "program" to mean a "computer program" which further added to the confusion).
- Business units with data management functions do not typically consider themselves to be "data programs"- data management is viewed as a means to an end rather than an end in and of itself.

The case study experience indicated that rather than trying to be overly prescriptive about what constitutes a data program, the assessment guidance should outline the types of functions that a data program may include. It should emphasize the fact that a logical data program may include representatives from business and IT groups that collaborate on different aspects of data management. It should also point out that the assessment process will likely be easier if the number of different data sets and/or systems considered is limited.

Assessment Structure, Elements and Tools

Data Value and Data Management Split. Dividing the self-assessment process into separate data value and data management components worked well. Where a set of related business areas and data programs were selected for the two types of assessments it was helpful to bring together both the data user and data manager perspectives. Data value/quality discussions tended to blend into observations about the management, efficiency and effectiveness of data applications and systems. And, data management conversations tended to include discussions about data quality.

Tools. Online questionnaires were used for the Data Management Assessment; a spreadsheet tool was used for the Data Value Assessment. Each method had advantages and

disadvantages. Participants found the online questionnaires easier to complete – and the results were easier to compile. However, the online questionnaires did not include a capability to automatically assign maturity levels to assessment elements based on the responses – this was a manual process performed at the assessment workshop. In contrast, the spreadsheets allowed participants to see how different responses impacted the ratings. The major advantage of the spreadsheet was that it provided an easy way to customize the questions for each group to reflect specific business activities and data sources.

Data Management Structure and Assessment Elements. An initial comprehensive questionnaire was designed reflecting the different dimensions of data management maturity identified in the first phase of this project. Based on feedback, a simpler version of the questionnaire was needed for groups that didn't have enterprise data management as part of their responsibilities. Multiple suggestions were made for additional improvements and further simplification. Specific areas for improvement are:

- **Granularity:** The assessment process guidance and/or the questionnaire needs to clarify how respondents should address this – e.g. answer based on the weakest system; most important systems; predominant practices, or allow for users to explicitly provide a system by system breakdown for some questions.
- **Governance – Stewardship:** The guidance needs to refine questions to ask about multiple types/levels of data stewards and owners; ask about whether the agency has defined classes of data based on the need for enterprise or cross business-unit management activities.
- **Governance – Customers and Value:** The assessment elements should emphasize awareness of data customers and value added by data; de-emphasize accounting for full costs of the data program. Participants felt strongly that data program costs need to be discussed in the context of current and potential value and the benefit of the data to the agency.
- **Governance – Preservation and Security:** The assessment tools should emphasize alignment of data archiving and preservation activities with business needs.
- **Data Architecture and Integration – Linkages/Standards:** The guidance should consider splitting geospatial linkages and standards off as a separate element.
- **Data Architecture and Integration – Change Management:** The guidance should emphasize that this element covers technical capabilities for data change analysis – include governance-related elements under stewardship.
- **Data Quality:** In the pilot studies, specific questions about accuracy, completeness, consistency and currency yielded similar responses for the most part. Additional work may be needed to better define terms and simplify this portion of the questionnaire.

Data Management Maturity Levels. Mapping questionnaire responses to levels proved to be a greater challenge than anticipated. There were several cases where multiple questions were associated with a single assessment element. As noted under “granularity” above, further thought is also needed to address situations where one component of a data program meets criteria for a higher maturity level but others do not.

Data Value Structure and Assessment Elements. The data value questionnaire was

structured around a set of specific business processes that consume data (customized for each group), and a set of data sources that feed those processes (also customized for each group). Final assessment tools should consider simplifying the questionnaire so that participants are asked to assign a single quality and usability rating for each data source (rather than for combinations of business process and data source.) This simplification would enable additional data sources to be listed (i.e. more than 10 if appropriate), which would provide a helpful greater level of granularity for the rating process.

Data Value Questionnaire Configuration. Data value questionnaire configuration involved translating a set of generic business activities into more specific activities meaningful to each data value team. For two of the data value teams, the generic activities were directly applicable and could easily be translated into specific activities. However for two others, the generic activities did not fit. While the intent of having generic activities was to allow for comparison of ratings for generic categories across business areas, the case studies showed that it may not be feasible to maintain the five generic categories of activities across multiple data value groups. This implies that Data Value Assessment results may need to be summarized at the group level rather than disaggregated by activity category.

Data Value Rating Scales. For the Data Value Assessment, each assessment element (data availability, data quality and data usability) was assigned a poor-fair-good-excellent rating, and then an overall rating was assigned (1-Initial, 2-Limited, 3-Basic, 4-Sufficient, 5- Advanced). Because the poor-fair-good-excellent ratings are more intuitive than the overall ratings, and because they shed some light on where the potential issues are, the research team's conclusion is that the assessment summary should show both the detailed as well as the summary ratings.

Data Value Rating Process. In comparison to the data management questionnaires that emphasized yes/no questions, there was greater room for subjectivity in the data value ratings. The facilitator role was crucial for arriving at a balanced set of consensus ratings - through clarifying the definitions of each rating level and asking individuals to reconsider their ratings when it was clear that they were not interpreting the ratings as intended.

Target Setting and Improvement Planning

Targets. Each of the assessment teams discussed target setting, but none actually set targets. In general, participants felt that targets could be helpful to provide the motivation for improvement, but that the process could actually work fine without targets – i.e. gaps can be identified based on the assessment, and actions can be proposed and prioritized to close the gaps. In order to set targets, participants wanted to understand the implications of moving to the next maturity level – what actions would be required, what level of effort would be involved, and what benefits would result. Participants understood that some advances would be relatively straightforward whereas others might be much more difficult given organizational or resource constraints.

Given variations in groups and organizations, it is unlikely that standard guidance could be developed on the specific investments required to move across maturity levels.

Some participants suggested that defining performance metrics for data management would be more meaningful than maturity levels. Example metrics were the time required to stand up a server, or the total number of data schemas requiring management.

Another group suggested using the concept of diminishing returns in the context of target setting. The definition of maturity levels for data management tend to lead an agency towards increasing amounts of time and effort for analysis, documentation, review, coordination, and technology investments as they move up the maturity levels. The objective of target setting could be to identify (for each data program and assessment element) what the point of diminishing returns would be – where additional investment of time and money would not produce sufficient benefits to be justified.

Improvement Planning. The case study experience indicated that different approaches to prioritization or ranking should be allowed. Possible approaches are: numbering each action from highest to lowest priority; assigning high-medium-low priorities to each action; or assigning a sequence number rather than a priority rank to the actions to identify which ones should be pursued first (with the assumption that the initial activity may lead to new information that will impact priorities and perhaps the actions themselves). As another variation, actions may be clustered into logical groupings which are then ranked or rated.

Implementation. Linking the data self-assessment process and resulting action plan to other initiatives seems to be an effective strategy for initiating and sustaining the effort. Both case study agencies had a number of ongoing and planned initiatives that were related to the identified gaps. The assessment process provided a way to validate these initiatives and provide further support for seeing them through. One of the assessment teams noted that they had an existing action planning and tracking process within their division that would be the appropriate mechanism to vet, prioritize and monitor data improvements.

Guidance. Participants requested that guidance be included on how to improve data management maturity. For example, information on the steps for staging and implementing data governance, as well as examples of more mature practices and suggestions on the roles and responsibilities of governing bodies would be useful. In addition, data stewardship models appropriate for the DOT environment would be useful - traditional definitions for data stewards may need some readjusting to take into account the widely diverse data programs, applications, uses, customers and systems that exist in transportation agencies.

4.0 ASSESSMENT GUIDANCE AND FINAL TOOLS

Based on the findings summarized above, the research team revised the assessment methodology and developed final assessment tools. This chapter summarizes the major changes that were made and provides an overview of the tools and associated guide for conducting the assessment.

4.1 Revisions to Assessment Methodology

Data Value Assessment

The data value tool that was tested in the case studies included:

- A Configuration screen for specification of a set of business activities and data sets (or collections of data sets)
- An Importance screen for rating the importance of each data set for each business activity
- An Availability screen for rating the overall availability of data for each business activity
- A Quality screen for rating the quality of each data set with respect to each business activity
- A Usability screen for rating the usability of each data set with respect to each business activity
- A Results screen that assigned a maturity level for each business activity based on the Availability, Quality and Usability ratings (which were on 4 point scale from Poor to Excellent) and the relative Importance ratings off different data sets for each business activity

Based on the case studies, this approach was simplified. Rather than asking assessment teams to rate Quality and Usability for each combination of business activity and data set, the new approach asks for ratings of each data set (across all of the business activities). The original approach was based on the premise that both quality and usability of a data set should be assessed in relation to a target use for that data set. However, in practice, participants in the case studies tended to provide the same quality and usability ratings for a given data set across different business activities. This was due in part to fatigue given the length of the assessment, and in part because it was difficult for participants (who had varying levels of familiarity and hands-on involvement in each activity) to distinguish differences in quality and usability across different activities.

A second major change in the data value tool was to split the data Quality rating into three specific aspects of data quality: Currency, Accuracy, and Completeness. This was done in order to pinpoint the specific types of quality issues that were present in each data set and provide a better basis for identifying gaps and improvements.

The third change in the data value tool was to drop the maturity scale, and simply present results based on the four point scale (Excellent, Good, Fair, Poor). The maturity scale was originally included in the data value tool in order to provide consistency with the data management portion of the assessment. However, the approach to maturity level assignment based on the availability, quality and usability ratings resulted in a loss of useful information relative to the original sets of ratings. In addition, the maturity levels didn't add new meaningful or actionable information. In fact, maturity scales are more applicable for processes than for describing the state of the data, which is fundamentally what the Data Value Assessment is for. Therefore, the research team decided to drop maturity level ratings for data value and accept inconsistency in ratings between the data value and management assessments.

Data Management Assessment

The Data Management Assessments in the case studies were based on two online survey tools—one for agency-wide data management processes and a second for program-specific data management processes. The first tool included additional questions on use of enterprise-wide data management approaches. The research team developed a mapping from questionnaire responses to specific maturity levels for the different assessment elements. In many cases, responses from more than one question were considered in assigning a maturity level. Once the responses were compiled, there were many cases where it was difficult to assign a single maturity level. The results showed that additional work was needed to tighten up the criteria for assigning maturity levels. It was clear that the methodology needed to shift from one in which an agency or data program must meet all criteria for a particular level to one that can assign a level even when criteria for multiple levels are partially met.

A different assessment approach was developed that involved defining a set of criteria for each maturity level for each assessment sub-element, and asking the assessor to indicate the extent to which they agreed with the set of criteria: Totally Agree, Somewhat Agree, Somewhat Disagree and Totally Disagree. Logic in the assessment tool was established to assign a single maturity level for each element based on the highest level for which a threshold agreement value was assigned. This threshold is defaulted in the tool to "Somewhat Agree". Thus, for the following hypothetical set of ratings, Level 4 would be assigned, because this is the highest level for which Somewhat Agree (or Totally Agree) was selected.

- Maturity Level 1 – Somewhat Agree
- Maturity Level 2 – Totally Disagree
- Maturity Level 3 – Somewhat Disagree
- **Maturity Level 4 – Somewhat Agree ← ← ----- Selected Level based on User Ratings**
- Maturity Level 5 – Somewhat Disagree

In addition to changing the overall approach to assigning maturity levels, an effort was made to streamline the set of assessment elements to reduce the assessment length and associated level of fatigue observed in the case studies. Elements were eliminated or combined where

possible. One notable area where consolidation occurred was for the Data Quality element. Separate elements for data quality measurement and improvement processes for Accuracy, Consistency, Completeness and Timeliness & Currency were collapsed into a single data quality measurement and improvement element.

General Assessment Approach

In the case studies, each assessment team member was asked to complete the assessment (the data value spreadsheet or the data management online questionnaire) prior to the initial workshop. At the first workshop, it became clear that there were varying levels of understanding and multiple interpretations of the assessment elements. Many of the team members changed their ratings after having the opportunity to discuss the meaning of each element and obtain clarification on what the different levels implied. Because of this, the assessment process recommended in the final guide does not involve asking assessment teams to complete ratings prior to the first workshop. At the first workshop, time is allocated for the facilitator to walk through definitions and make sure that there is a common understanding of the elements and assessment levels. Then, participants are provided with time to complete each section of ratings individually (at the workshop) and then to discuss them and achieve consensus ratings with other team members. This approach should provide a much improved process. It eliminates the need for team members to complete up front work on a task that they may not fully understand. It also eliminates the effort required to distribute assessment tools, track completion of the assessment, follow-up with stragglers, and compile results prior to the first workshop.

4.2 Final Tools

Three assessment tools were developed – all in spreadsheet form:

- A Data Value Assessment Tool
- A Data Management Assessment Tool for agency-wide data management processes
- A Data Management Assessment Tool for specific data programs

The tools are intended to be completed by assessment teams as a group – therefore no capability for combining individual responses was developed. Agencies conducting Data Value Assessments for multiple business areas would use a separate spreadsheet for each business area, and then compile results from each area from the Results tabs of the different spreadsheets. A similar process would be used to compile results from different Data Management Assessments for multiple data programs (and for agency-wide processes.)

Descriptions and screen shots for the Data Value Assessment Tool and the Agency-Wide Data Management Assessment Tool are shown below. The Program-Specific Data Management Assessment Tool is very similar to the Agency-Wide tool; it includes a slightly different set of elements, and different criteria for some of the elements.

The tools are designed for use on a PC computer.

4.3 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. The following elements are considered:

- **Data Availability** – addresses whether the agency has the right kinds of data in place to meet its information needs.
- **Data Quality** – addresses whether the data that are available are good enough to meet the agency's information needs.
- **Data Usability** – addresses whether the agency's data are being provided in a convenient form for analysis and interpretation.

Availability is assessed with respect to specific business activities. *Quality* and *Usability* are assessed for each of the major data sources that are used for performing the selected business activities. The assessment categories ratings are: Excellent, Good, Fair and Poor. In order to calculate weighted ratings across a variety of data sources, the assessment also asks team members to rate the importance of each data source to each business activity.

The following sections describe each of the tabs in the Data Value Assessment Tool.

Menu Tab

The **Menu** tab of the Data Value Assessment Tool provides navigation buttons for each of the screens in the tool.



Figure 1. Data Value Tool – Menu Tab

Instructions Tab

The **Instructions** tab provides an overview of how to configure the tool and complete the assessment.

Instructions

Overview

This tool allows the agency to self-assess whether and how its data is adding value to a particular business unit based on four elements:

Data Importance
Data Availability
Data Quality
Data Usability

For the business unit in question, the user first provides a list of particular activities and data sources related to that business unit. For example, the user may provide the following inputs for the Maintenance Management business unit:

<u>Activities</u>	<u>Data Sources</u>
Maintenance Budgeting	Maintenance work orders, resource utilization, accomplishment
Maintenance Activity Tracking	Maintenance level of service/maintenance feature quality
Equipment Management	Fleet/equipment inventory and utilization data
Materials Management	Materials inventory and utilization data

The tool is then populated according to these activities and data sources, and the user answers questions related to each element specifically for the chosen activities and data sources.

Once the user has finished providing input, the data value results are summarized for the assessment elements at the levels of data source, activity, and business unit.

Workflow

Throughout this spreadsheet tool, cells that require an input from the user will have the following color: This is an input

The first tab after the instructions is the **Configuration** tab, which allows the user to do the following:

Enter the business unit to be analyzed.
Enter the activities to be analyzed within that business unit.
Enter the data sources to be analyzed within that business unit.

Once the user has entered these inputs, the user must press the Apply button to apply the changes to the rest of the spreadsheet. If these inputs are adjusted and the user does not press the Apply button before changing tabs, the user will be prompted to apply the changes.

The next tab is the **Example Lists** tab, where the user can see sample business units and associated activities, as well as sample data types and associated data sources.

Next are the four element tabs, where the user provides input for the activities and data sources associated with the business unit. The user must answer all input questions on each element tab to complete the assessment. However, users do not have to complete the elements in consecutive order and do not have to finish one entirely before moving on to the next one. The color of the element tab will dynamically adjust from orange to dark green based on the user's progress.

Once the user is finished, the **Results** tab shows the data value results for the assessment elements at the levels of data source, activity, and business unit.

Next

Figure 2. Data Value Tool – Instructions Tab

Example Lists Tab

The **Example Lists** tab can assist the assessment team members in identifying inputs for the **Configuration** tab. The tab includes a standard list of activities for six typical DOT business areas, and data sources for ten typical data types.

Sample Business Areas and Associated Activities						Sample Data Types									
Business Area	Activity	Program/Process	System/Tool	Reporting/Output	Frequency	Transportation System Management	Travel Economic and Forecast Data	Real Estate Data	Financial Report Data	Travel Data	General Data	Financial, Program, and Project Data	Performance & Measurement Data	Environmental Data	Regulatory and HR Data
Transportation System Management	System Performance Monitoring														
Travel Economic and Forecast Data	Travel Economic and Forecast Data	Travel Economic and Forecast Data	Travel Economic and Forecast Data	Travel Economic and Forecast Data	Travel Economic and Forecast Data	Travel Economic and Forecast Data	Travel Economic and Forecast Data	Travel Economic and Forecast Data	Travel Economic and Forecast Data	Travel Economic and Forecast Data	Travel Economic and Forecast Data	Travel Economic and Forecast Data	Travel Economic and Forecast Data	Travel Economic and Forecast Data	Travel Economic and Forecast Data
Real Estate Data	Real Estate Data	Real Estate Data	Real Estate Data	Real Estate Data	Real Estate Data	Real Estate Data	Real Estate Data	Real Estate Data	Real Estate Data	Real Estate Data	Real Estate Data	Real Estate Data	Real Estate Data	Real Estate Data	Real Estate Data
Financial Report Data	Financial Report Data	Financial Report Data	Financial Report Data	Financial Report Data	Financial Report Data	Financial Report Data	Financial Report Data	Financial Report Data	Financial Report Data	Financial Report Data	Financial Report Data	Financial Report Data	Financial Report Data	Financial Report Data	Financial Report Data
Travel Data	Travel Data	Travel Data	Travel Data	Travel Data	Travel Data	Travel Data	Travel Data	Travel Data	Travel Data	Travel Data	Travel Data	Travel Data	Travel Data	Travel Data	Travel Data
General Data	General Data	General Data	General Data	General Data	General Data	General Data	General Data	General Data	General Data	General Data	General Data	General Data	General Data	General Data	General Data
Financial, Program, and Project Data	Financial, Program, and Project Data	Financial, Program, and Project Data	Financial, Program, and Project Data	Financial, Program, and Project Data	Financial, Program, and Project Data	Financial, Program, and Project Data	Financial, Program, and Project Data	Financial, Program, and Project Data	Financial, Program, and Project Data	Financial, Program, and Project Data	Financial, Program, and Project Data	Financial, Program, and Project Data	Financial, Program, and Project Data	Financial, Program, and Project Data	Financial, Program, and Project Data
Performance & Measurement Data	Performance & Measurement Data	Performance & Measurement Data	Performance & Measurement Data	Performance & Measurement Data	Performance & Measurement Data	Performance & Measurement Data	Performance & Measurement Data	Performance & Measurement Data	Performance & Measurement Data	Performance & Measurement Data	Performance & Measurement Data	Performance & Measurement Data	Performance & Measurement Data	Performance & Measurement Data	Performance & Measurement Data
Environmental Data	Environmental Data	Environmental Data	Environmental Data	Environmental Data	Environmental Data	Environmental Data	Environmental Data	Environmental Data	Environmental Data	Environmental Data	Environmental Data	Environmental Data	Environmental Data	Environmental Data	Environmental Data
Regulatory and HR Data	Regulatory and HR Data	Regulatory and HR Data	Regulatory and HR Data	Regulatory and HR Data	Regulatory and HR Data	Regulatory and HR Data	Regulatory and HR Data	Regulatory and HR Data	Regulatory and HR Data	Regulatory and HR Data	Regulatory and HR Data	Regulatory and HR Data	Regulatory and HR Data	Regulatory and HR Data	Regulatory and HR Data

Figure 4. Data Value Tool – Example Lists Tab

Importance Tab

The **Importance** tab is used to rate the importance of each data set for conducting each of the business activities. In completing these ratings, assessment team members should consider whether the business activity could be conducted without each type of data – and how the results or effectiveness of the activity would be impacted if the data were not available. The importance ratings are as follows:

- **High Importance:** Essential; can't perform this activity without it
- **Medium Importance:** Valuable; could do without it, but would impact value or credibility of results
- **Low Importance:** Helpful; but could do without it
- **NA:** Don't use this data source for this activity

The Importance ratings are used to calculate overall ratings by activity – for example, so that a poor data quality rating is weighted more heavily for a highly important data set than for one that has relatively low importance. A more detailed description of this methodology is provided in the **Results** tab section.

All of the assessment tabs include the following features:

- A comments block that can be used to record rating explanations or exceptions raised during the assessment workshop.
- Five buttons on the top panel:
 - **Menu:** return to the menu screen
 - **Back:** go to the previous screen
 - **Next:** go to the next screen
 - **Clear:** clear all answers on the Tab
 - **Config:** return to the configuration screen
- A chart showing progress towards completion for each assessment element.

Menu	Data Importance Rating: Maintenance Management <i>How would you rate the importance of each data set for the following activities?</i>	Clear
Back	High Importance: Essential; can't perform this activity without it	Config
Next	Medium Importance: Valuable; could do without it, but would impact value or credibility of results	
	Low Importance: Helpful; but could do without it	
	NA: Don't use this data source for this activity	

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

Comments:

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

Comments:

Figure 5. Data Value Tool – Importance Tab

Availability Tab

The **Availability** tab asks the assessment team members to think about each of the business activities and whether there is sufficient data available to perform that activity efficiently and effectively. The Availability ratings are as follows:

- **Poor:** Little or no data available to support this activity
- **Fair:** Limited data available - large gaps remain
- **Good:** Basic data are available - some gaps remain
- **Excellent:** Sufficient data are available to meet needs - no gaps
- **NA:** Don't know - not enough information

The **Availability** tab deliberately does not ask about each of the different types of data – this is an opportunity to consider other types of data that may be valuable for each activity beyond those that have been identified in the Configuration. For example, there may be an unmet need for a type of data that the agency does not currently collect.

A brief description of the calculation of results is discussed in the **Results** tab section.

Menu

Back

Next

Data Availability Rating: Maintenance Management

How would you rate the availability of data to help you with these activities?

Poor: Little or no data available to support this activity

Fair: Limited data available - large gaps remain

Good: Basic data are available - some gaps remain

Excellent: Sufficient data are available to meet needs - no gaps

NA: Don't know - not enough information

Clear

Config

Activity	Availability Rating
Maintenance Budgeting	Fair
Maintenance Activity Tracking	Excellent
Equipment Management	Good
Materials Management	Good
Winter Maintenance - Snow Route Planning, Snow Plow Tracking	Fair ▼

Comments:

Figure 6. Data Value Tool – Availability Tab

Quality Tab

The **Quality** tab asks assessment team members to rate three different dimensions of quality of the different data sets, considering the most demanding needs across the collection of business activities listed on the **Configuration** tab.

The three different dimensions of data quality are as follows:

- **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.

The Quality ratings are as follows:

- **Poor:** Data aren't detailed, current, reliable or complete enough to be useful
- **Fair:** Data are useful but lack of currency, reliability 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

A brief description of the calculation of results is discussed in the **Results** tab section.

Menu

Back

Next

Data Quality Rating: Maintenance Management

How would you rate the quality of available data for these activities - considering level of detail, timeliness, completeness and consistency?

Poor: Data aren't detailed, current, reliable or complete enough to be useful

Fair: Data are useful but lack of currency, reliability 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

Clear

Config

Data Source	Quality Rating		
	Currency	Accuracy	Completeness
Maintenance Work Orders, resource utilization, accomplishment	Fair	Fair	Excellent
Maintenance level of service/maintenance feature quality	Poor	Poor	NA
Fleet/Equipment Inventory and Utilization Data	Good	Good	Good
Snow Plow Routes	Excellent	Excellent	Excellent
Materials Inventory and Utilization Data	Fair	Good	Excellent
Incident Data - type, status, response time	Excellent	Good	Good
Road Weather Data	Excellent	Excellent	Excellent
Oversize/Overweight Permitting Data	Excellent	Excellent	Excellent
Outdoor Advertising Data	Fair	Good	Good
Permits (access, right-of-way, utility, detour, etc.)	Excellent	Excellent	Excellent

Comments:

Figure 7. Data Value Tool – Quality Tab

Usability Tab

The **Usability** tab asks assessment team members to rate the amount of effort required to transform each type of data into a form that is useful, considering the most demanding needs across the collection of business activities listed on the **Configuration** tab. The Usability ratings are as follows:

- **Poor:** Data are available but require substantial effort to translate into usable form
- **Fair:** Data are available but require moderate effort to translate into usable form
- **Good:** Data are available in a usable form but improvements to reporting capabilities would be helpful
- **Excellent:** Data are available in a usable form - no improvements needed
- **NA:** Don't know - not enough information

A brief description of the calculation of results is discussed in the **Results** tab section.

Menu

Back

Next

Data Usability Rating: Maintenance Management

How would you rate the ease of accessing, summarizing and using the available data?

Poor: Data are available but require substantial effort to translate into usable form

Fair: Data are available but require moderate effort to translate into usable form

Good: Data are available in a usable form but improvements to reporting capabilities would be helpful

Excellent: Data are available in a usable form - no improvements needed

NA: Don't know - not enough information

Clear

Config

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

Comments:

Figure 8. Data Value Tool – Usability Tab

Results Tab

The **Results** tab provides a synthesis of the ratings – by data source, by activity, and for the business area as a whole.

For the Data Source summary:

- The Importance rating shown is the highest rating for the data source across all business activities;
- The Quality and Usability ratings are taken directly from the Quality and Usability tabs

For the Activity summary:

- The Availability rating is taken directly from the Availability tab
- The Quality ratings are calculated based on the methodology outlined below:
 - For each data source, the overall quality rating is determined – as the limiting (lowest) quality rating among currency, accuracy and completeness.
 - An overall data quality rating for an activity is calculated from the overall quality ratings for each data source by weighting the quality rating of each data source by the importance rating of the data source for that activity. Data sources with values of “NA” for either the importance rating or the overall quality rating (i.e. the assessment team members picked “NA” for currency, accuracy, and completeness) are not included in the calculation.
- The Usability ratings are calculated based on the same methodology – though the first step is not required because there is only a single rating per data set.

For the Business Area:

- Two sets of ratings are calculated based on the Availability, Quality and Usability ratings for each business activity. A “conservative” rating is based on the limiting factor or activity with the lowest rating. An “optimistic” rating is based on an average of the ratings across the different activities.

The **Results** tab includes two bar charts. The one in the lower left corner of the tab shows the percent complete for each assessment element. The one in the lower right corner shows the Business Area ratings.

The **Results** tab also includes a Print button to print the Results tab to one page.

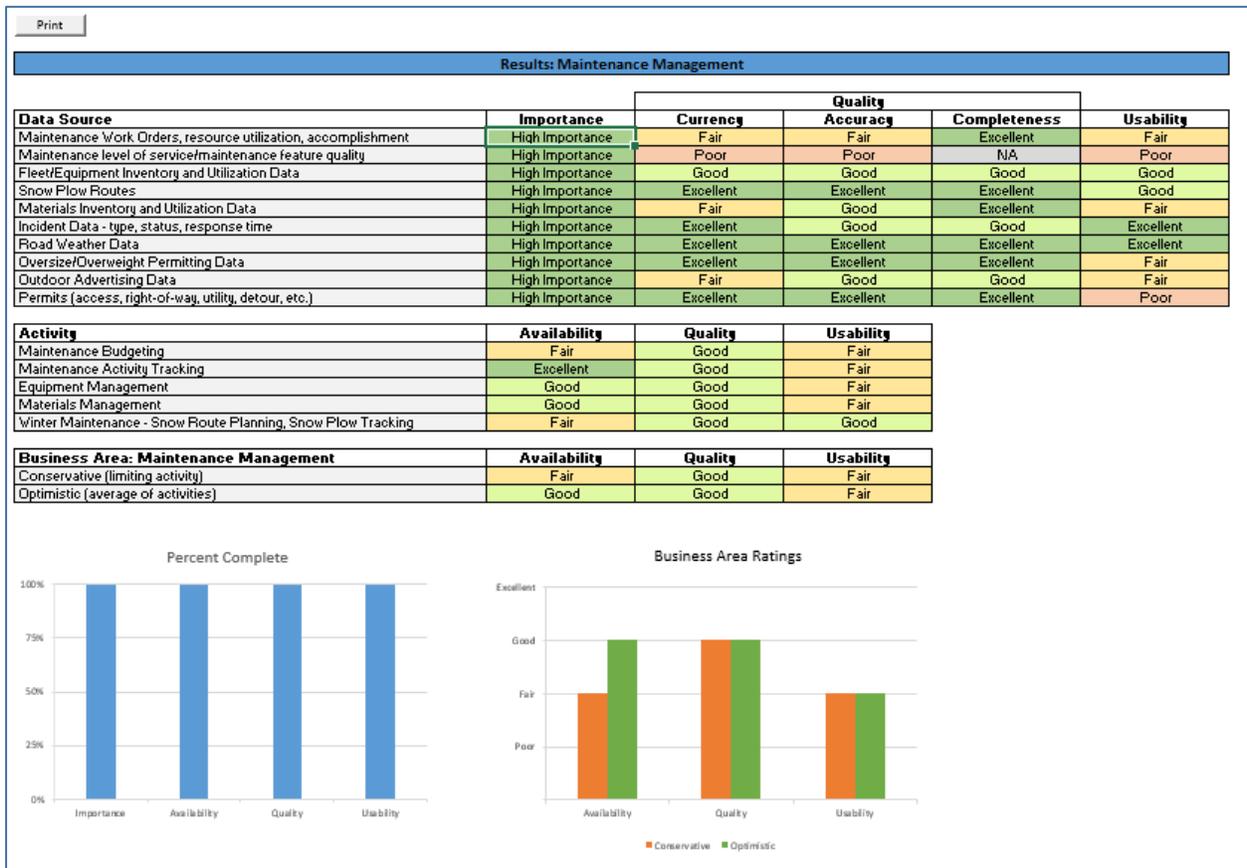


Figure 9. Data Value Tool – Results Tab

4.4 Data Management Assessment Tool

Overview

Two data management tools are available – one for an agency-wide assessment, and a second for an assessment of a specific data program or data category. These tools are very similar in structure, but have some variations 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. The following elements are considered:

- **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. It 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. It 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, assure quality as new data are acquired, and improve quality of existing data.

Each of these areas is broken down 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 a well-defined approach that is well-defined, documented and institutionalized within the agency or data program.

The following sections describe each of the tabs in the Agency-Wide Data Management tool. The Program-Specific Data Management tool is very similar to the Agency-Wide tool; it includes a slightly different set of elements and different criteria for some of the elements, but is otherwise identical.

Menu Tab

The **Menu** tab of the Data Management Assessment tool provides navigation buttons for each of the screens in the tool.

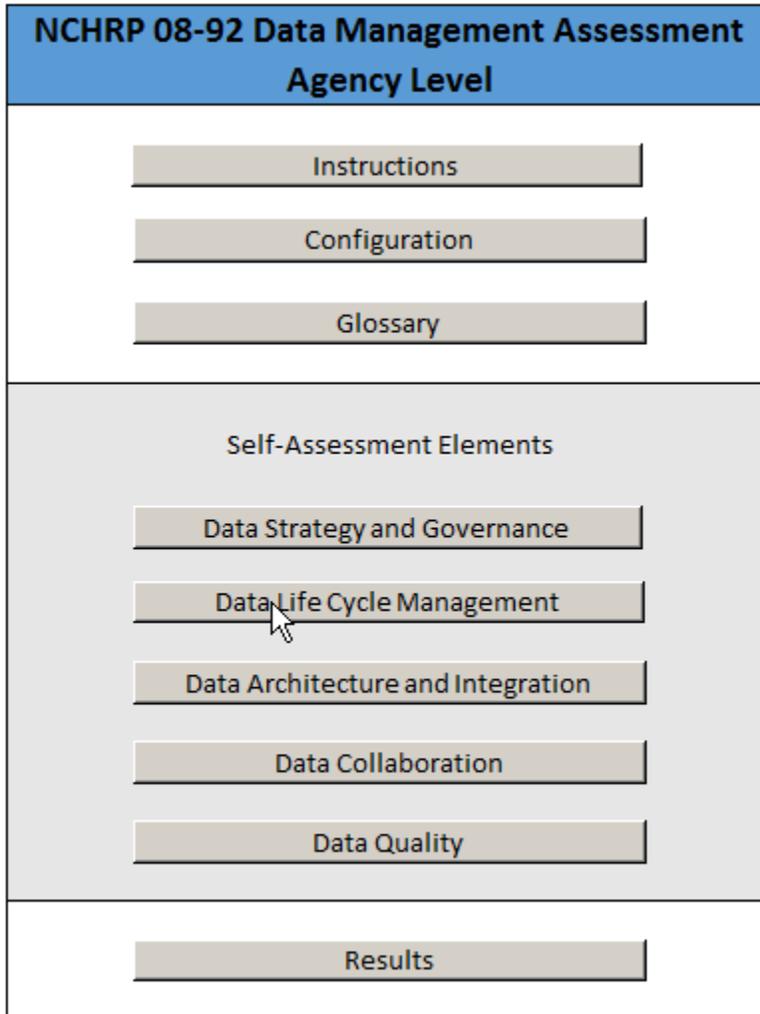


Figure 10. Data Management Tool – Menu Tab

Instructions Tab

The **Instructions** tab provides an overview of how to configure the tool and complete the assessment ratings.

Instructions

Overview

This tool allows the agency to self-assess the maturity level of its data management practices based on five elements:

- Data Strategy and Governance*
- Data Life Cycle Management*
- Data Architecture and Integration*
- Data Collaboration*
- Data Quality*

Each of these elements is further divided into sub-elements. For example, *Data Quality* is divided into the following sub-elements:

- Measurement and Reporting*
- Assurance and Improvement*

For each sub-element, there are five possible maturity levels with the following definitions:

<i>1: Initial</i>	<i>Ad-Hoc and event driven, success due to heroic efforts of individuals</i>
<i>2: Developing</i>	<i>Recognized need for improvement, pilot initiatives underway</i>
<i>3: Defined</i>	<i>Defined & documented processes not yet stabilized or widely socialized</i>
<i>4: Functioning</i>	<i>Implemented processes - operating and adding value</i>
<i>5: Sustained</i>	<i>Evaluated & improved processes, sustained over time</i>

The specific criteria for a particular maturity level depends on the sub-element.

Based on the provided criteria, the user is asked to rate how well the agency meets the criteria as follows:

- Totally Disagree*
- Somewhat Disagree*
- Somewhat Agree*
- Totally Agree*

Once the user has finished providing input, an overall maturity level is shown for each element and for the data management program as a whole.

Workflow

Throughout this spreadsheet tool, cells that require an input from the user will have the following color: This is an input

The first tab after the instructions is the **Configuration** tab, which allows the user to do the following:

- Adjust the weights between sub-elements used to calculate the maturity level of an element. *Setting a sub-element weight to 0% excludes it from the analysis.*
- Adjust the weights between elements used to calculate the maturity level of the data management program as a whole. *Setting an element weight to 0% excludes it from the analysis.*
- Adjust the threshold value that the user must provide for a maturity level to be a candidate for ranking.

The next tab is the **Glossary**, where the user can review the meaning of technical terms.

Next are the five element tabs, where the user provides input for each criteria.

The user must provide input for all criteria, except if an element or sub-element has been excluded from analysis in the configuration tab.

However, users do not have to complete the elements in consecutive order and do not have to finish one entirely before moving on to the next one.

The color of the element tab will dynamically adjust from orange to dark green based on the user's progress.

Once the user is finished, the **Results** tab shows the overall maturity level for each element and for the data management program as a whole.

Next

Figure 11. Data Management Tool – Instructions Tab

Configuration Tab

The **Configuration** tab allows the assessment team members to make two types of configurations: weights on elements and sub-elements, and 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 have 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%. Assessment team members can set weights to 0% for those sub-elements or elements that they do not wish to include, which will remove those sub-elements or elements from the calculation of results. The spreadsheet tool includes validation to ensure that the weights add up to 100%.

Adjusting the threshold value. The threshold value impacts how maturity levels are assigned to each sub-element. In order to assess the maturity level for each sub-element, the data management tools present a series of five descriptions of current practice – corresponding to the five maturity levels. Members of the assessment team are asked to indicate the extent to which they feel that each description reflects the current state of agency practice. Options are:

- Totally Disagree
- Somewhat Disagree
- Somewhat Agree
- 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 Somewhat Agree. The default value of this threshold can be changed to another rating value.

The tab includes five buttons:

- **Back:** go to the previous screen
- **Next:** go to the next screen
- **Return:** go to the last screen that was viewed prior to the configuration screen
- **Reset:** reset the configuration
- **Global Reset:** reset the entire worksheet, including the configuration and all inputted answers

Configuration

Element	Sub-Element	Weights	
		Sub-Element	Element
Data Strategy and Governance	Strategy and Direction	20%	20%
	Roles and Accountability	20%	
	Policies and Processes	20%	
	Data Asset Inventory and Value	20%	
	Data Management Sustainability	20%	
Data Life Cycle Management	Data Updating	17%	20%
	Data Access Control	17%	
	Data Findability and Documentation	17%	
	Data Backups and Archiving	17%	
	Data Change Management	17%	
	Data Delivery	17%	
Data Architecture and Integration	Location Referencing	25%	20%
	Geospatial Data Management	25%	
	Data Consistency and Integration	25%	
	Temporal Data Management	25%	
Data Collaboration	Internal Agency Collaboration	50%	20%
	External Collaboration	50%	
Data Quality	Measurement and Reporting	50%	20%
	Assurance and Improvement	50%	

Threshold
Somewhat Agree

This is the minimum response that the user must provide for a maturity level to be a candidate for ranking.

Figure 12. Data Management Tool – Configuration Tab

Glossary Tab

The **Glossary** tab includes a glossary of technical terms used throughout the assessment.

Glossary of Terms	
Business rule	A formally stated constraint governing the characteristics or behavior of an object or the relationship between objects (entities). Example: the width of an Interstate lane is 12 feet.
Change management	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.
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.
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, reports) including descriptive information on what is included and how to access, compiled for the purpose of facilitating discovery and understanding.
Data classifications	A set of categories used to distinguish key characteristics of a given data resource such as level of sensitivity or degree of importance, used to determine appropriate governance policies.
Data dictionary	A place where a limited set of "data about the data" or meta data are stored. It may include technical meta data including column names and formats and/or business meta data such as data definitions, business rules, etc.
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.
Data governance	The processes, practices, policies and decision making authorities in place to plan, monitor, and enforce the management of data assets.
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 enhancement.
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.
Data program	A business unit whose primary responsibilities include collection, processing, reporting and/or distribution of a particular class of data. DOT examples include: Traffic Monitoring, Crash Records, HPMS/Road Inventory, etc.
Data quality	The degree to which data is accurate, complete, timely and consistent with requirements and business rules and relevant for a given use.
Data set	A collection of data made available for access or download in one or more formats. Examples: a state's crash records for a single year; a database with roughness measures for pavement segments on the state highway system.
Data stewardship	The formal, specifically assigned and entrusted accountability for business (non-technical) responsibilities ensuring effective control and use of data and information assets.
Data visualization	Techniques for graphical representation of trends, patterns and other information.
Findability	The degree to which relevant information is easy to find when needed; findability is improved through application of metadata, taxonomies and other organizing tools, and search technologies.
Geospatial data	Data that include location, specified with explicit geographic positioning information.
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 along the network.
Master data	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, etc.
Meta data	Data describing content, content, and structure of documents and records and the management of such documents and records through time. Literally, data about data.
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.
Quality assurance	Processes to ensure that data meets specified requirements.
Quality control	Processes to detect defects in collected data and take appropriate action.
Sensitive data	Data that is confidential, privileged, or proprietary that should be protected from unauthorized disclosure, loss, misuse, or corruption in order to avoid serious consequences to the organization that owns it.

Figure 13. Data Management Tool – Glossary Tab

Data Strategy and Governance Tab

The **Data Strategy and Governance** tab contains the maturity level criteria for the five sub-elements of the Data Strategy and Governance element of the agency-wide data management maturity model. Assessment Team members indicate their degree of agreement with each of the five possible maturity levels for each sub-element (Totally Agree, Somewhat Agree, Somewhat Disagree, Totally Disagree). A maturity level is assigned for each sub-element based on the highest maturity level receiving the threshold value (Somewhat Agree by default) or higher.

All of the assessment tabs include the following additional features:

- A Comments block that can be used to record any details about the evaluation raised during the assessment workshop.
- Six buttons on the top panel:
 - **Menu:** return to the menu screen
 - **Back:** go to the previous screen
 - **Next:** go to the next screen
 - **Clear:** clear all answers on the tab
 - **Config:** return to the configuration screen
 - **Glossary:** return to the glossary screen
- A progress tracking bar chart on the right side panel. This provides a useful visual cue indicating whether some of the rating items have been skipped. To see more of the right side panel, the assessment team can decrease the zoom level.
- A radar chart showing the interim assessment results on the right side panel.

Menu	Back	Next	Clear	Config	Glossary
1. Data Strategy and Governance - Agency Level Evaluation <i>Leadership and management practices to manage data as a strategic agency asset.</i>					
1.1 Strategy and Direction <i>Leadership commitment and strategic planning to maximize value of data to meet agency goals.</i>					
1.1.1 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 on a reactive or opportunistic basis:					Totally Disagree
1.1.2 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 in an informal manner:					Somewhat Disagree
1.1.3 Executive leadership has communicated their expectation that business units and information technology functions should collaborate on identifying and implementing data improvements that are of agency-wide benefit: Data improvement needs have been systematically reviewed, assessed and documented:					Somewhat Agree
1.1.4 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:					Somewhat Disagree
1.1.5 Data governance and planning activities are continually refined to focus on key risks and opportunities - and eliminate activities without demonstrated payoff: Data governance and planning activities would have a high probability of continuing through changes in executive leadership:					Totally Disagree
Comments:					
1.2. Roles and Accountability <i>Clear roles, accountability and decision making authority for data quality, value and appropriate use.</i>					
1.2.1 Accountability for the quality, value, and appropriate use of data has not been clearly established:					Somewhat Disagree
1.2.2 One or more individuals have been identify 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 their role haven't been spelled out:					Somewhat Disagree
An agency-wide data governance body has been established with representation from information technology and business functions, and has defined its charter:					

Figure 14. Data Management Tool – Data Strategy and Governance Tab

Data Life Cycle Management Tab

The **Data Life Cycle** tab contains the maturity level criteria for the six sub-elements of the Data Life Cycle element of the agency-wide data management maturity model. It functions in the same manner as the **Data Strategy and Governance** tab.

Menu	Back	Next	Clear	Config	Glossary
2. Data Life Cycle Management - Agency Level Evaluation <i>Practices for managing data throughout its life cycle from collection to archiving or deletion.</i>					
2.1. Data Updating <i>Well-defined and coordinated data update cycles.</i>					
2.1.1 Data updating cycles and business rules for data updates have not been defined:				Totally Disagree	
2.1.2 Updating cycles have been established but have not been documented:				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:				Totally Agree	
2.1.4 Updating cycles are being consistently followed: Business rules for data updating are embedded in and enforced by applications (where applicable):				Somewhat Agree	
2.1.5 Data updating methods are periodically reviewed to identify opportunities for improved efficiencies:				Somewhat Disagree	
Comments:					
2.2. Data Access Control <i>Well-defined policies and guidelines for managing access to data sets.</i>					
2.2.1 There are no established policies for determining if access to data sets should be limited:				Answer	
<i>A process of defining what data is sensitive and needs to be protected is underway:</i>					

Figure 15. Data Management Tool – Data Life Cycle Management Tab

Data Architecture and Integration Tab

The **Data Architecture and Integration** tab contains the maturity level criteria for the four sub-elements of the Data Architecture and Integration element of the agency-wide data management maturity model. It functions in the same manner as the **Data Strategy and Governance** tab.

Menu	Back	Next	Clear	Config	Glossary
3. Data Architecture and Integration - Agency Level Evaluation <i>Technical standards, processes, tools and coordination mechanisms to maximize data integration and minimize duplication.</i>					
3.1. Location Referencing <i>Common location referencing methods across agency data sets.</i>					
3.1.1 The agency does not have a single common LRS: Data sets including location elements cannot be spatially integrated with other agency data sets:				Totally Disagree	
3.1.2 The agency is working towards establishing a single common LRS: Representation of location information is in the process of being standardized:				Totally Disagree	
3.1.3 The agency has developed a single common LRS: Quality standards for the LRS have been established with input from a variety of 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:				Somewhat Agree	
3.1.4 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):				Totally Disagree	
3.1.5 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 closely with agency GIS staff and proactively work to improve their data sets' consistency with agency-wide standards:				Totally Disagree	
Comments:					
3.2. Geospatial Data Management <i>Standardized approach to collection, management and integration of geospatial data.</i>					
3.2.1 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 information technology functions				Answer	
3.2.2 The agency has designated responsibilities for enterprise-wide planning and support for managing geospatial data				Answer	

Figure 16. Data Management Tool – Data Architecture and Integration Tab

Data Collaboration Tab

The **Data Collaboration** tab contains the maturity level criteria for the two sub-elements of the Data Collaboration element of the agency-wide data management maturity model. It functions in the same manner as the **Data Strategy and Governance** tab.

Menu		Back	Next	Clear	Config	Glossary
4. Data Collaboration - Agency Level Evaluation						
<i>Internal and external collaboration to maximize data sharing and avoid duplication of effort.</i>						
4.1. Internal Agency Collaboration						
<i>Collaboration across agency business units to leverage opportunities for efficiencies in data collection and management</i>						
4.1.1	Most data collection efforts in the agency are independent - there have been little or no efforts to coordinate across business units: The agency does not have information about the extent to which data are duplicated:					Totally Disagree
4.1.2	The agency has done an assessment of 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:					Somewhat Agree
4.1.3	The agency has implemented a data collection effort involving coordination from 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:					Somewhat Agree
4.1.4	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:					Somewhat Agree
4.1.5	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:					Somewhat Agree
Comments:						
4.2. External Collaboration						
<i>Partnerships with external entities to share data and avoid duplication</i>						
Individual business units obtain and use publicly available data from external entities as needs and opportunities arise:						
4.2.1	The agency has acquired single "point in time" data sets from external entities: External data requests are handled on a piecemeal basis					Answer
4.2.2	The agency is currently exploring partnerships with other public and private sector organizations to share data on an ongoing basis:					Answer

Figure 17. Data Management Tool – Data Collaboration Tab

Data Quality Tab

The **Data Quality** tab contains the maturity level criteria for the two sub-elements of the Data Quality element of the agency-wide data management maturity model. It functions in the same manner as the **Data Strategy and Governance** tab.

Menu		Back		Next		Clear		Config		Glossary	
<h3>5. Data Quality - Agency Level Evaluation</h3> <p><i>Standards and practices to ensure that data are of sufficient quality to meet user needs.</i></p>											
<h4>5.1. Data Quality Measurement and Reporting</h4> <p><i>Metrics and reporting to ensure user understanding of current data quality</i></p>											
5.1.1 There are no agency-wide activities related to data quality measurement and reporting:										Somewhat Agree	
5.1.2 The agency is exploring establishment of common data quality metrics for shared data elements:										Somewhat Disagree	
5.1.3 The agency has defined common data quality metrics across data programs in order to integrate data (e.g. locational accuracy):										Totally Disagree	
5.1.4 The agency has implemented data quality standards, verification techniques, and reports for common data elements:										Somewhat Disagree	
5.1.5 The agency proactively identifies new areas where common data quality metrics across data programs would be beneficial:										Totally Disagree	
<p>Comments:</p>											
<h4>5.2. Data Quality Assurance and Improvement</h4> <p><i>Practices for improving quality of existing data and assuring quality of newly acquired data</i></p>											
5.2.1 Data quality is assessed and improved in the context of individual data programs. No agency-wide support is provided:										Totally Agree	
5.2.2 Limited technical assistance is available for data program managers or business units on fundamental data quality concepts and practices:										Somewhat Agree	
5.2.3 The agency has established guidelines for determining spatial accuracy requirements and appropriate collection methods for new data collection efforts:										Somewhat Disagree	
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:											
5.2.4 The agency provides standard tools for gathering and tracking response to user feedback on data quality issues:										Somewhat Disagree	
The agency has deployed data profiling and cleansing tools and makes use of these tools to proactively identify (and, where possible) correct data quality issues:											
5.2.5 The agency provides tools for specification, maintenance and management of business rules:										Totally Disagree	

Figure 18. Data Management Tool – Data Quality Tab

4.5 Guide Development

A major deliverable of this project is a Guide that transportation agencies can use to carry out the data self-assessment. This Guide was created based on direction from the project panel and input from the focus group and case study participants. It provides an overview of the assessment process and step by step guidance for each of the three major phases: Prepare, Assess, and Improve. It also includes several appendices with resources for conducting the assessment:

- A glossary of terms
- A DOT data program inventory – that provides a starting point for agencies to compile a list of their major data programs as part of the Prepare phase of the assessment
- Documentation of the data management maturity model, with a description of each element and sub-element, criteria for assigning maturity levels to each sub-element, and discussion of the potential benefits agencies can realize by moving up the maturity scale.
- A data improvement catalog including a list of data improvements that agencies can consider for the Improve phase of the assessment, and a rich set of examples and resources that agencies can draw upon as they move forward with planning and implementing data improvements.

The Guide is published as *NCHRP Report 814: Data to Support Transportation Agency Business Needs: A Self-Assessment Guide*.

5.0 NEXT STEPS

Several steps are recommended for moving the work completed in NCHRP Project 08-92 into implementation. These are:

- Disseminate Project Results
- Develop a Strategy for Ongoing Stewardship
- Conduct Additional Testing of the Assessment Methodology and Tools
- Refine the Tools and Support Materials

Each of these steps is discussed below.

Disseminate Project Results

The first step in implementing the products of NCHRP 08-92 is to conduct outreach to ensure that agencies are aware of the self-assessment methodology and understand its value. Appendix C of this report includes a set of presentation slides that can be used in future outreach activities. Appendix D includes a flyer that provides an “at a glance” summary of the methodology.

Opportunities can be sought to disseminate project results at future upcoming TRB and AASHTO meetings – both as formal program presentations and as agenda items at committee meetings. In addition, members of the research panel can be requested to seek opportunities to brief others within their agencies as well as colleagues in peer agencies. Finally, one or more webinar presentations can be scheduled following publication of the final products to introduce the methodology and tools to the target audience. Webinars can be conducted and publicized either through TRB or AASHTO channels.

Develop a Strategy for Ongoing Stewardship

Maximizing value of the data self-assessment will require ongoing support to test, refine, market and support its application. Efforts to date to develop the transportation data self-assessment methodology have been the result of support and leadership from the Transportation Research Board Data Section committees, and the AASHTO Standing Committee on Planning Data Subcommittee. These are the logical entities to provide leadership for continued development and implementation support. Key stewardship activities include:

- Conducting periodic (e.g. quarterly) conference calls to coordinate ongoing outreach activities to ensure that DOTs are aware of the self-assessment tools and understand the benefits from using them;
- Guiding and funding further testing, development and evolution of the methodology and tools – based on input and feedback from target agencies;
- Establishing and maintaining a home web page for the self-assessment methodology, initially including links for downloading the products of NCHRP Project 08-92, and an

area where individuals could share comments on their experiences in implementing the assessment approach and tools and offer suggestions for improvement. Additional features could be added over time such as an online version of the tools, support for agencies to benchmark themselves against peer agencies, and a space to share resources on implementation of data management improvements.

Conduct Additional Testing of the Assessment Methodology and Tools

The data self-assessment methodology underwent a number of fairly substantial changes over the course of the project. An initial set of questions was tested in focus groups and then expanded for the case studies. Following the case studies, additional refinements were made to produce the current set of tools. Because these tools have not been tested in their current form, an additional cycle of application in a set of pilot states, followed by tool refinement would be beneficial. One way to accomplish this would be to initiate a pilot project through AASHTO in which 2-4 states would volunteer to test the methodology. Following the testing period, a peer exchange could be held to share results, challenges, and identify suggestions for refinement to the tools and the process guidance.

While many aspects of the existing methodology and tools are generic and have broad applicability, the primary target audience was state DOTs. In order to broaden the reach of the data self-assessment methodology, the testing process could seek to involve interested transit agencies and MPOs. Testing of the existing tools in these agencies would provide a basis for identifying the types of refinements or customizations that would add value for these agencies.

Refine the Tools and Supporting Materials

Results of the testing process could be used to develop additional refinements to the methodology, tools and guidance materials. Testing may also identify the need for additional support materials to facilitate implementation. The following ideas for tool refinements and support materials provide a starting point for scoping future enhancements to the assessment process. These should be validated and augmented following the testing process.

- Refine the maturity level criteria for the Data Management Assessment based on further testing to ensure that assessment teams understand the criteria and are able to clearly identify which level of maturity that their agency (or data program) should be placed;
- Supplement the maturity level criteria for the Data Management Assessment with additional examples of practice for each level – in “story form”;
- Develop a set of templates for the Data Value Assessment, potentially including:
 - templates that incorporate customized criteria for assessing data availability, quality and usability of data that reflect the types of questions and concerns raised in the executive interviews held at the start of this project – for example, “understanding mobility and safety from a multi-modal perspective”, or “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.”
- templates that draw upon existing assessment tools for transportation asset management, safety, transportation operations, and transportation performance management. For example, the Transportation Asset Management Gap Analysis tool references a number of asset management practices – such as tradeoff analysis, root cause analysis for asset failures, and data-driven prioritization. Each of these capabilities could be featured as a business activity, with a specific list of data sources that would be required to carry out the activity.

Agencies could use these specific templates to track their progress in improving the value of their data in support of these key practices;

- Develop customized versions of the tools for MPOs and Transit Agencies;
- Develop an online training course for assessment facilitators that covers the assessment process, the assessment elements, and techniques for managing assessment teams to stay on track and foster productive discussion;
- Develop model data improvement plans that can be used to facilitate the “Improve” phase of the assessment – with links to the improvement catalog information currently presented in Appendix C of the Guide in Volume 2;
- Create online (web) versions of the tools to replace the existing spreadsheet versions. This would facilitate tool application, and would allow for periodic tool refinement and enhancement. An online version of the tools could include a capability for agencies to create custom versions based on selection of specific transportation business areas or data programs. It could also support benchmarking of results across agencies.
- Create a “quick version” of the assessment perhaps in the form of checklists or online questionnaires that can be applied by agencies that are unable to devote the time to a full scale assessment.

REFERENCES

- [1] Secret, C., K. Schneweis, and G. Yarbrough. *NCHRP 08-36, Task 100 Transportation Data Self-Assessment Guide*. (August 2011) 56 pp. [Online]. Available: <http://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=2815>
- [2] Sawyer, M., N. Lefler, J. Soika, D. Carter, R. Scopatz, F. Gross, H. Rothenberg, J. Miller, G. Bahar, and K. Eccles. *United States Roadway Safety Data Capabilities Assessment*, Report No: FHWA-SA-12-028, Federal Highway Administration (July 31, 2012) 51 pp. [Online]. Available: http://safety.fhwa.dot.gov/rsdp/downloads/rsdp_usrsdca_final.pdf
- [3] Enterprise Data Management Council, *Data Maturity Model - unpublished review draft* – available to reviewers on request from: <http://www.edmcouncil.org/>
- [4] IBM Corporation, *The IBM Data Governance Council Maturity Model: Building a roadmap for effective data governance*, (2007) 15 pp. [Online]. Available: http://www-935.ibm.com/services/us/cio/pdf/leverage_wp_data_gov_council_maturity_model.pdf
- [5] National Association of State Chief Information Officers, *Data Governance Part II: Maturity Models - A Path to Progress* (March 2009) 30 pp. [Online]. Available: <http://www.nascio.org/publications/documents/nascio-datagovernanceptii.pdf>
- [6] Aiken, P., M. D. Allen, B. Parker, A. Mattia. “Measuring Data Management Practice Maturity: A Community's Self-Assessment.” *IEEE Computer* 40(4): pp. 42-50 (2007) [Online]. Available: <http://www.irmac.ca/0708/Measuring%20Data%20Management%20Practice%20Maturity.pdf>
- [7] American Association of State Highway, and Transportation Officials, *AASHTO Transportation Asset Management Guide*, Washington, D.C. (2002)
- [8] American Association of State Highway, and Transportation Officials. *AASHTO Transportation Asset Management Guide: A Focus on Implementation*. AASHTO (2011)
- [9] Cambridge Systematics, Inc., Boston Strategies International, Inc., Gordon Proctor and Associates, and M. J. Markow, *NCHRP Report 666: Target-Setting Methods and Data Management to Support Performance-Based Resource Allocation by Transportation Agencies: Volume I: Research Report, Volume II: Guide for Target-Setting and Data Management*. Transportation Research Board of the National Academies, Washington, D.C. (2010) 133 pp. [Online]. Available:

[http://www.trb.org/Main/Blurbs/TargetSetting Methods and Data Management to Suppo 164178.aspx](http://www.trb.org/Main/Blurbs/TargetSetting%20Methods%20and%20Data%20Management%20to%20Support%20164178.aspx)

- [10] B. Kahn, D. M. Strong, and R. Y. Wang, "Information Quality Benchmarks: Product and Service Performance." *Communications of the ACM*, Vol. 45, No. 4 (April 2002) pp. 184-192. [Online]. Available: <http://web.mit.edu/tdqm/www/tdqmpub/KahnStrongWangCACMApr02.pdf>
- [11] Creating an Effective Program to Advance Transportation System Management and Operations. Primer. FHWA-HOP-12-003 (January 2012) 14 pp. [Online]. Available: <http://ops.fhwa.dot.gov/publications/fhwahop12003/fhwahop12003.pdf>.
- [12] *Data Use for Improved Learning Needs Assessment*, Assessment and Accountability Comprehensive Center (2010). <http://datause.cse.ucla.edu/needs.php>> Accessed July 30, 2013.
- [13] NCHRP Project 20-90 web page: <<http://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=2981>>, accessed July 24, 2013
- [14] Olsen, M. J., J. D. Raugust and G. V. Roe, *NCHRP Synthesis 446: Use of Advanced Geospatial Data, Tools, Technologies, and Information in Department of Transportation Projects*. Transportation Research Board of the National Academies, Washington, D.C. (2013) 95 pp. [Online]. Available: <http://www.trb.org/Main/Blurbs/169202.aspx>
- [15] Harrison, F. *Meeting Critical Data Needs for Decision Making in State and Metropolitan Transportation Agencies – Summary of a Conference*. Transportation Research Board Conference Proceedings on the Web 9, Transportation Research Board of the National Academies, Washington, D.C. (2013) 114 pp. [Online]. Available: <http://onlinepubs.trb.org/onlinepubs/conf/CPW9.pdf>
- [16] Cambridge Systematics, *NCHRP Report 706: Uses of Risk Management and Data Management to Support Target-Setting for Performance-Based Resource Allocation by Transportation Agencies*. Transportation Research Board of the National Academies, Washington, D.C. (2011) 59 pp. [Online]. Available: <http://www.trb.org/Main/Blurbs/166250.aspx>
- [17] Flintsch, G. W. and J. Bryant. *Asset Management Data Collection for Supporting Decision Processes*. FHWA (May 2006) 96 pp. [Online]. Available: http://www.fhwa.dot.gov/asset/dataintegration/if08018/assetmgmt_web.pdf
- [18] Cambridge Systematics, *CDOT Performance Data Business Plan*. (December 2011) [Online]. Available: http://www.camsys.com/pubs/CDOT_Performance_Data_Business_Plan.pdf

- [19] Minnesota Department of Transportation Data Business Plan (2008). 92 pp.
- [20] Cambridge Systematics, *Data Governance, Standards, and Knowledge Management White Paper of the Alaska Data Business Plan*. (2009)
- [21] Virginia Department of Transportation System Operations Directorate Data Business Plan (2008)
- [22] Transportation Research Board Committee on Strategies for Improved Passenger and Freight Travel Data, *TRB Special Report 304: How We Travel: A Sustainable National Program for Travel Data*. Transportation Research Board of the National Academies, Washington, D.C. (2011). [Online]. Available:
[http://www.trb.org/Main/Blurbs/ How We Travel A Sustainable National Program for 164998.aspx](http://www.trb.org/Main/Blurbs/How_We_Travel_A_Sustainable_National_Program_for_164998.aspx)
- [23] Cempel, E. *A CFO's Handbook on Performance Management*. AASHTO [2010] [Online]. Available: http://www.transportation-finance.org/pdf/0211_performance_management_briefing_paper.pdf
- [24] Lawrence, M. D. Cartwright-Smith, J. Mans, P. Nguyen, and N. Lefler. *Benefit-Cost Analysis of Investing in Data Systems and Processes for Data-Driven Safety Programs*. Report No: FHWA-SA-12-029. Federal Highway Administration, (August 2012) 63 pp. [Online]. Available: <http://safety.fhwa.dot.gov/rsdp/downloads/bcareport.pdf>
- [25] Data Management Association, *Data Management Body of Knowledge Functional Framework* (2008).
- [26] J. Dyché "Five Models for Data Stewardship" (2009),
<<http://www.dataflux.com/Login.aspx?ReturnUrl=%2fResources%2fDataFlux-Resources%2fWhite-Paper%2fFive-Models-for-Data-Stewardship.aspx>>, Accessed July 23, 2013
- [27] FHWA Highway Performance Monitoring System Field Manual. (March 2013), 279 pp. [Online]. Available:
<http://www.fhwa.dot.gov/policyinformation/hpms/fieldmanual/hpms2013.pdf>
- [28] *Model Minimum Uniform Crash Criteria. 4th Edition*. Governors Highway Safety Associate, Federal Highway Administration, Federal Motor Carrier Safety Administration, and National Highway Traffic Safety Administration (2012)
- [29] Lefler, N., F. Council, D. Harkey, D. Carter, H. McGee, and M. Daul. *Model Inventory of Roadway Elements*. Report No: FHWA-SA-10-018. Federal Highway Administration (2010)

- [30] Fiedler, R., K. Eccles, N. Lefler, A. Fill, and E. Chan. *MIRE Fundamental Data Elements Cost-Benefit Estimation*. Report No: FHWA-SA-13-018. Federal Highway Administration (2013)
- [31] Brecher, A., J. Hassol, and S. Sloan, *State-of-the-Practice and Lessons Learned on Implementing Open Data and Open Source Policies*, FHWA Office of Advanced Travel Management. Report FHWA-JPO-12-030 (May, 2012) 76 pp. [Online]. Available: http://ntl.bts.gov/lib/47000/47200/47239/FHWA-JPO-12-30_Final_Pkg_v2.pdf
- [32] Knodler, M. A., R. Riessman, H. Rothenberg, and J. Kennedy. *Improving Crash Data Quality in Massachusetts*. Report No: UMTC 08-03. Massachusetts Executive Office of Transportation and Public Works (2008)
- [33] Boodlal, L., J. Emery, and R. Souleyrette. *Crash Data Improvement Program Guide*. Federal Highway Administration (April 2010) 58 pp. [Online]. Available: <http://safety.fhwa.dot.gov/cdip/finalrpt04122010/finalrpt04122010.pdf>
- [34] Pierce, L. M., G. McGovern, and K. A. Zimmerman, *Practical Guide for Quality Management of Pavement Condition Data Collection*. FHWA (2013) [Online]. Available: http://www.fhwa.dot.gov/pavement/management/qm/data_qm_guide.pdf
- [35] *Recommended Framework for a Bridge Inspection QC/QA Program*. FHWA web page <<http://www.fhwa.dot.gov/bridge/nbis/nbisframework.cfm>> Accessed July 29, 2013
- [36] Battelle, Cambridge Systematics, Inc., and Texas Transportation Institute, Texas A&M University, *Traffic Data Quality Measurement – Final Report*. FHWA Office of Highway Policy Information. (September 15, 2004) [Online]. Available: <http://isddc.dot.gov/OLPFiles/FHWA/013402.pdf>
- [37] FHWA Traffic Monitoring Guide (2001, Updated 2013)
- [38] Bremmer, D., K. Cotton, D. Cotey, C. Prestrud, and G. Westby, *Measuring Congestion: Learning From Operational Data* (2004) 18 pp. [Online]. Available: http://www.wsdot.wa.gov/NR/rdonlyres/E8FBCE6A-BF80-4787-9AF7-8D1799EA0327/0/WSDOT_Measuring_Congestion.pdf
- [39] Ahn, K., H. Rakha and D. Hill, *Data Quality White Paper*. FHWA Office of Operations (June, 2008) 42 pp. [Online]. Available: http://ops.fhwa.dot.gov/publications/fhwahop08038/pdf/dataqual_whitepaper.pdf
- [40] Oasis Consulting Services, *A Transportation Executive's Guide to Organization Improvement*. Prepared for NCHRP Project 20-24(42) Guidelines for State DOT Quality Management Systems, June, 2006.

APPENDIX A – FOCUS GROUP RESULTS

Chapter 2 summarized the information gathering phase of the projects, which included focus groups in five state DOTs. This Appendix provides additional detail on the types of data improvements suggested by the focus group participants, and the feedback on the assessment framework.

Data Improvements

As noted in Chapter 2, one of the focus group exercises involved asking participants to talk about their “wish lists” for data improvements. This question was intended to get participants thinking about the types of suggested improvements that they would expect to come out of a data self-assessment process.

Table A.1 summarizes the types of data improvements suggested by focus group participants, organized by data program type, as well as by the original Data Management Assessment dimensions based on the NCHRP Project 8-36, Task 100 report.

Table A.1. Focus Group Data Improvements

Data Program/ Function	Improvement Dimension	Improvement Type
Pavement	Alignment-Data Utilization	Improve data analysis efforts to develop predictive models
	Quality-Accuracy	Implement additional quality assurance checks on vendor provided pavement data
Bridge	Alignment – Data Utilization	Link bridge data with STIP project and cost data
	Alignment – User Needs	Provide access to as-built plans highlighting variations from construction plans
	Alignment – Standards	Transition to use of new National Bridge Elements
	Quality – Timeliness	Provide faster turn-around time between inspection and data provision to users
Asset data (multiple)	Alignment – User Needs	Get agreement on sustainable standards for what data should be collected for each asset + define improvements & data governance policies
	Data Management – Integration	Implement spatial data integration including LRS (route and mile point) updating across systems.
	Data Management – Technology Upgrades	Improve system for updating bike map info (shoulder widths, surface types, etc.)
Maintenance	Alignment – Data Utilization	Improve analysis of life cycles, cost-effectiveness for pavement striping options

Data Program/ Function	Improvement Dimension	Improvement Type
	Alignment – Data Utilization	Enhance data integration and tools for tradeoffs, ROI, life cycle costs and include abilities to link asset condition and work management data
Road inventory	Alignment - Standards	Complete MIRE subset data
	Alignment – User Needs	Improve local road data
	Alignment – Data Utilization	Improve packaging of information for decision support – involves integration with asset, maintenance, program, financial data.
	Alignment – Data Utilization	Improve accessibility of the data
	Data Management – Technology Upgrade	Replace mainframe, upgrade software for automating road network updating
	Data Management – Definitions	Improve data documentation
	Alignment – User Needs	Use mobile LiDAR to get accurate data on side slopes, other features.
Real time traffic & incidents	Alignment – Data Utilization	Improve management and ability to aggregate/link different data sets and present to meet user needs
	Alignment – User Needs	Collect additional speed data to calculate system delay performance measures
Traffic counts	Alignment – User Needs	Collect more ADTs on local roads
	Data Management – Technology Upgrade	Develop a comprehensive management system for integrating multiple data collection and processing processes
	Data Management – Program Tradeoffs and Costs	Reduce safety risks associated with data collection
Crash	Alignment – Data Utilization	Improve tools for user access & reporting
	Alignment – User Needs, Data Utilization	Improve data on slopes; provide ability to overlay inventory data with crash data
	Quality – Accuracy	Improve accuracy and completeness of data
	Quality - Timeliness	Improve timeliness – from crash to availability in system
	Data Management – Technology Upgrades	Upgrade technology to facilitate data updating, access, query, mapping and visualization features
System performance	Alignment – User Needs, Data Utilization	Provide travel time trends for arterials and interstates based on real-time information

Data Program/ Function	Improvement Dimension	Improvement Type
	Data Management – Integration	Automate data extraction from source systems
Local programs	Alignment – Data Utilization	Improve data query, mapping and visualization tools (for state funded local projects) for external stakeholder use
Planning, Travel Modeling, Performance	Alignment – Sources, Users and Uses	Implement methods to better understand how data are being used and impacts of not having the data
	Alignment – User Needs	Provide better info on non-highway modes – bike/pedestrian, intermodal Improve freight data
	Alignment – User Needs	Provide better info on commercial vehicle travel in urban areas
Motor Carrier – IFTA & IRP	Data Management – Technology Upgrade	Rewrite/replace major legacy system
Financial data	Alignment – Data Utilization	Improve integration across budget, expenditure, procurement, contracts, program, federal program systems to produce meaningful statistics
	Alignment – Sources, Users and Uses	Facilitate better understanding of how data are being used, and needs for training or tools to support use
Capital Program data	Alignment – Data Utilization	Improve integration of federal financial data (obligations, sources) with capital program data
Multiple – district engineer perspective	Alignment – Data Utilization	Present a variety of information on system condition and performance in a convenient way to help us evaluate the system and develop improvement projects.
GIS data	Data Management – Governance Policies	Implement more formal data update cycles
	Data Management – Integration	Improve integration with source data systems to automate creation/updating of spatial data
Business Intelligence, Strategic Data Management	Alignment – Sources, Users and Uses	Provide resources to deliver added value to data Provide information to help users better understanding of value, strengths, weaknesses
	Alignment – Sources, Users and Uses	Provide additional information to better understand what technology investments needed to align with business needs and move forward
	Alignment – User Needs	Provide better data to support the programming function

Data Program/ Function	Improvement Dimension	Improvement Type
	Data Management – Ability to Segregate, Aggregate and Analyze	Establish, clarify and/or strengthen data architect roles to improve integration & efficiency
	Data Management – Technology Upgrades, Governance Policies	Manage expanding data resources – and adopt and implement storage, archiving and retention strategies

Focus Group Feedback on Data Self-Assessment Elements

The following summarizes the suggestions by focus group participants about elements to include in the assessment process, organized by the original three categories from the NCHRP Project 8-36, Task 100 report: strategic alignment, data quality and data administration/management.

Strategic Alignment

- Data Program Alignment with Individual Business Unit Needs
 - Do business units have the data needed to: understand existing conditions/performance, predict future conditions, and understand impacts of investments or policy changes on predictions?
 - Distinguish the multiple factors that may prevent a data program from meeting the needs: not providing the right data; not having the ability or resources to collect more data; not having the tools needed to store, manage, analyze or distribute the data; not actually integrating data to meet decision needs (e.g. crash + traffic volumes + road inventory + GIS); not providing needed level of granularity or aggregation and trend data.
 - Are data programs meeting the essential business needs of the unit –if not, what specific needs are not being met?
 - Is there regular communication between data providers and users to ensure alignment of data with user needs?
 - Is the data program responsive to user needs and suggestions for improvement?
- Data Accessibility and Usability
 - How are data accessed (on request, via data mart, GIS map, etc.)?
 - To what extent do people know where/how to access data?
 - To what extent are people expected or required to use the data, trained to use it, and held accountable for using it?
 - Distinguish accessibility for improving transparency vs. accessibility for reducing internal staff time for data processing and analysis.
 - Recognize tradeoffs between providing open access to data and the potential for misinterpretation or lack of consistent use.
- Data Program Alignment with Strategic Priorities Across the Organization

- The maturity progression for strategic data alignment would include: (1) understanding that there is a data need, (2) precise/formal definition of the need, (3) understanding of what data exists, (4) establishing data ownership, (5) defining required data elements, (6) setting up data access channels, (7) preparing a strategic plan for data, (8) conducting annual reviews of needs, (9) establishing a feedback loop, (10) organizational capacity for change/improvements.
- Is there a shared understanding of the available data and its value across the silos?
- Has a data mapping (from data sets/programs to business processes) been done?
- Is the agency able to identify and prioritize data gaps?
- Is the agency able to identify data that needs to be shared across the agency?
- Has the agency identified “mission critical” data programs – i.e. programs that would not be cut if staffing levels had to be reduced?
- Does the agency understand the extent to which data are being used and the value added by use (both strategic and operational)? This is an important dimension of strategic alignment and could be used for prioritization.
- Are there examples of situations where too much data is being collected given value being provided?
- What level of sponsorship exists for different data programs?
- How much is being invested in different data programs?
- Ask if the organization assesses or understands the risks of not addressing the identified data gaps.
- Leveraging Data to Meet Multiple Needs
 - How willing and able are different business units to repurpose a data set for different users/uses?
 - Note that there may be differences in accuracy requirements across business areas, and that it is not always necessary (or cost-effective) for a data program to meet multiple needs.
 - What is the agency’s comfort level with using external (e.g. private sector) data sets? (May be a need to overcome trust, “black-box” issues.)
- Meeting External Requirements
 - What specific data requirements need to be met (e.g. MAP-21, HPMS, state legislation, etc.)? Are these already in place or anticipated?

Data Quality

- Timeliness and Currency
 - Some participants weren’t familiar with the distinction between timeliness and currency. They are related but not the same: timeliness is based on how long it takes from collection to delivery, whereas currency is based on the age of the data. From a user requirement standpoint, currency is the primary concern; data program owners may be more concerned about timeliness.
 - Have data collection, updating and reporting cycles been established based on currency needs?

- Has the data program owner or steward assessed user needs or expectations for static or “frozen” data sets – as opposed to current data sets?
- Quality Assurance and Improvement Processes
 - Are data validation and reconciliation processes used?
 - Have business rules been defined and documented for valid values (field and record-level)?
 - Are there staff resources with the time, expertise, and training to perform effective QA (including the ability to spot data quality issues)?
 - Is there an active feedback loop from data users to providers to report and respond to quality concerns?
 - Do data collectors understand how data are to be used and by whom?
- Establishing Expectations and Standards for Quality
 - Is there a governance body in place to make decisions about quality standards when there are multiple potential users for a data set, each with differing requirements and priorities? For example, planning and engineering users typically have different data requirements. How are decisions made about whether to capture the level of detail, accuracy, and currency required for engineering use? How does the agency factor in the need to anticipate potential future uses for the data sets in these decisions – where requirements cannot be fully understood?
 - Are cost and sustainability implications of different accuracy levels analyzed?
 - Are confidence levels for data calculated, used to establish sample sizes and published as metadata?
 - How do data owners, stewards and users perceive data quality and the costs of improving quality? Many data users will only have an anecdotal understanding of data quality.
 - How do perceptions about quality line up with available quality measurements?

Data Program Administration/Management

- Data Documentation and Program Sustainability
 - How is data lineage tracked – to identify the original master source and the changes applied along the way? How are derivative data products managed? Are they automated?
 - Is there training on what data exists and how to use it?
 - To what extent are data management roles and practices sustained over time? Is there a succession plan for data managers/experts, manuals/documentation, retention schedules and archives, ability to weather changes in leadership?
- Governance
 - Are owners/stewards identified for each data set? Need to distinguish person with accountability and decision making authority for improvement from the person with the best technical understanding of the data.
 - Are there support systems in place for using, sorting, reporting data (documentation, training, technical assistance) – to ensure appropriate usage?

- Is there communication among data producers, owners and users?
- Are there Service Level Agreements (SLAs) for stewardship roles & responsibilities? SLA process can be used to identify customer needs and set expectations.
- Are governance structure and processes in place, including establishment of managed source systems? Is governance extended to external contractors/consultants that collect data or provide systems?
- To what extent are governance policies periodically reviewed and revised/improved?
- To what extent does the agency measure awareness, reach, effectiveness and costs of data governance?
- What level of organizational buy-in exists for enterprise solutions vs. independent invention of tools and methods (criteria for the “aware” level?)
- Distinguish *support* for data use from *governance* of data use.
- Maturity progression: siloed, single purpose, reactive improvements, ad-hoc processes, no documentation, supplied on request to codified and automated processes, documentation, marketing, disclaimers, self-serve, iterative improvements.
- Technology and Automation
 - Does the agency has a data warehouse program in place?
 - Does the agency has sufficient network connectivity and bandwidth to enable remote data access from its field offices?
 - Does the agency have current tools and technology for data management?
 - Are managers aware of risks or missed opportunities associated with use of older technology?
 - How able are data program managers & agency as a whole to keep up with needs give constant changes in technology?
 - Are data analysis tools available? If available, to what extent are they used?
 - What is the current efficiency of data management processes – collection, loading, updating, validation, synchronization, aggregation, distribution, etc.?
- Data Security
 - Are there corporate policies on data privacy and sensitivity? Does the agency periodically verify that these policies are being followed?
 - Is a disaster recovery plan in place for the data program?
- Change Management
 - Is information about changes to data collection, processing and reporting methods provided to data users?
 - Does a formal data change management process exist? A change management process would include assessing impacts of changes to data structures or master data content, consultation and communication with affected data owners and users, and propagating the changes across databases as needed. How efficient is the process?

Focus Group Feedback on the Data Self-Assessment Process

Focus group participants were also asked for their input on how the assessment process would be carried out. These suggestions are summarized below, organized according to the three major steps in the process: Prepare, Assess, and Improve.

Prepare Step

- Focus on a manageable area; don't try to "boil the ocean"
- Set a clear timeframe for the assessment effort – e.g. 3-9 months to maintain interest and focus and minimize impacts of staff turnover and competing initiatives. Timeframe for improvements would be longer, but well-defined (e.g. 1-3 years)
- Executive sponsorship is important for this initiative to succeed – begin with executive-led team to identify focus areas, then form specific teams as appropriate (e.g. based on key performance areas such as safety, mobility, preservation or based on fund managers); identify a small core implementation team to manage and track the overall assessment process.
- Include representatives of users, collectors, data system owners, analysis tool owners; consider external stakeholders as well (e.g. local traffic engineers, MPOs).
- Provide executives with examples of what this process might achieve (moving towards data-driven decision making, making better use of funds, reducing risks), let them know that it could be implemented at different scales; not necessarily a major effort.
- Ensure that participants understand what they will get out of the process – ensure that they have an incentive to spend the amount of time that will be required.
- Include separate management and technical sessions, then a joint session where they provide each other feedback.
- The screening/scoping process should be included to help agencies identify where to focus – e.g., target problems that are causing the biggest issues. Agencies could use a combination top down and bottom up approach to do this. The bottom-up part could involve self-reporting to describe key gaps and their impacts.
- The data inventory activity should be designed so that it can be assembled without much effort – keep it general. Provide list of data categories that are broad enough for consideration of gaps in existing data. Include examples of data types belonging in each category.
- Keep in mind that participants may not be able to answer some of the assessment questions without going through a data mapping exercise -- which could be a substantial effort in and of itself. However, the results of a mapping exercise would provide participants with a common understanding of where and how data used in business processes.

Assess Step

- The assessment should go beyond measuring what the current state is; it should also provide insight into why things are the way they are.

- Set up a framework to show data alignment with substantive policy or program areas – e.g. financial effectiveness, asset management, performance management.
- It would be useful to do assessment at the business area level.
- The assessment can be used to communicate data issues to upper management – i.e., what being data-centric entails.
- Keep in mind that there are multiple perspectives that need to be represented and integrated: job/role, work unit, division, agency internal, external customer. On the questionnaires, specify which perspective the individual providing ratings should be taking.
- Questionnaires will need to be framed differently for particular data program owners vs. section managers responsible for multiple programs. Also, answers will vary based on what specific type(s) of data within a program area you are talking about.
- Communicate the purpose of the assessment to those completing the questionnaire.
- Include hyperlinks within assessment questionnaires that provide respondents with additional explanatory information.
- Be sure to include a comment block so that people completing questionnaires can explain their answers.
- Consider setting aside a specific time for people to complete questionnaires to ensure focused attention and timely responses.
- Consider using an independent reviewer (rather than self-reporting by business or data units) for some areas to provide an objective view.
- Consider continuous scale for ratings in the questionnaire (e.g. percentage confidence in data quality).
- Gap analysis should help agency to set priorities for what is most important to focus on. Keep in mind that many of the needs will be expensive and time-consuming technology projects – difficult to figure out where to start and how much to take on at once. Factors might include: program size, data criticality, dependencies (e.g. roadway data provide foundation for other data programs).
- Consider including customer satisfaction surveys – either as part of the assessment or as one of the techniques to consider to move up on the maturity scale.
- Use focus groups/round tables to validate questionnaire results and arrive at consensus ratings and discuss targets; this was a successful approach based on the safety data assessment.
- Don't be overly prescriptive about how to set targets, just provide guidance on factors to consider.
- Perhaps use a SWOT analysis to establish focus areas for improvements.
- It is important to bring information on benefits and costs of closing gaps into the target setting process.

Improve Step

- The assessment process needs to provide agencies with more than an understanding of what the gaps are – it should ideally include guidance, examples for improvement –

benchmarks from other agencies. Gaps identified should lead to improvements for consideration.

- This step needs to achieve executive buy-in for improvements, involve a discussion of how improvements are to be resourced.
- Be realistic about what can be achieved; don't try to do too much. Caution about trying to architect the perfect solution; go for "easy wins" and lightweight solutions.
- Recognize that often the highest priority initiatives from a need perspective involve major technology investments that are difficult to fund and implement.
- Produce an action plan with owners and timeframes – but consider working at the strategy level, and delegating detailed tactical planning so as not to bog down the effort.
- Include overall monitoring responsibility and plan to check in every year to see what has been accomplished and revise actions. Emphasize cyclical nature of process.
- If possible, integrate actions into business plan rather than creating a separate document and tracking process.
- Ensure that the action plan meshes with existing strategic planning efforts in the agency.

APPENDIX B – CASE STUDY RESULTS

Chapter 3 summarized the data self-assessment case studies conducted at Michigan DOT and Utah DOT, focusing on the process that was followed and the feedback that was received. Substantive results of the two pilots are provided below. This material was used in the development of the Assessment Guide, provided as a separate Volume to this Final Report.

Michigan DOT

Mobility/Congestion Data Value Assessment

The Mobility/Congestion Team consisted of both users and producers of travel-related data. This group focused on the adequacy of existing travel data to meet the needs of key consumers of the data. It included people representing asset management, pavement design, and safety – in addition to people with planning and congestion management responsibilities. The group also included people who might be classified more as data producers rather than users – though as was noted in the early focus groups for this project, the lines between data users and producers at DOTs are often not clear cut.

The Data Value Assessment asked respondents to rate the value of available data sources for conducting specific business activities. Before the questionnaire was distributed, configuration was required to identify data sources and business activities related to each group. The configured business activities and data sources for the MDOT mobility and congestion team are shown below in Table B.1.

Table B.1. Michigan DOT Congestion/Mobility Data Value Activities and Data Sources

Business Area Activity Definition		
For each standard category, identify a specific activity for this business area		
ID	Business Category	Purpose of the Activity
A1	Monitor results or performance against established objectives	Transportation system performance monitoring
A2	Track expenditures, resources used and accomplishments	Scoping and design of candidate projects
A3	Assess future needs – for budgeting or lining up new/different resources	Corridor and long-range planning, multi-modal planning
A4	Diagnose root causes for limited performance or inefficiencies	Real-time traffic and incident management
A5	Plan, prioritize or schedule actions to be taken	Improvement program development/prioritization

Data Source Definition		
Identify up to 10 data sets that you believe are the most important sources for this collection of activities		
ID	Data Program or System	Data Set(s)
S1	Road Inventory	LRS & Road Characteristics
S2	Traffic Monitoring	Counts, AADT, Classification
S3	Bike/Pedestrian	Bike Routes, Paths, Counts
S4	Real Time Travel	Current volume/occupancy, travel time
S5	Capital Program	Construction project scope and status

Assessment Results

Eleven team members completed the questionnaire. Individual questionnaire results were compiled and shared at the meeting. There were substantial variations across participants in ratings due in part to varying interpretations of questions, and in part due to varying perspectives and levels of involvement in different activities. The group was able to come to a set of consensus ratings, which are shown below in Table B.2.

Table B.2. Michigan DOT Congestion/Mobility Data Value Team Consensus Ratings

Activity	Summary Rating	Scores: 1=Poor, 2=Fair, 3=Good, 4=Excellent		
		Availability	Quality	Usability
A1. Transportation system performance monitoring	D-Sufficient	3.0	2.6	2.5
A2. Scoping and design of candidate projects	D-Sufficient	3.0	3.0	2.8
A3. Corridor and long-range planning, multi-modal planning	D-Sufficient	3.0	2.8	3.0
A4. Real-time traffic and incident management	B-Limited	2.0	2.9	2.9
A5. Improvement program development/prioritization	D-Sufficient	3.0	3.0	3.0

The team validated that with the exception of real-time traffic and incident management, data availability, quality and usability were “sufficient”. Real-time traffic and incident management received a lower rating due to limited geographic coverage, and difficulty of making use of the real-time data that are collected.

Gaps and Consequences

Table B.3 summarizes the gaps and associated consequences of these gaps as articulated by team members.

Table B.3. Michigan DOT Congestion/Mobility Data Value Team Gaps and Consequences

Gap Description	Consequences/Impacts
Missing traffic, pavement and roadway inventory data coverage for non-state maintained public roads.	<ul style="list-style-type: none"> • Impacts MDOT’s ability to meet federal reporting requirements for HPMS, MAP-21 and Safety (the MIRE Fundamental Data Elements) – if not met, could jeopardize federal funding. • Additional data would provide opportunity to answer questions about road performance comparisons from state legislators. • Additional data would provide opportunity to better plan work zones and detours based on an understanding of off-system capacity and utilization.
Insufficient speed data coverage between sensors (outside of urban areas)	<ul style="list-style-type: none"> • Note: not a priority gap – MDOT continues to monitor evolving technologies for speed data collection.
Need for additional permanent traffic counting sites for volume and classification on federal aid roads not under state jurisdiction.	<ul style="list-style-type: none"> • Meet Traffic Monitoring Guide requirements for seasonal factors (currently use factor of 1) • Improve quality of traffic modeling outputs used for corridor and long-range planning • Support MPO and local agency needs for understanding of traffic flows on local networks.
Need for greater availability of weekend traffic counts – almost all of the current counts are all done on weekdays.	<ul style="list-style-type: none"> • Meet Traffic Monitoring Guide requirements for weekend traffic data. • Improve information available for incident response, and MOT planning (construction is primarily taking place nights and weekends) and operations planning. • Improve quality of traffic assignment models – in many cases weekday volumes are similar; weekend travel patterns are the major distinguishing factor across different road sections.
Need for greater availability of traffic count data for seasonal/special events (e.g. festivals)	<ul style="list-style-type: none"> • Improve information available for incident response, detour planning for special events, and project scoping and design.
Lack of real-time traffic information off of the state system.	<ul style="list-style-type: none"> • Improve ability to identify bottlenecks and hotspots as part of congestion management planning • Provide ability to identify and evaluate effectiveness of detour routes • Improve availability of real-time traffic condition information to travelers • Provide improved basis for program development and prioritization of candidate projects prioritization (based on bottleneck locations)
Need for expanded incident data	<ul style="list-style-type: none"> • Meet new FHWA Traffic Monitoring requirements for incident data • Provide a basis for improving traffic data quality via comparison of volumes against historical incident information • Improve incident management capabilities
Need for additional information on bike and pedestrian volumes	<ul style="list-style-type: none"> • Meet new Traffic Monitoring Guide requirements for bike/pedestrian counts • Provide information to improve bike and pedestrian planning • Provide improved basis for prioritizing safety and facility improvements for bikes and pedestrians

Gap Description	Consequences/Impacts
Road inventory data not detailed enough for bike/pedestrian planning: shoulder type/width, sidewalk location, parking locations – for local system plus level of aggregation on state system	<ul style="list-style-type: none"> Assessing sufficiency of facilities for pedestrian and bicycle use
Ability to use ITS data for volume – ability to validate for this purpose; account for detector malfunctions	<ul style="list-style-type: none"> System performance monitoring – opportunity to increase coverage of volume information; potential to save \$\$ on traffic counting and improve data quality.
Time lag in availability of data after collection – need information on (final AADT), crash, and inventory for annual network surveillance (May)	<ul style="list-style-type: none"> Transportation system performance monitoring- end up using prior year’s data
No “final” open to traffic date included in project programming and scheduling data	<ul style="list-style-type: none"> Improvement program and project prioritization – want to remove locations from analysis, also helps with traffic data validation.
Insufficient information in project descriptions to know if bike/pedestrian considerations are included	<ul style="list-style-type: none"> Impacts accuracy and level of effort for preparation of required Act 51 year end summary and accomplishment reports (Complete Streets) Act 51 requires 1% of state funding on bike/pedestrian improvements.
Lack of central location for accessing traffic data – e.g. need to contact locals for data.	<ul style="list-style-type: none"> Increases data compilation burden for transportation system performance monitoring People need to go to multiple places – MPO, counties, etc. Customer service issue.
Availability of tools for accessing and integrating road inventory data (for some potential users) - ability to combine traffic, inventory, etc. Sufficiency file, RoadSoft data	<ul style="list-style-type: none"> Increases data compilation burden for transportation system performance monitoring Limits data value and usability for field staff
Data on facility condition/sufficiency are spread across multiple repositories and in different formats: PMS, IMS, sufficiency file, etc.	<ul style="list-style-type: none"> Increases data compilation burden for transportation system performance monitoring

Gap Description	Consequences/Impacts
Difficult to extract and use data from construction plans to populate/update asset and system inventory data	<ul style="list-style-type: none"> • Missed opportunity to save money on data collection by leveraging existing data sources

Priority Actions

The following priority actions were identified to address the identified gaps:

- Start doing seven day or weekend volume (and possibly class) counts – as existing staff/equipment allows; priority in areas where there are issues in identifying clusters and factors.
- Place permanent traffic recorders on a sample of local roads (e.g. five) to obtain clustering characteristics for seasonal adjustments. Work with Asset Management Council to identify representative sites.
- Add Weigh in Motion sites on local roads (one to three locations) – to identify pavement improvement strategies for local governments
- Incorporate probe data into the Regional Integrated Transportation Information System (RITIS) - expanding to non-freeways; assess quality and validate reliability
- Collect Fundamental Data Elements (FDEs) on Federal Aid roads – involve Asset Management Council; explore use of student interns for this.
- Develop plan to obtain improved inventory information for non-motorized infrastructure (e.g. sidewalk locations, shoulder type/width) from existing data sets/project plans
- Modernize the Traffic Data Management System (TDMS) – provide central location for local traffic detection data in regions
- Offer training to staff on how to access and integrate data from various sources at MDOT
- Expand short-term non-trunkline federal aid traffic data collection – volume + additional commercial data to meet federal requirements for factoring

Facilities Management Data Value Assessment

The **Facilities Management Team** members were from a business unit responsible for facility assessment, scoping, design, construction and maintenance. The facilities management unit manages 400 buildings, 14 welcome centers and 67 rest areas. The unit oversees a facility inspection process – buildings are assessed on a three year cycle. Facilities data are managed in five databases/spreadsheets for 1) inventory and inspection data; 2) financial data; 3) project planning; 4) photos; 5) work order tracking.

The configured business activities and data sources for the MDOT facilities data value team are shown below in Table B.4. For this group, area-specific labels fit well with the intent of the standard categories.

Table B.4. Michigan DOT Facilities Management Data Value Activities and Data Sources

Business Area Activity Definition		
For each standard category, identify a specific activity for this business area		
ID	Business Category	Purpose of the Activity
A1	Monitor results or performance against established objectives	Track facility inventory and condition (Includes both buildings and system/components).
A2	Track expenditures, resources used and accomplishments	Track facility capital and maintenance expenditures and work accomplishment
A3	Assess future needs – for budgeting or lining up new/different resources	Identify candidate projects for rehabilitation, replacement and expansion/addition
A4	Diagnose root causes for limited performance or inefficiencies	Diagnose causes of high maintenance costs or inefficiencies
A5	Plan, prioritize or schedule actions to be taken	Prioritize candidate projects and develop resource-constrained improvement program
Data Source Definition		
Identify up to 10 data sets that you believe are the most important sources for this collection of activities		
ID	Data Program or System	Data Set(s)
S1	Facilities	Facility Inventory
S2	Facilities	Facility Inspection/Condition
S3	Facilities	Maintenance Records
S4	Facilities	Facility Improvement Program
S5	Financial	Budgets and Expenditures

Assessment Results

Table B.5 shows the summary assessment results for the Facilities Team.

Table B.5. Michigan DOT Facilities Data Value Team Consensus Ratings

		Scores: 1=Poor, 2=Fair, 3=Good, 4=Excellent		
Activity	Summary Rating	Availability	Quality	Usability
A1. Track facility inventory and condition (Assume: includes both buildings and system/components).	1-Initial	1.0	2.2	1.0
A2. Track facility capital and maintenance expenditures and work accomplishment	1-Initial	1.0	1.6	1.0

		Scores: 1=Poor, 2=Fair, 3=Good, 4=Excellent		
Activity	Summary Rating	Availability	Quality	Usability
A3. Identify candidate projects for rehabilitation, replacement and expansion/addition	2-Limited	2.0	2.0	1.0
A4. Diagnose causes of high maintenance costs or inefficiencies	1-Initial	1.0	1.0	1.0
A5. Prioritize candidate projects and develop resource-constrained improvement program	2-Limited	2.0	1.9	1.0

Gaps and Consequences

Table B.6 summarizes the gaps and associated consequences of these gaps as articulated by team members.

Table B.6. Michigan DOT Facilities Data Value Team Gaps and Consequences

Gap Description	Consequences/Impacts
No data on building components (e.g. HVAC) are available. (Capabilities of current data systems and tools limit ability to capture and store this info.)	<ul style="list-style-type: none"> Forced to manage maintenance & capital improvements reactively rather than pro-actively. No basis for preventive maintenance planning and scheduling; foregoing potential savings from improved life cycle management. Limited ability to track warranties in effect and manufacturer's updates – increases risk of voiding existing warranties. Without component data available it would be difficult to know what needed to be replaced if a disaster occurred.
Insufficient level of detail on historical maintenance costs and accomplishments. It is not possible to track expenditures by different types of component improvements.	<ul style="list-style-type: none"> Improved decisions about replacement of components – may spend more money on repairs than it would to replace. Lack of information on nature of repairs, causes of repairs, repeat work on individual components, extent of repair work covered under warranties, cost of particular repair types – limits diagnostics that could be used to identify potential cost savings. (see IT call for projects)
Lack of data on the replacement value of facilities (lower priority)	<ul style="list-style-type: none"> Impacts quality of decision making regarding facility capital investments and maintenance.
Can't currently update asset condition data based on work completed. Some of the data from work done by regions is not captured.	<ul style="list-style-type: none"> Reports on asset condition don't reflect improvements made since the inspection. Makes a stronger case for improvement needs if the data are more current. Limits ability to identify needs & prioritize candidate projects. Opportunity to inspect less frequently and save money.
Need better ability to capture and store photos.	<ul style="list-style-type: none"> Limits ability to ensure that work is carried out as intended.

Gap Description	Consequences/Impacts
Current form for maintenance allows work to be assigned and prioritized, but doesn't allow for efficient work tracking and reporting after the fact	<ul style="list-style-type: none"> • Lack of ability to track and report on work done. • Extensive manual effort needed to summarize the information. • Lack of trend information.
Insufficient level of detail and consistency of data on maintenance and capital project needs - related to lack of facility component data	<ul style="list-style-type: none"> • Currently preparing the "needs" database is labor-intensive. • Opportunity to provide more accurate basis for needs estimates by having existing component (and related repair) info.
Facility maintenance records are not complete, detailed or reliable enough to track condition, (e.g. available systems don't provide sufficient structure to ensure completeness).	<ul style="list-style-type: none"> • Limits ability to understand existing facility condition.
There is a need for continued improvements to inspector training and guidance to facilitate consistent ratings by inspectors.	<ul style="list-style-type: none"> • Limits confidence in current facility condition and facility condition trends
Data systems and tools were developed in-house. Data format and query capabilities are very much a manual process and can be difficult to navigate without a high degree of knowledge.	<ul style="list-style-type: none"> • It is a manual and time-consuming process to pull information together to identify priorities and prepare improvement programs. Staff resources who know how and are capable of doing this effort are limited and put it at high risk.
Data in current systems are not easily shared, accessed or understood within the Facilities group. Additional knowledge management/transfer, skill building and succession planning are needed. Access by external users not feasible due to lack of security provisions in existing system.	<ul style="list-style-type: none"> • Risk of future information loss or data misuse

Priority Actions

The following priority actions were identified to address the identified gaps:

- Pursue an information technology (IT) project to develop, procure or implement an enterprise facilities management system
- Gather business requirements – as part of the facilities management system development process
- Have staff write reference manuals on how to use existing systems for reporting

- Identify the major facility components and key attributes (biggest cost and risk items)

Real Estate Data Management Assessment

The Real Estate Team members were from a business unit responsible for property acquisition in conjunction with highway projects, and management of MDOT-owned parcels. Team members were responsible for managing several data systems including:

- RESALE data base for excess property management.
- REMIS – a large legacy system that manages parcel data from acquisition to disposal.
- FileNet -- a document repository (e.g. stores deeds); used to manage workflow.
- ProjectWise -- a department wide system used for engineering content management and team collaboration.

Assessment Results

The team’s assessment results are summarized in Table B.7.

Table B.7. Michigan DOT Real Estate Data Management Team Assessment Results

	1-Initial: Ad-Hoc	2-Aware: Recognize Need for Improvement	3-Defined: Developed & Documented Processes	4-Proficient: Processes Implemented as Defined	5-Optimizing: Strategies Evaluated and Improved
DATA GOVERNANCE					
Data Stewardship	X				
Understanding of Costs and Value	X Cost			X Value	
Customer Relationships	X				
Data Preservation				X	
Findability/ Documentation		X			
DATA ARCHITECTURE AND INTEGRATION					
Data Accessibility/ Usability			X Other		X ProjectWise
DATA QUALITY					
Data Quality Assurance	X interaction with users		X QA/QC processes		
Accuracy	X standards	X understand/			

	1-Initial: Ad-Hoc	2-Aware: Recognize Need for Improvement	3-Defined: Developed & Documented Processes	4-Proficient: Processes Implemented as Defined	5-Optimizing: Strategies Evaluated and Improved
		quantify & improve			
Completeness	X standards	X understand/ quantify & improve			
Currency	X standards, processes to improve	X understand/ quantify			
Consistency			X		

As can be seen in the table, in some instances multiple levels were assigned to a given element due to variations in current practices for different data subsets, or variations in levels implied by individual questions related to an element. These results were used to refine the final set of maturity level descriptions.

Gaps and Associated Consequences

Table B.8 summarizes the gaps and associated consequences identified by the team during the second assessment workshop.

Table B.8. Michigan DOT Facilities Data Management Team Gaps and Consequences

Gap Description	Consequences/Impacts
Lack of single point of accountability across all of the real estate-related data systems	<ul style="list-style-type: none"> Lack of data standardization, lack of understanding about how to access the data. Potential for data misuse or under-use if there isn't a clearly identified point person who understands what data are available, their limitations and their uses. Having an identified data steward (as opposed to a system business owner) becomes even more important as data are shared across systems, and as experienced staff turns over.
There may be a lack of executive-level understanding about the importance and value of real estate data - though GIS has helped to elevate the importance of Right of Way (ROW) data	<ul style="list-style-type: none"> Results in under-investment in RE data/systems. ROW maps are out of date, RE data systems are out of date and require duplicative data entry across forms.
Lack of ongoing outreach to understand current and potential future data user needs. (Outreach is conducted as part of requirements gathering for system improvement or replacement. More proactive, ongoing outreach is difficult given resource/time constraints)	<ul style="list-style-type: none"> Lack of proactive outreach may result in underutilized data.

Gap Description	Consequences/Impacts
Data value not well understood by offices with data entry responsibilities.	<ul style="list-style-type: none"> • Data quality can be impacted when those entering data are not using it and don't have a full understanding of how others are using it.
For some systems, mechanisms for data customer feedback/input are ad-hoc rather than proactive.	<ul style="list-style-type: none"> • Lack of proactive outreach may result in underutilized data.
Department policies for archiving data exist but focus on paper records and are not uniformly understood. Archiving and retention policies haven't been established for ROW data based on understanding of customer requirements for retrieval of both project and program-level data. (Exception: ProjectWise)	<ul style="list-style-type: none"> • Inconsistent interpretations and implementations of existing policy • Risk that important historical information may not be easy to retrieve – particularly at the program (as opposed to project) level • Older microfilm records are starting to degrade – risk of permanent information loss
Updates to data dictionary information including data definitions are needed. (Data dictionary for REMIS exists, but may not always be updated as fields change.)	<ul style="list-style-type: none"> • As existing RE staff retire, understanding of data definitions may be lost.
RE Data are not available at an aggregate level – partially due to data quality issues as well as data aggregation and reporting capabilities.	<ul style="list-style-type: none"> • It is difficult to get a system, regional or broader view of the overall real estate picture – e.g. understand total square footage/acreage owned by MDOT.
Lack of integration across data sources.	<ul style="list-style-type: none"> • Time consuming to obtain data in usable form; impedes use by casual users.
Inefficient/Duplicate data entry – older system not set up to auto-populate fields based on prior entries	<ul style="list-style-type: none"> • Inefficient use of staff time for data entry tasks.
Lack of ability to enter and update data from mobile devices.	<ul style="list-style-type: none"> • Opportunity to improve staff efficiency if there was an ability to enter information about a parcel from a field device (and then print immediately and provide to property owner.)
Haven't standardized geographic referencing to enable mapping and integration with other agency data – involves addressing accuracy issues as well as differences in how locations are specified. (Note: Looking at adding spatial indexing to ProjectWise)	<ul style="list-style-type: none"> • Opportunity to integrate ROW data with other data in GIS platform and improve accessibility and therefore data value – e.g. would like ability to call up all relevant documents within a selected location scope.
Limited definition and standardization of data validation rules in REMIS.	<ul style="list-style-type: none"> • Risk of unreliable data; additional staff time required for data cleaning and quality assurance.
Limited quality assurance standards and processes for REMIS	<ul style="list-style-type: none"> • Risk of unreliable data; additional staff time required for data cleaning and quality assurance.

Gap Description	Consequences/Impacts
Limited quality assurance standards and processes for Resale	<ul style="list-style-type: none"> Risk of unreliable data; additional staff time required for data cleaning and quality assurance.
Timeliness issues – because the system is slow and frustrating, people wait until the end of the acquisition process to enter information due to frustrating data entry experience. New system that captures data as part of workflow is needed.	<ul style="list-style-type: none"> Providing an improved user experience and capturing data as part of the workflow will improve data quality.
Data completeness issues in REMIS below the general parcel level data. Related to frustrating data entry process.	<ul style="list-style-type: none"> Impacts ability to meet FHWA reporting requirements when data are not complete. Increases level of staff time needed for report preparation.

Priority Actions

The following priority actions were identified to address the identified gaps:

- Develop or purchase a new system to replace REMIS and RESALE
- Standardize systems, eliminate redundancies, remove non-critical fields (if first action can't be done) – with data controls to ensure complete entry of these fields
- Train staff on new simplified version of REMIS
- Develop formal policies and procedures for data archiving and preservation
- Provide GIS access to parcel data (to increase visibility of this data)
- Define data validation rules for each column (in conjunction with system replacement)
- Establish a data stewardship group or community of interest

Aeronautics Data Management Assessment

Aeronautics team members were from a business unit responsible for managing the state's 238 public use airports and heliports. Functions include capital programming; project management; environmental review; tall structure permitting; inspection, licensing and/or registering airports, flight schools, aircraft, and aircraft dealers; and state and federal reporting. Aeronautics data programs serve a diverse customer base including FHWA, airport managers, pilots, cities, MPOs, and internal Intermodal Management System (IMS) and Finance data consumers. Aeronautics has several major data systems including:

- ASM (Airport System Manager) – a COTS application that includes information on the physical inventory features of airports, as well as inspection, licensing and permitting data. Inventory and Inspection data being managed in ASM meets FAA reporting requirements – it flows through FAA standardized reporting structures via the GCR software.
- AeroPM – includes data and information on airport projects
- The Intermodal Management System (IMS) - part of MDOT's overall Transportation Management System (TMS). The IMS manages statewide modal data (including

physical inventory and usage) for connectivity and long-range planning. IMS is fed data by ASM.

Assessment Results

The team’s assessment results are summarized in Table B.9.

Table B.9. Michigan DOT Aeronautics Data Management Team Assessment Results

	1-Initial: Ad-Hoc	2-Aware: Recognize Need for Improvement	3-Defined: Developed & Documented Processes	4-Proficient: Processes Implemented as Defined	5-Optimizing: Strategies Evaluated and Improved
DATA GOVERNANCE					
Data Stewardship	X No single individual has overall responsibility				X Specific data area/ business process maturity is high
Understanding of Costs and Value	X Understand Cost				X Understand Value
Customer Relationships			X		
Data Preservation			X		
Findability/ Documentation		X ASM	X AeroPM	X IMS	
DATA ARCHITECTURE AND INTEGRATION					
Data Accessibility/ Usability					X some mobile data on airports for pilots and airport managers
DATA QUALITY					
Data Quality Assurance		X General quality (but varies by data system)		X QA/QC processes	
Accuracy		X Understand accuracy	X Standards for accuracy + efforts to improve accuracy		

	1-Initial: Ad-Hoc	2-Aware: Recognize Need for Improvement	3-Defined: Developed & Documented Processes	4-Proficient: Processes Implemented as Defined	5-Optimizing: Strategies Evaluated and Improved
Completeness		X Understanding – for ASM	X Standards and targets for completeness + efforts to improve completeness	X Understanding of completeness for IMS, inspection and Aero PM project data	
Currency			X Targets are set by others or internally -- but still work to do on getting agreement	X understand or know currency + efforts to improve	
Consistency			X		

Gaps and Associated Consequences

Table B.10 summarizes the gaps and associated consequences identified by the team during the second assessment workshop.

Table B.10. Michigan DOT Aeronautics Data Management Team Gaps and Consequences

Gap Description	Consequences/Impacts
Lack of formal (role-based) definition of accountability and responsibility for data quality for each type of aeronautics data	<ul style="list-style-type: none"> • Potential for data quality issues as data are moved across systems. • Things work now due to informal relationships and collaboration and strong system stewards. When existing key staff retire, there are risks that data quality will decline due to lack of formalized accountability.
Limited ability to fully understand and communicate the value proposition for aeronautics data	<ul style="list-style-type: none"> • As functions are automated, it is important to know how much time is spent managing versus analyzing data – savings from centralization of data
There may be a lack of executive understanding about the importance and value of aeronautics data – including the business processes they support.	<ul style="list-style-type: none"> • If executives don't understand how the data are being used, it may result in lack of priority or attention to aeronautics data issues.

Gap Description	Consequences/Impacts
There is a lack of defined (written) protocols for sharing, using and publishing aeronautics data beyond IMS. (Data of general interest is in IMS, which has established proactive customer relationships.)	<ul style="list-style-type: none"> Not a high priority gap – but may result in data being published/shared without data steward review to ensure accuracy.
No documented business requirements or policies for retention and access to historical data	<ul style="list-style-type: none"> Some data customer needs may not be met; more research would be required to identify what these needs are and what the impacts might be.
There are some holes in data dictionary information for 3 rd party apps (ASM - GCR)	<ul style="list-style-type: none"> Potential for misinterpretation of data elements.
No formal change notification process from DTMB to IMS when ASM (COTS system) code changes occur – impacts IMS update scripts	<ul style="list-style-type: none"> Some delay in IMS updates from ASM, may result in perception of lack of IMS data currency.
Need to further automate process to assign LRS reference to airport points of access. Airports are located based on lat-long (for pilots), and PR number (for drivers)	<ul style="list-style-type: none"> Opportunity for minor improvement in data management efficiency and currency/completeness of location info needed for mapping. (Not a significant concern.)
Data quality assessment and improvement – ASM Facility and communication to IMS	<ul style="list-style-type: none"> More proactive approach could enhance data use and therefore value
Data quality assessment and improvement – ASM Aircraft Registration	<ul style="list-style-type: none"> More proactive approach could enhance data use and therefore value
Data quality assessment and improvement – ASM Tall Structures	<ul style="list-style-type: none"> More proactive approach could enhance data use and therefore value
Data quality assessment and improvement – AeroPM	<ul style="list-style-type: none"> More proactive approach could enhance data use and therefore value

Priority Actions

The following priority actions were identified to address the identified gaps:

- Create document with business requirements for historical data
- Populate Aeronautics data element dictionary entries in InfoSphere
- Document who (role and person) is responsible for each type of data – for all Aeronautics data entities (includes determining appropriate level of granularity for data entities)
- Create process documentation for maintaining each type of data (possibly following IMS model - checklists) - for each module of ASM + one for AeroPM. (can be combined with item 1 above)

Agency-Wide Data Management Assessment

The agency-wide data assessment team consisted of the members of the Michigan DOT’s Data Governance Council. This group was established to address data management throughout the department. The group who participated in the workshop included representatives of both business and data management functions in the Department, as well as several individuals from the Michigan State Department of Technology, Management and Budget (DTMB).

Assessment Results

The team’s assessment results are summarized in Table B.11.

Table B.11. Michigan DOT Agency-Wide Data Management Team Assessment Results

	1-Initial: Ad-Hoc	2-Aware: Recognize Need for Improvement	3-Defined: Developed & Documented Processes	4-Proficient: Processes Implemented as Defined	5-Optimizing: Strategies Evaluated and Improved
DATA GOVERNANCE					
Data Ownership Stewardship		X Approach + Processes for adding or changing data	X Support + Current practices – some may be level 4		
Data Governance (Q7-11)		X Governance principles in place + Implementation of standards + Change management		X Governance structure in place+ Current practices	
Understanding of Costs and Benefits		X Data costs + Use and value for primary customers + Plans and budgets	X New initiatives		
Customer Relationships	X-no written SLAs)			X - for some programs, e.g. MAP, Pavement	
Findability/ Documentation	X– no catalog	X		X – data dictionaries using InfoSphere	

	1-Initial: Ad-Hoc	2-Aware: Recognize Need for Improvement	3-Defined: Developed & Documented Processes	4-Proficient: Processes Implemented as Defined	5-Optimizing: Strategies Evaluated and Improved
Data Preservation	X – some core systems not in enterprise databases		X – retention policies but focused on paper		
Data Security			X– not fully tested for some databases	X – for enterprise databases	
DATA ARCHITECTURE AND INTEGRATION [Group did not complete these – based on survey results]					
Geospatial Referencing				X (but large spread in responses)	
Data Architecture		X			
Data Integration		X			
Change Impact Analysis		X			
Data Collection Coordination & Data Sharing			X		
Data Usability		X (varies by data program)	X (varies by data program)		
DATA QUALITY [Group did not complete these – based on survey results]					
Data Quality		X (support QA processes and tools across programs)		X (standard database development practices)	

Gaps and Associated Consequences

Table B.12 summarizes the gaps and associated consequences identified by the team during the second assessment workshop.

Table B.12. Michigan DOT Agency-Wide Data Management Team Gaps and Consequences

Gap Description	Consequences/Impacts
Not all core data programs have identified business data stewards.	<ul style="list-style-type: none"> Without business data stewards, there is no mechanism for implementing data governance policies. There is also risk of poor data quality impacting business decisions.
Processes for ensuring that new data collection efforts are not duplicative or result in inconsistent data definitions and formats.	<ul style="list-style-type: none"> Without such processes, there is increased risk that poor data design decisions would be made (e.g. duplication), which would increase IT costs.
There is a need for additional definition of policies and standards, informed by a shared understanding of data governance processes and success factors.	<ul style="list-style-type: none"> Without a shared understanding of data governance, it is difficult to reach consensus on an appropriate set of processes that would best fit MDOTs needs.
Single authoritative sources have not been identified for some data programs/elements	<ul style="list-style-type: none"> Potential for “multiple versions of the truth”
Metrics for communicating value and success of data governance activities have not yet been determined. Data isn’t considered as part of the agency’s strategic plan.	<ul style="list-style-type: none"> Data governance is not sustainable unless its value is recognized and communicated.
Some of the agency’s core business data systems are not in enterprise databases.	<ul style="list-style-type: none"> Data aren’t discoverable or shareable There are risks of data loss because data may not be backed up.
Consistent documentation of data collection, updating and reporting cycles for most data programs is not in place. (Note: many data sets do have excellent documentation, but differences in collection and reporting methods across regions is not captured)	<ul style="list-style-type: none"> Users can’t necessarily understand the derivation of the data that might be available.
The full value (costs/benefits) of data programs are not consistently tracked.	<ul style="list-style-type: none"> Limits ability to communicate value or assess potential efficiency improvements.
There is limited training and succession planning to sustain staffing resources for key data programs. (Note: succession management challenges aren’t limited to data programs)	<ul style="list-style-type: none"> Risk of inefficiency and need to re-invent processes and systems due to loss of knowledge.

Gap Description	Consequences/Impacts
Data mapping has not been completed to identify dependent business processes for each data set; some data program managers do not have a good understanding of who their customers are.	<ul style="list-style-type: none"> Limits ability to know who might be impacted by changes or problems. Limits ability to understand full value of data and data quality requirements. May be opportunities to eliminate data sets/schemas that are redundant or not being used.
Lack of ability to tap into core data for assets that is stored in design plans.	<ul style="list-style-type: none"> Opportunity to make better use of existing data sources and avoid additional data collection costs.
Lack of processes for identifying what data exists and what its quality is.	<ul style="list-style-type: none"> Can't determine suitability for purpose
Feedback from users is not generally sought on adequacy of available data definitions.	<ul style="list-style-type: none"> Without meaningful definitions (confirmed by user feedback), users may not understand the data they are trying to use.
There is no agency-wide data catalog. (Note: catalog information exists for Oracle corporate databases, but not yet for other data.)	<ul style="list-style-type: none"> Employees may not know what data are available and how to access them.
Data retention and archiving policies based on business needs have not been established. (applies mostly to electronic data)	<ul style="list-style-type: none"> The default is keeping all data – which may mean that the agency is paying too much for storage. Without a conscious design process to provide access to historical data, it may be difficult to work with older data sets where coding has changed.
A change management process has been developed for the MAP database but has not yet been implemented for other systems. (Note: this is an area where a lot of progress has been made. Identification of data stewards for other systems will allow for full rollout of change management.)	<ul style="list-style-type: none"> Lack of change management can result in data modifications to meet one user requirement without consideration of how this may impact others. Lack of change management can cause system interfaces and reports to “break”.
While MDOT has implemented a standard LRS based on the state framework, geospatial data standards may not be consistently followed, and many applications are not spatially enabled. Some differences in projections exist between the Center for Shared Solutions [CSS] spatial data and MDOT spatial data.]	<ul style="list-style-type: none"> Lack of ability to map existing data or integrate data based on spatial referencing.

Gap Description	Consequences/Impacts
Guidelines for determining spatial accuracy requirements and appropriate collection methods have not been established. [Note: there is no presumption that all data must meet the most stringent accuracy requirements, but that there is an explicit understanding of the spatial accuracy needed for each business purpose, and the data collection devices and methods required to meet each accuracy level.]	<ul style="list-style-type: none"> • May limit ability to use data that are collected for multiple purposes. • Data may be collected without sufficient accuracy to meet their intended business purpose.
Master data has not yet been standardized and shared/synchronized across business units and applications	<ul style="list-style-type: none"> • Duplication of data and work – more effort required to maintain different systems • More dependencies across databases increases complexity of system maintenance and increases risk that data will get out of sync, potentially disrupting dependent business processes
There is no agency-wide data integration solution in place (e.g. data warehouse, enterprise service bus, etc.). Expertise and/or tools are not available 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 to meet specific business needs.	<ul style="list-style-type: none"> • Opportunity to reduce time needed to obtain answers to business questions. • Potential for erroneous data joins when users attempt to integrate data themselves. • Performance burden on transactional systems to support reporting needs.
Opportunities for coordinating data collection efforts across business units are pursued on an ad-hoc basis but there is limited enterprise-wide, proactive data collection coordination. (Note: MDOT is currently pursuing opportunities for coordinating data collection – e.g. LiDAR, sufficiency pilots)	<ul style="list-style-type: none"> • There may be opportunities for combining efforts and gaining efficiencies.
Agency employees do not have the ability to easily visualize trend information on asset condition, capital and maintenance expenditures, traffic, crash rates, and other important agency performance indicators.	<ul style="list-style-type: none"> • Agency may be optimizing decision making within silos but not across silos.
Agency field staff cannot access information about assets, projects or work orders on mobile devices	<ul style="list-style-type: none"> • Work is not reported on an activity level or asset/location specific • Lost opportunity to improve stakeholder communication • Lost opportunity to provide needed information to newer field employees (reduce need for knowledge transfer)

Priority Actions

The following priority actions were identified to address the identified gaps:

- Undertake an agency-wide effort to populate/maintain data definitions within InfoSphere
- Develop a data integration strategy concept document (e.g. data warehouse)
- Produce a high level business architecture that maps data to business processes
- Identify and assess risk of core data sets not currently housed in enterprise systems
- Address responsiveness to field data needs – to reduce motivation for maintaining separate data sets (related to risk assessment)
- Investigate opportunities to reduce the number of schemas
- Define process for improved communication and change management with respect to changes to data coding or requirements
- Develop a Training Strategy for data management to build common terminology and understanding of basic concepts in the DMBOK. Potential first step is an offsite meeting for the Data Governance Council
- Develop presentation to sponsors about data management goals, strategies and definitions
- Develop process for harvesting asset data from CADD files
- Develop records retention policy for electronic data (including delete policy)

Utah DOT

Maintenance Data Value Assessment

The **Maintenance Data Value Team** consisted of individuals from central and regional maintenance offices. The team also included the GIS manager for the department. The team represented producers and users of the data.

The Utah DOT Maintenance Data Value Assessment questionnaires identified broad standard categories of work that were then paired with the data sources and systems that help accomplish the activities. Table B.13 summarizes Utah DOT maintenance data value activities, data sources and data systems.

Table B.13. Utah DOT Maintenance Data Value Activities and Data Sources

Business Area Activity Definition		
For each standard category, identify a specific activity for this business area		
ID	Standard Category	Area-Specific Label (activity)
A1	Monitor results or performance against established objectives	Track maintenance level of service (outcomes) in Maintenance Management Quality Assurance (MMQA) system
A2	Track expenditures, resources used and accomplishments	Track maintenance expenditures, resources used and accomplishments (outputs) in Operations Management System (OMS)
A3	Assess future needs – for budgeting or lining up new/different resources	Develop future year maintenance budget requests

A4	Diagnose root causes for limited performance or inefficiencies	Identify opportunities for improved efficiency
A5	Plan, prioritize or schedule actions to be taken	Plan, prioritize and schedule work
Data Source Definition		
Identify up to 10 data sets that you believe are the most important sources for this collection of activities		
Source ID	Data Program or System	Data Set(s)
S1	Road Inventory	Road Inventory - Highway Referencing System
S2	OMS	Feature Inventory
S3	OMS	Maintenance Work Orders
S4	OMS	Funding Allocations
S5	Snow Plow Planning	Snow Plan
S6	Pavement Management System (dTIMS)	Plan for Every Section (PFES)

Assessment Results

The team’s assessment results are summarized in Table B.14.

Table B.14. Utah DOT Maintenance Data Value Team Assessment Results

		Scores: 1=Poor, 2=Fair, 3=Good, 4=Excellent		
Activity	Summary Rating	Availability	Quality	Usability
A1. Track maintenance level of service (outcomes) (MMQA)	D-Sufficient	3.0	3.0	2.8
A2. Track maintenance expenditures, resources used and accomplishments (outputs) (OMS)	D-Sufficient	3.0	3.4	3.8
A3. Develop future year maintenance budget requests	D-Sufficient	3.0	3.2	3.2
A4. Identify opportunities for improved efficiency	D-Sufficient	3.0	2.9	2.8
A5. Plan, prioritize and schedule work	D-Sufficient	3.0	3.0	3.1

Gaps and Associated Consequences

Table B.15 summarizes the gaps and associated consequences identified by the team during the second assessment workshop.

Table B.15. Utah DOT Maintenance Data Value Team Gaps and Consequences

Gap Description	Consequences/Impacts
MMQA – Need to determine adequacy of sampling (study underway with the University of Utah)	<ul style="list-style-type: none"> • Opportunity for better understanding of data collection needs for maintenance budgeting
MMQA – need to improve ability to address management questions about cost to achieve performance targets – covers 16 of 116 activities, relies on historical data. Activities have changed, so historic data not as useful. Many of the high cost elements of the maintenance budget aren't included.	<ul style="list-style-type: none"> • Better estimates of budget needed to improve level of service. But already spending 1% of budget on data collection. Useful to have information to make the case for what it costs to achieve, harder to tell story, more reactive approach.
MMQA - difficult to interface with dTIMS to update pavement quality elements. Need to tap into more specific pavement distress info related to what maintenance can control.	<ul style="list-style-type: none"> • Opportunity for improved efficiency; can't calculate an MMQA grade for hard surfaced pavements (ability to move data and calculate score); currently not able to integrate the info into maintenance budgeting process.
Want to link OMS and dTIMS. Want to push work planning info from ePM and OMS back into dTIMS. If state forces do a chip seal, this info gets entered into OMS as a work order, but doesn't get back into dTIMS PFES.	<ul style="list-style-type: none"> • Inefficient management of programs, loss of credibility with upper management, not efficient with data entry resources.
MMQA – alignment with feature inventory needs more work to use consistent units of measure and eliminate need for use of quantity overrides in MMQA (trust in data quality is also an issue) – disconnect for signs; guardrail is in OMS. Drainage collected using a separate effort (quality issues); overrides don't necessarily reflect feature inventory.	<ul style="list-style-type: none"> • Confusion, difficult to answer questions about feature inventory with confidence, hard to provide solutions.
Feature Inventory – some gaps in data; not updated based on work completed (including following construction projects), no ability to update in the field using mobile devices (tablet app under development)	<ul style="list-style-type: none"> • Loss of credibility when feature inventory doesn't match field observation • May impact ability to plan and prioritize work.
dTIMS – need more attention on validating condition assessment and treatment recommendations	<ul style="list-style-type: none"> • Loss of credibility when condition and recommendations don't appear reasonable

Gap Description	Consequences/Impacts
OMS – expenditures and accomplishments – tracked by station rather than more granular location (route-section)- snowplan info	<ul style="list-style-type: none"> • Impedes ability to analyze patterns of repeat work and correlate to other location-specific data • Lack of ability to tell local jurisdictions what maintenance costs are by section
OMS – need for improved reports targeted at users outside of Maintenance, titles not informative	<ul style="list-style-type: none"> • Difficult and time consuming to use, not getting full value from the information
Construction projects – need for more consistent current information about status. Design-build projects create most problems.	<ul style="list-style-type: none"> • It is sometimes difficult to find out the status of construction projects or whether a specific project has been delayed – important to know for scheduling of maintenance work, e.g. striping.
Need better ability to analyze information about how to optimize deployment of resources for snowplowing.	<ul style="list-style-type: none"> • Potential for cost savings based on understanding of past experience (e.g. which products work best) – move to a more data-driven, consistent approach.

Priority Actions

The following priority actions were identified to address the identified gaps:

- Develop process to get sign information into OMS Feature Inventory
- Train additional staff in how to use OMS; enhance support function
- Improve OMS reporting – streamline and simplify
- Require contractors to provide feature inventories – define data elements and format
- Update feature inventory database and tie to MMQA – reduce need to override totals – link to data updates in field
- Fix link between dTIMS and OMS – implement web-based Plan for Every Section (PFES)
- Build 2-way interface between dTIMS and OMS

Project Scoping and Design Data Value Assessment

The Project Scoping/Design Data Value Team included central and regional office preconstruction and design managers. Representatives from traffic operations and the Director of Business Information Technologies also participated in the meetings. Having all of these perspectives in the room strengthened discussion and enhanced results.

The focus of the team was on the data needed to manage projects, prepare design plans, and perform environmental reviews, review existing conditions (including as-built data) and draft concept plans. Table B.16 outlines how business area activities were defined for the Project Scoping/Design data value questionnaire and the principal data sources and systems associated with managing and accomplishing the work.

Table B.16. Utah DOT Project Scoping and Design Data Value Activities and Data Sources

Business Area Activity Definition		
For each standard category, identify a specific activity for this business area		
ID	Business Category	Purpose of the Activity
A1	Monitor results or performance against established objectives	Project Management/Project Control (tracking with Executive Dashboard)
A2	Track expenditures, resources used and accomplishments	Prepare Design Plans
A3	Assess future needs – for budgeting or lining up new/different resources	Environmental Review
A4	Diagnose root causes for limited performance or inefficiencies	Review Existing Conditions/Identify Needs (using As-Builts)
A5	Plan, prioritize or schedule actions to be taken	Create Concept Reports
Data Source Definition		
Identify up to 10 data sets that you believe are the most important sources for this collection of activities		
Source ID	Data Program or System	Data Set(s)
S1	Roadway Inventory	LRS & Road Geometry
S2	Traffic Monitoring	Counts, AADT, Classification
S3	UGate/UPlan	Asset Inventory, Concept Reports, Environmental Data
S4	Safety Management System	Crash and Fatality Data (OSR)
S5	ePM	Contract Management, Project Management
S6	ePM	ROW
S7	ePM	Executive Dashboard

Assessment Results

The team's assessment results are summarized in Table B.17.

Table B.17. Utah DOT Project Scoping and Design Data Value Team Assessment Results

		Scores: 1=Poor, 2=Fair, 3=Good, 4=Excellent		
Activity	Summary Rating	Availability	Quality	Usability
A1. Project Management/Project Control (tracking with Executive Dashboard)	D-Sufficient	3.0	2.9	2.9
A2. Prepare Design Plans	C-Basic	3.0	2.4	2.6
A3. Environmental Review	D-Sufficient	3.0	2.8	2.8
A4. Review Existing Conditions/Identify Needs (using As-Builts)	B-Limited	2.0	2.4	2.6
A5. Create Concept Reports	C-Basic	3.0	2.4	2.6

Gaps and Associated Consequences

Table B.18 summarizes the gaps and associated consequences identified by the team during the second assessment workshop.

Table B.18. Utah DOT Project Scoping and Design Data Value Team Gaps and Consequences

Gap Description	Consequences/Impacts
ROW Data: Would like to geo-reference and make available in UPlan (working on this – current projects and historical as resources allow)	<ul style="list-style-type: none"> Ability to understand where ROW boundaries are, help in scoping projects, help to interact with partners (utilities, environmental resource agencies), help maintenance, allow resident engineer to understand ROW status – identify which parcels haven’t been cleared and who is responsible for next step. Benefit to private survey also, ability to publish to Google Earth – improve transparency with public, not possible without geo-referencing.
As-Builts: Uneven availability and quality across regional offices; inconsistent application of standards and QA to ensure accuracy.	<ul style="list-style-type: none"> Difficult to understand underground infrastructure; 3D as-built would help to capture this. Need to re-discover. Could just validate instead at lower level of effort. Want to move from investigation to validation at the scoping process. Less potholing needs.
As-Builts: Can’t retrieve from UPlan (working on this – auto-generator, link from ProjectWise)	<ul style="list-style-type: none"> More convenient access – inefficiencies (from minutes-hours to seconds-minutes); ability to understand what is the latest.
Utility Data: Difficult to obtain up to date info.	<ul style="list-style-type: none"> Greater expense for subsurface exploration, potential change orders and construction delays. Less accuracy.
UPlan: May be need to improve user familiarity with capabilities and uses – new tools and features being added to enhance usability. Quality and availability issues, need to sustain it.	<ul style="list-style-type: none"> Potential benefits not being realized. Potential for loss of credibility if tools are released too quickly before audience is ready.

Gap Description	Consequences/Impacts
Crash: would like to improve timeliness – make available sooner. Issue primarily with local jurisdiction data – 1 year lag. Have near real time from UHP. Also locational accuracy has been off.	<ul style="list-style-type: none"> • Ability to better utilize available safety funds – based on most current information. • Ability to use system-level summary data (rather than anecdotal info) – would impact results of safety analysis (frequency and severity).
Crash: Need to improve ease of mapping and analysis; few individuals with skills to mine and analyze data (working to improve)	<ul style="list-style-type: none"> • Impacts efficiency of operational safety reviews
Executive dashboard – performance improvements (currently calculations are regenerated each time); queries directly against database rather than “trusted source” – to be addressed through data warehouse project	<ul style="list-style-type: none"> • Increased level of frustration for users; may inhibit use • May end up with inconsistencies between dashboard and ePM reports.
Traffic: Lack of QA/QC on TMC traffic detectors	<ul style="list-style-type: none"> • Need for experienced person interpreting data • Risk of using incorrect data
Traffic: Multiple sources of traffic data – not consistent, requires interpolation, need to call traffic unit for information – not “self-serve” (working on this – new self-serve capability almost complete)	<ul style="list-style-type: none"> • Staff time for traffic engineers to answer inquiries • May limit staff use of traffic information
LRS: not all region personnel know it exists or how to use it. There are some timeliness issues with getting LRS changes propagated into different business systems.	<ul style="list-style-type: none"> • Timeliness issues – business systems may be out of sync, creates inconsistent data.
Feature inventory not being updated frequently enough	<ul style="list-style-type: none"> • Potential loss of confidence in the system, impacting willingness to use it.
Still have isolated pockets of “stand-alone” information not on central GIS. Mostly a concern with project-level data (controlled out of regions). Also applies to agreement databases currently in MS Access	<ul style="list-style-type: none"> • Lack of ability to share data. • May have multiple datasets that are inconsistent, not sure what the most accurate or authoritative, authenticated source is. • May result in loss of confidence in reports where inconsistencies exist.

Priority Actions

The following priority actions were identified to address the identified gaps:

- Provide improved access to metadata about whether GIS data is the temporary or authenticated version, and date of last update.
- Provide more user-friendly ways of querying or browsing GIS data to make it more findable given increasing volume.
- Implement a retention schedule for GIS maps.

- Form working group to investigate how to update feature inventory based on permitting (might add turn lanes, etc.), maintenance, construction – consider: flagging asset inventory when a project has been done; use of static LiDAR (or phone LiDAR), extraction of 3D design data, creating pay item for contractors to provide update.
- Update uCAT (user cost calculations, maintenance of traffic planning) to include PeMS and signal performance data, ID and train additional expert traffic users.
- Populate traffic data into data warehouse to improve accessibility – include preliminary LOS calculations – help to ID hotspots and choke points.
- Improved data management for GIS data.
- Develop 3D utility management (in process).

Asset Management Data Management Assessment

The Asset Management Team consisted of individuals responsible for some aspect of asset-related data. Many of these individuals were also data users. UDOT asset management includes pavements, bridges, safety assets (MIRE), ITS elements, traffic data, roadway features, HPMS and other associated data programs needed to make investments and maintain existing facilities. Much of the asset data that UDOT manages is provided via a LiDAR-based data collection effort. Base map data is provided by an outside vendor and environmental data is derived from a variety of external sources.

Assessment Results

The team’s assessment results are summarized in Table B.19.

Table B.19. Utah DOT Asset Management Data Management Team Assessment Results

	1-Initial: Ad-Hoc	2-Aware: Recognize Need for Improvement	3-Defined: Developed & Documented Processes	4-Proficient: Processes Implemented as Defined	5-Optimizing: Strategies Evaluated and Improved
DATA GOVERNANCE					
Data Stewardship	X distributed stewardship model				
Understanding of Costs and Value			X other assets/features (understanding of value varies by data type)		X pavement & structures
Customer Relationships			X Understand user needs, but not written down		

	1-Initial: Ad-Hoc	2-Aware: Recognize Need for Improvement	3-Defined: Developed & Documented Processes	4-Proficient: Processes Implemented as Defined	5-Optimizing: Strategies Evaluated and Improved
Data Preservation		X specific requirements for historical data access need definition	X varies by program		X UGate, UPlan
Findability/ Documentation		X Traffic and crash data – on request rather than “self-serve”			X UGate provides access & documentation, asset & traffic data available via mobile apps, some APIs available
DATA INTEGRATION & ARCHITECTURE					
Data Accessibility/ Linkages		X more work needed to link different data sets			X data highly accessible via UGate
DATA QUALITY					
Data Quality Assurance		X varies based on type of data			X pavement
Accuracy		X for Mandli data set	X Varies based on data		
Completeness	X Standards for completeness	X Processes to improve completeness	X Understand completeness		
Currency		X Exploring different ways to improve	X Stewards understand and some users know timeliness, don't have governance for how to document and share timeliness info with users.		
Consistency			X Varies		

Gaps and Associated Consequences

Table B.20 summarizes the gaps and associated consequences identified by the team during the second assessment workshop.

Table B.20. Utah DOT Asset Management Data Management Team Gaps and Consequences

Gap Description	Consequences/Impacts
Accountability for asset data quality, value and appropriate use has not been formalized; lack of formal definition of stewardship responsibilities	<ul style="list-style-type: none"> • Trust in data suffers • Inconsistent levels of quality • Potential for reporting inaccurate information • Employee transitions may result in inconsistencies • Potential to not address all user needs
Need for improved ability to assess, document and communicate the value proposition for asset data (note that this may take time - e.g. cross-drainage data example)	<ul style="list-style-type: none"> • Improved ability to answer questions about whether the department is collecting the right data and whether they are getting full value from the data they collect. • Not being aware of the hidden costs to not collecting the data – can’t compare to cost of data collection. • Not being aware of “the cost of ignorance” – e.g. ability to do bike plan with awareness of assets. • Supports stewardship – ability to justify staff time to manage the data.
Potential need for additional tracking of data program costs.	<ul style="list-style-type: none"> • Ability to evaluate different approaches to data updating. • Ability to identify opportunities for improved efficiencies in data acquisition and management.
Need for more consistent documentation of data sets (how and when collected, how updated, locational accuracy, known quality issues) and communication of this information to data customers.	<ul style="list-style-type: none"> • Customers don’t understand age of data or lineage of data • People might not know who to contact to get information on a dataset • Potential for misunderstanding or misuse of data • Potential for data to not get full use
Need for additional data roll-out processes as new data comes on line to help customers understand how the data can be used to improve decision making and gain efficiencies in work processes	<ul style="list-style-type: none"> • May be collecting data that aren’t needed • Limits return on data investment • Can’t control the message
Few documented business requirements or policies for retention and access to historical data (to be addressed via data warehouse initiative)	<ul style="list-style-type: none"> • Increased staff time to retrieve data in response to requests • Limited ability to conduct safety analysis • Limited documentation of where assets once existed
Need for systematic data quality assessment and cleansing (Oracle EDQ tool acquired for this purpose)	<ul style="list-style-type: none"> • Limited ability to check for duplicates and inconsistencies • Limited ability to cleanse and transform source data sets for reporting and analysis • Limited ability to provide data validation rules to data collection contractors to improve data quality

Gap Description	Consequences/Impacts
Potential need for additional formal data quality assurance processes – including field spot checks of “ground truth”, comparison across data sets.	<ul style="list-style-type: none"> • Potential for improving data quality and trust in data using documented procedures
Need for better understanding of data customer needs for completeness, accuracy, currency – beyond what is needed for federal reporting.	<ul style="list-style-type: none"> • Missed opportunities to meet customer needs (without additional cost) • Missed opportunities to save money by reducing accuracy
Need for improved way of presenting information about currency of data elements in UGate	<ul style="list-style-type: none"> • People may make incorrect assumptions about data accuracy
Need for improved ability to link data on asset condition with data on work completed	<ul style="list-style-type: none"> • Provide ability to update condition based on work completed, or at least flag that the condition info is no longer accurate • Lack of trust in data; potential for duplicate data collection • Lack of ability to understand treatment effectiveness
Need for additional documentation of data definitions – as they evolve over time.	<ul style="list-style-type: none"> • People may make incorrect assumptions about data meaning

Priority Actions

The following priority actions were identified to address the identified gaps:

- Document accuracy and QA process for different data elements
- Update metadata descriptions for traffic monitoring stations & formalize procedures for de-commissioning old stations
- Form Data Governance Board to determine responsibilities for accuracy & retention, identify customers for each type of data
- Develop common data dictionary accessible to everyone – including source, steward
- Form group to identify business case and approach to updating the feature inventory (including attributes) as work is done
- Task governance board with how to operationalize “data as an asset” – emphasize life cycle management
- Communicate data governance implementation plan agency-wide

As-Built Plan Data Management Assessment

The As-Built Data Management Team consisted of individuals representing the statewide (central office) construction function. One resident engineer was included in the group. “As-builts” data refer to data on any changes in a construction or improvement project that are made in the field after the construction plan has been approved, bid and awarded. The original policy for construction as-builts was written in 2002 and updated in 2013. Region design engineers are responsible for ensuring that as-built data are provided by contractors at the time of project close-out. The central construction group serves as a clearinghouse for the pass through of as-built changes from the regions. Data received are entered into ProjectWise. There are specifications for the submittal of data. Currently, as-built changes are received in a variety of formats from scanned plan sheets to CADD drawings.

Assessment Results

The team’s assessment results are summarized in Table B.21.

Table B.21. Utah DOT As-Built Plans Data Management Team Assessment Results

	1-Initial: Ad-Hoc	2-Aware: Recognize Need for Improvement	3-Defined: Developed & Documented Processes	4-Proficient: Processes Implemented as Defined	5-Optimizing: Strategies Evaluated and Improved
DATA GOVERNANCE					
Data Stewardship	X				
Understanding of Costs and Value	X don’t quantify costs – value of doing this not clear	X understanding of value			
Customer Relationships	X Respond as requests come in				
Data Preservation			X All plans are in Project Wise sorted by project		
Findability/ Documentation		X limited documentation of quality; limited ability to find by location			

	1-Initial: Ad-Hoc	2-Aware: Recognize Need for Improvement	3-Defined: Developed & Documented Processes	4-Proficient: Processes Implement- ed as Defined	5-Optimizing: Strategies Evaluated and Improved
DATA INTEGRATION & ARCHITECTURE					
Data Accessibility/ Usability		X Variable formats; contractors need to re-create plan in 3D; relation to other data sets not established		X Available in ProjectWise; some in UGate	
DATA QUALITY					
Data Quality Assurance	X How do you address data quality		X How do you ensure quality		
Accuracy	X Processes to improve accuracy	X Understand accuracy	X Standards for location accuracy		
Completeness	X Standards for completeness; No systematic process as written for filling in the gaps	X Some understanding but cannot quantify	X Processes for improving completeness Inspectors do check		
Currency				X Available within 90 days of project close-out Standards in place Processes are in place	
Consistency	X Aware there are inconsistencies but ROI of addressing them hasn't been established				

Gaps and Associated Consequences

Table B.22 summarizes the gaps and associated consequences identified by the team during the second assessment workshop.

Table B.22. Utah DOT As-Built Plan Data Management Team Gaps and Consequences

Gap Description	Consequences/Impacts
<p>No single person accountable for the as-built data program; lack of designated owner for the data. As-builts are produced as part of construction, but primary users are in other units - Maintenance, Design, ROW</p>	<ul style="list-style-type: none"> Without someone responsible, the as-builts provide limited benefits for future users. Accuracy suffers with nobody in charge. If they aren't reliable or consistent, people don't use them. It isn't clear how to find them.
<p>As-built policies and specifications exist, but there is uneven awareness and adherence to these policies and specs.</p>	<ul style="list-style-type: none"> Data quality suffers; inconsistencies across as-built plans
<p>There is limited awareness and understanding of current and potential customers for as-builts.</p>	<ul style="list-style-type: none"> Spotty compliance with standards and perceived value of as-builts – if customers were known, it would elevate priority of data quality
<p>Policies for retention and archiving may not reflect all customer needs</p>	<ul style="list-style-type: none"> May be storing too much data, or not enough. May not have info when a need arises
<p>It is not easy to find the most recent set of as-builts for a given location—or verify that a given set of as-builts is the most recent. Dates on records may not reflect the actual completion date.</p>	<ul style="list-style-type: none"> Potential for unneeded re-surveying when data already exists, because designers don't have confidence that the information exists or is current
<p>Majority of as-builts are submitted as scanned images in PDF.</p>	<ul style="list-style-type: none"> As-builts can't be searched or used for asset extraction.
<p>Additional work is needed to attribute as-builts based on route-milepost LRS</p>	<ul style="list-style-type: none"> Difficult to find the as-builts
<p>Difficult to locate features within as-builts on the LRS (convert from stationing)</p>	<ul style="list-style-type: none"> Inconvenient and time-consuming to use the information
<p>The relationship between Mandli asset/road feature data collection and as-built information has not been defined - e.g. approach to providing a complete picture of the current roadway (including subsurface and ROW info)</p>	<ul style="list-style-type: none"> Collecting information twice; detracts from trust in feature inventory
<p>It is difficult to ascertain whether utility information on as-builts is still current given lack of information provided to UDOT by utilities on their work.</p>	<ul style="list-style-type: none"> Lack of trust in the utility information – re-surveying of subsurface utility. Utility-related delays and repair costs

Gap Description	Consequences/Impacts
There is limited understanding of what data and level of accuracy is required by different as-built data users.	<ul style="list-style-type: none"> As-builts may not meet the needs of some customers – see above.
There may be a need for additional guidance and/or training in order to achieve greater data consistency (e.g. what level of accuracy is needed, what specific measurements are required – top of pipe vs. invert)	<ul style="list-style-type: none"> Data not reliable, increased variability, lack of trust, potential for need to re-survey
Accuracy and completeness (including blue sheets) of as-builts is inconsistent; as-built quality is not a current area of emphasis for contractors since it may require carrying designers/surveyors for the project duration. (Possible action: make as-builts a pay item – but this may impact bid costs)	<ul style="list-style-type: none"> Same as above

Priority Actions

The following priority actions were identified to address the identified gaps:

- Form committee to develop business case for improving as-built data management and strategy to present to technical committee
- Identify who needs the as-built data and what is important to them – via survey or interviews/ meetings
- Identify single point of accountability and define responsibilities for as-built program
- Analyze retention process and determine needs: Identify customers who handle documentation and send to retention processes, records retention reps, survey needs
- Develop specification with data attributes and accuracy requirements, make as-builts a pay item and define improved process to ensure compliance
- Evaluate changes in utility permitting process to require submittal of info
- Make as-builts available for easy retrieval (e.g. based on route, accessed via project map)

Agency-Wide Data Management Assessment

The agency-wide data assessment team consisted of representatives from Information Technology, GIS, project development, and finance.

Assessment Results

The team’s assessment results are summarized in Table B.23.

Table B.23. Utah DOT Agency-Wide Team Assessment Results

	1-Initial: Ad-Hoc	2-Aware: Recognize Need for Improvement	3-Defined: Developed & Documented Processes	4-Proficient: Processes Implemented as Defined	5-Optimizing: Strategies Evaluated and Improved
DATA GOVERNANCE					
Data Ownership Stewardship		X	X		
Data Governance	X No data governance goals and objectives	X Informal leadership body meets Limited documentation of changes There is an inventory of applications	X Governance standards for LRS and UGate		
Understanding of Costs and Benefits	X No processes for identifying overall priorities	X Both level 1 & 2 are true			
Customer Relationships		X Varies by data program	X UGate and UPlan has done more		
Findability/ Documentation		X Other	X UGate – data catalog exists		
Data Preservation			X Retention schedules in place but need customer validation		

	1-Initial: Ad-Hoc	2-Aware: Recognize Need for Improvement	3-Defined: Developed & Documented Processes	4-Proficient: Processes Implemented as Defined	5-Optimizing: Strategies Evaluated and Improved
Data Security		X Disaster recovery plan exists, not fully tested	X Have classification at system vs. data level; stewards manage access; encryption used for data "at rest"		
DATA ARCHITECTURE & INTEGRATION					
Geospatial Referencing			X Refining definitions		X Referencing used for all new systems
Data Architecture		X no data architect but data warehouse project manager provides some capability here Some master data have been identified			
Data Integration			X Some progress with UGate; data initiative to further this		
Change Impact Analysis		X In process of implementing Oracle ODI to provide some capabilities here			

	1-Initial: Ad-Hoc	2-Aware: Recognize Need for Improvement	3-Defined: Developed & Documented Processes	4-Proficient: Processes Implemented as Defined	5-Optimizing: Strategies Evaluated and Improved
Data Collection Coordination & Data Sharing				X LiDAR data collection was a coordinated effort Data are being shared among agencies – AGRC clearinghouse	
Data Usability		X Exploring additional ways of providing data		X Several tools in place for data access and analysis	
DATA QUALITY					
Data Quality		X no agency- wide approach beyond standard software development process			

Gaps and Associated Consequences

Table B.24 summarizes the gaps and associated consequences identified by the team during the second assessment workshop.

Table B.24. Utah DOT Agency-Wide Data Management Team Gaps and Consequences

Gap Description	Consequences/Impacts
Not all core data programs have identified business data owners and stewards with well-defined responsibilities.	<ul style="list-style-type: none"> Provide a focal point for data improvement, ensure that data are meeting needs, no clear accountability for quality, can create confusion about who to go to. No defined source of truth.
Need to reduce risk of specialized data knowledge loss in the event of key employee departure.	<ul style="list-style-type: none"> Inefficiency – move backwards on the learning curve, impedes progress on long-term initiatives.
Need to better understand extent to which existing data are being fully utilized and leveraged	<ul style="list-style-type: none"> Don't know who to talk to get requirements, may be collecting data that aren't needed. Hard to articulate value of data. May be missing opportunities to use data to gain efficiencies.

Gap Description	Consequences/Impacts
Data mapping has not been completed to identify dependent business processes and customers for each data set. (Starting to do mapping as part of DW project)	<ul style="list-style-type: none"> • Same as above
May be a need to better understand cost of data programs in order to evaluate outsourcing opportunities and weigh against value	<ul style="list-style-type: none"> • Difficult to evaluate options for gaining efficiencies
Need to better understand and articulate value of data programs for achieving established business goals.	<ul style="list-style-type: none"> • May not apply an appropriate level of resources to data (e.g. under-investment)
There are no established forums (e.g. communities of interest) for data users across business units to articulate needs and participate in defining data improvements. (Exception: GIS user group.)	<ul style="list-style-type: none"> • Missed opportunity for discussion and consensus about improvement needs
Data retention and archiving policies exist but have not necessarily been validated based on an understanding of business needs.	<ul style="list-style-type: none"> • Risk of needing to re-generate data (e.g. for legal purposes); If data is inadvertently destroyed, may not be able to recover if not discovered immediately. Needs for snapshots in time aren't met. (Discussion with users resulted in no identified needs for archiving). Lack of understanding of the potential risk of loss.
There is limited ability to conduct an analysis of how a data format or definition change in one system will impact other systems.	<ul style="list-style-type: none"> • Staff time impacts to fix; risk of reporting inaccurate information due to inconsistent data across systems. (due to undefined change management processes)
Responsibilities and processes for data change management have not been defined and documented.	<ul style="list-style-type: none"> • Same as above
A single authoritative source system of record has not been identified for each data element of agency-wide interest.	<ul style="list-style-type: none"> • Lack of ability to integrate data and provide "single version of the truth." Limited commitment and accountability to data quality considering downstream uses (broader constituency). Can't understand lineage of data copied from other systems.
Mapping has not yet been completed to identify which systems include information for key corporate data entities (e.g. projects, road segments, vendors).	<ul style="list-style-type: none"> • Potential for redundancies and inconsistencies.
Lack of end-user awareness of data set characteristics – how collected, updated, etc.	<ul style="list-style-type: none"> • Lack of trust of data
Disaster recovery plans are in place but have not been tested.	<ul style="list-style-type: none"> • Potential for data loss.

Gap Description	Consequences/Impacts
A data architect role has been established, but sufficient resources and support processes are not in place to ensure data consistency and integration across systems.	<ul style="list-style-type: none"> Lack of commitment to data management, data not being used to its full value, lack of ability to make progress.
Master data has not yet been standardized and synchronized across business units and applications	<ul style="list-style-type: none"> Same as above
Lack of specific data marts addressing business needs (to be addressed via data warehouse initiative); need to address broader data usability needs (beyond geospatial). Need to define audiences – needs and capabilities.	<ul style="list-style-type: none"> Not getting full value from data, not able to provide answers to business questions
Lack of agency-wide tools and processes for identifying what (non- spatial) data exists and what its quality is (opportunity to address with new Oracle tools)	<ul style="list-style-type: none"> Can't determine suitability for purpose

Priority Actions

The following priority actions were identified to address the identified gaps:

- Formalize (document) data stewardship responsibilities – including understanding of data customers
- Identify enterprise-wide conformed dimensions and authoritative sources (e.g. region boundaries, maintenance stations)
- Define single source of truth for data elements
- Identify agency-wide data program steering team
- Define the different data programs
- Identify how data management activities are to be resourced – develop a value statement – end product
- Make data management part of the culture – way of doing business. Demonstrate business value through concrete successes
- Document roles for data stewardship and communicate these
- Understand users for each data set – map data needs to business processes - start with mapping business processes
- Complete one of the data warehouse/BI life cycle loops and demonstrate value.

Synthesis of Gaps and Improvements – Data Value

The following summarizes the kinds of gaps that were identified by the four Data Value Assessment teams:

Availability Gaps

- Data on system characteristics for non-state maintained roads, incidents, real-time traffic, speed, vehicle classification and freight movement
- Roadway inventory data for features such as shoulders, sidewalks, utilities, culverts and right of way
- Timely data on roadway changes
- Bicycle/pedestrian data including counts and ability to identify related improvements in roadway projects
- Weekend and holiday travel data
- Information from construction plans on changes to road/asset inventory
- Information to support performance management – e.g. costs to achieve performance targets.

Quality Gaps

- Validation of traffic data accounting for sensor malfunctions
- Consistency and standardization of data element coding
- Currency of project status information
- Specificity of inspection information to enable understanding of required maintenance or rehabilitation actions
- Consistency of data – need for improved inspector training

Usability Gaps

- Finding and integrating information from multiple sources for needs analysis, scoping, prioritization and programming
- Lack of ability to roll up data from different geographic locations and/or data systems
- Insufficient capabilities to track and diagnose issues and root causes of issues
- Lack of standardized reports to meet common information needs; lack of ad-hoc reporting tools
- Not all data needed for business decisions is in electronic formats or easily available to users
- Data duplication – not clear which is the authoritative source
- Difficulty making use of archived data due to inconsistencies in the way historical data were collected and managed
- Inflexible information systems that data maintenance, access, sharing, reporting and use difficult
- Education - not all staff know how to access and use available data and tools (e.g. GIS)

Improvements

- Expand traffic, bike and pedestrian data collection activities to meet FHWA reporting requirements, performance monitoring, modal planning and other business needs.
- Improve roadway inventory data collection when there are pressing business needs for the data – and look for tools (e.g. real-time field collection) that improve collection efficiencies.
- Offer training to staff on how access, integrate and use available data.
- Replace antiquated data systems with systems that are easier to use and maintain when there are clear business needs and benefits exceed costs.
- Move core data into enterprise systems that provide easier access and use.
- Provide additional documentation or reference manuals on how to use existing data systems.
- Improve and streamline reporting functionalities to address user needs.
- Provide training in how to use data viewer and analysis tools and provide more user-friendly ways of querying and browsing data.

Synthesis of Gaps and Improvements – Data Management

The following summarizes the kinds of gaps that were identified by the Data Management Assessment teams:

Data Governance Gaps

- Ownership and Stewardship
 - Accountability for data quality and management has not been defined for all core data programs.
 - Single, authoritative sources have not been identified for some data elements.
 - Metrics for communicating the value and success of data governance are not widely understood or used.
 - There may be a lack of executive-level understanding about the importance and value of some data. There is a need to improve capabilities to assess, document and communicate the value proposition for core data.
 - Data governance policies and standards may not exist.
 - There is limited training and succession planning to sustain staffing resources for key data programs to reduce the risk of specialized data knowledge loss in the event of key employee departures.
- Data Costs/Benefits
 - The full value or costs/benefits of data programs are not consistently tracked.
- Customer Relationships
 - There are few established forums for data users across business units to articulate needs and participate in defining data improvement.

- There is no systematic outreach to understand current and future data user needs, and understand the extent to which existing data are being fully utilized and leveraged.
- In some cases staff are entering data without a good understanding of how they are being used.
- Data Documentation and Findability
 - Data documentation is done as time allows – and is often lacking due to competing priorities.
 - Users are not able to easily identify what data exists and what its quality is.
 - There are opportunities to refine data documentation in terms of data definitions, dictionaries, catalogs, metadata and contact information.
 - Department policies for archiving data exist but focus on paper records and are not always well understood.
 - Retention schedules focus on legal requirements and may not be based on an understanding of user needs for archived data.
- Security
 - Disaster recovery plans are not always in place for core data. When they have been established they may not have been fully tested.

Data Architecture and Integration Gaps

- Linkages and standards
 - Data exist in separate systems and functional areas without an overall strategy or architecture for data integration.
 - There are not always sufficient resources to ensure data consistency and integration across systems.
 - “Master data” have not consistently been standardized and shared/synchronized across business units and applications.
 - Not all data has standardized geographic referencing to enable mapping and integration with other agency data.
 - Duplicative data entry may exist due to lack of automation or system integration.
 - It is not always easy to visualize trend information on asset conditions, capital and maintenance expenditures, traffic, crash rates and other important performance indicators.
- Change Management Processes
 - Responsibilities and processes for data change management have not been defined, documented or consistently implemented.
 - There is limited ability to conduct an analysis of how a data format or definition change in one system will impact another.
 - Changes made in one system do not always ripple through other systems that use or consume the data.
 - Change notification processes are not consistently in place to alert users and consumers of data changes.
 - While data architect roles have been established,

- Data Collection Coordination and Sharing
 - Opportunities for coordinating data collection efforts across business units are pursued on an ad-hoc basis, with limited enterprise-wide proactive data collection coordination.
 - New data collection efforts may not be consistently reviewed to minimize duplication with existing data collection activities.

Data Quality Gaps

- Processes for understanding and documenting user needs for accuracy, timeliness and completeness may not exist.
- Guidelines for determining spatial accuracy requirements and appropriate collection methods have not been formally established for all data programs.
- Staff do not have access to consistent and complete information about the quality of available data.
- Systematic data quality assurance and cleansing processes may not exist.
- Data review and acceptance processes may not exist.
- Limited data validation rules are built in to source data systems.
- Data consistency issues exist due to changes in coding conventions over time.

Improvements

Data Governance

- Ownership and Stewardship
 - Create a data management steering committee and/or more fully clarify data the roles and responsibilities of existing governance boards.
 - Identify data area stewards and establish a data stewardship group.
 - Work towards making data management part of the agency culture. For example, develop a training strategy for data management to build common terminology and understanding of basic concepts. Develop a communication strategy and presentations that can be shared with managers, business areas and data program managers on data management goals, strategies and definitions.
 - Identify how data management activities are to be resourced.
 - Identify and assess the risks of core data sets not currently housed in enterprise systems.
 - Implement data roll-out processes as new data come on line to help customers understand how the data can be used to improve decision making and gain efficiencies in work processes.
 - Identify enterprise-wide conformed dimensions and authoritative sources for boundaries, building locations and other enterprise data elements.
 - Define a single source of truth for data elements.
- Customer Relationships
 - Develop methods to bring data users together with IT and data providers to discuss current and future opportunities, issues and needs.

- Map data needs for core business processes.
- Documentation and Findability
 - Create a data catalog to outline what data are available, where they are located, how to access the data and who to contact for additional information.
 - Develop a common data dictionary that is accessible to everyone in the agency.
 - Ensure that metadata and other data documentation are routinely reviewed and updated.
- Data Security and Retention
 - Develop formal policies and processes for data archiving and preservation.

Architecture and Integration

- Linkages and Standards
 - Develop a data integration strategy.
 - Produce a high level business architecture that maps data to business processes.
 - Spatially enable core business data.
 - Develop or purchase new data management systems when business needs, benefits and risks warrant.
- Change Management
 - Define processes for improved communications and change management.
 - Build interfaces between data systems so that changes that occur in one system can ripple onto others that consume the data.

Data Quality

- Document accuracy and Q/A processes for different data elements.
- Standardize systems to eliminate redundancies.
- Train personnel in how to enter data and use the systems.

APPENDIX C – SUMMARY PRESENTATION

NCHRP 8-92: Implementing A Transportation Agency Data Program Self-Assessment

What is a Transportation Data Self-Assessment?

A methodology, guide and tools to help agencies answer four key 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, etc. is needed?
- Are we making best use of our data collection and management resources?
 - Are we being efficient about how we collect and manage the data?
- Are we getting full value from the data that we have?
 - Are users able to access, integrate and analyze it?

2

Why Conduct the Self-Assessment?

- **Data is an essential transportation agency asset**
 - Transportation agencies are increasingly data driven
 - Data needs are growing in number and complexity
 - Performance management
 - Asset management
 - System operations and traveler information
- **Data is expensive to collect and maintain**
 - Important to derive full value from data investments
 - Systematically identify opportunities to improve efficiencies and adjust the data portfolio to better meet agency needs

3

Operationalizing AASHTO's Data Principles

Data Principle

1. VALUABLE: Data is an asset
2. AVAILABLE: Data is open, accessible, transparent and shared
3. RELIABLE: Data quality and extent is fit for a variety of applications
4. AUTHORIZED: Data is secure and compliant with regulations
5. CLEAR: There is a common vocabulary and data definition
6. EFFICIENT: Data is not duplicated
7. ACCOUNTABLE: Decisions maximize the benefit of the data

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.

4

Benefits of a Data Self-Assessment

MOVE FROM:

- Data as a burden
- Seat of the pants decisions
- We don't trust the data
- We don't have that information
- Multiple inconsistent answers to your question
- Multiple uncoordinated data silos

TO:

- Data as an asset
- Data-driven decisions
- We rely on the data
- Look that up on your dashboard
- One single authoritative answer to your question
- Integrated data systems

5

Without Strong Data Management...

- Data may be collected but not well utilized because of insufficient quality, access or documentation
- Data may not be easily integrated to provide for meaningful analysis
- Data may be duplicated resulting in inefficiencies, inconsistencies and conflicting information
- Data may be collected that no longer adds value while other more pressing data needs go unmet
- Staff resources may lack tools and systems to effectively and efficiently respond to critical information requests

6

Background on NCHRP 8-92

The NCHRP 8-92 project included:

- A literature review of strong data assessment methods and frameworks
- Interviews with state and regional transportation executives and AASHTO and US DOT representatives
- Focus groups in five state transportation agencies – Colorado, Kentucky, Maryland, Minnesota and Oregon
- Case studies in two state transportation agencies – Utah and Michigan piloted the data value and data management assessment tools at several business area levels and at an enterprise level

7

NCHRP 8-92 Research Results

The NCHRP 8-92 project resulted in:

- A refined model and methodology for conducting agency data self-assessments
- A Guide that provides step-by-step guidance for agencies wishing to do transportation data-self assessments
- A set of assessment tools for gauging the maturity of data management practices and the quality of data for meeting business functions
- Examples of what agencies can do to “step up” and advance data management maturity levels and the value that can be derived from such actions

8

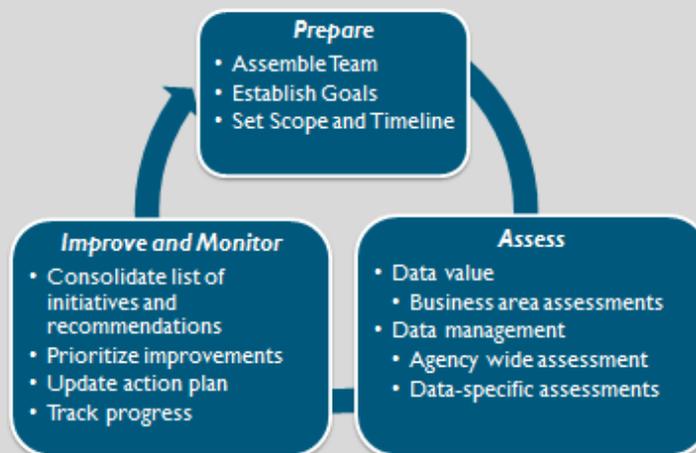
Assessment Overview

The data self-assessment framework features two assessment tools to examine current needs and practices:

- **Data Value Assessment** – assesses the degree to which *data users feel that data are providing value and meeting business needs*
- **Data Asset Management Maturity Assessment** – assesses the current level of *agency capabilities* for managing data assets to maximize their value

9

Assessment Process



10

Implementing a Transportation Agency Data Self-Assessment

NCHRP 08-92

Prepare Assess Improve

QUESTIONS

Do we have the right data to make good decisions and meet reporting requirements?

Agency Leaders

How can our agency make it quicker and easier to access and analyze data so that we can do our jobs more efficiently and effectively?

Data Users

Is our data good enough? Do we need to improve its level of accuracy, precision or timeliness?

Data Stewards

Are we managing our data to maximize its value and ensure its integrity?

Data Managers

Prepare

- ASSEMBLE TEAM**
Assemble a broad-based team to guide the effort.
- ESTABLISH ASSESSMENT GOALS**
Set a clear direction for what is to be accomplished.
- SET SCOPE AND TIMELINE**
Select data programs and assessment elements to include and establish a scope and a schedule for the effort.

Assess

- ASSESS DATA VALUE**
Assess current data availability, quality, and usability.
- ASSESS DATA MANAGEMENT**
Assess maturity level for current data management processes.
- DETERMINE GAPS**
Identify gaps between current state and desired state and identify candidate actions to close gaps.

Improve

- PRIORITIZE IMPROVEMENTS**
Analyze the results and prioritize actions for improvement.
- DEVELOP ACTION PLAN**
Develop a plan of specific actions to address the priority gaps.
- IMPLEMENT PLAN**
Assign responsibilities, allocate resources and track implementation.

ASSESSMENT



ACTIONS

- Data Consolidation and Standardization
- Data Collection, Processing, and Quality Improvements
- Data Management Staffing and Responsibilities
- Data Policies, Procedures, and Standards
- Data Mapping and Documentation
- Data Presentation and Analysis Improvements
- Information System Improvements

Data Value Elements

Element	Rating
Data availability – are 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 are available - some gaps remain
	● Excellent. Sufficient data are available to meet needs
Data quality – are data sufficiently accurate, credible and current to support decision making?	● Poor. Quality not sufficient – data not useful
	● Fair. Lack of currency, reliability or completeness limits value
	● Good. Acceptable but needs improvement
	● Excellent. Sufficient to meet needs
Usability – can data be easily integrated, analyzed and presented as needed to support decision making?	● Poor. Requires <i>substantial</i> effort to get data into usable form
	● Fair. Requires <i>moderate</i> effort to get data into usable form
	● Good. In usable form but reporting improvements helpful
	● Excellent. In usable form, no improvement needed

Data Value Sample Results

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

13

Data Value Results – Sample Gaps

Type of Data	Gaps	Business Impacts
Maintenance Work History	<i>Availability:</i> We need historical information for budgeting, but we only have aggregate expenditures, not costs by activity	Better data would improve ability to link budget estimates with expected outputs
Sign Inventory	<i>Quality:</i> Sign inventory is 3 years old and doesn't reflect recent work done	Districts won't use the inventory because they don't trust the data. They will spend time re-collecting data.
Traffic	<i>Usability:</i> We must submit a request to IT in order to get the traffic data reports we need	Strains IT resources and limits business value of the data

14

Data Management Assessment Elements

- **Data Strategy and Governance:** how decisions are made about what data to collect and how to manage and deliver it -- including roles, accountability, policies and processes.
- **Life-Cycle Data Management:** how data are maintained, preserved, protected, documented and delivered.
- **Data Architecture and Integration:** practices to standardize and integrate data to minimize duplication and inconsistencies, including spatial referencing
- **Data Collaboration:** processes to coordinate data collection and management with internal and external users
- **Data Quality Management:** practices to define, validate, measure and report data quality

15

Data Management Maturity Levels

Maturity Level Name	Definition
1 – Initial	Processes, strategies and tools are generally ad-hoc rather than proactive or enterprise-wide; successes are due to individual efforts
2 - Developing	Widespread awareness of more mature data management practices; recognition of the need to improve processes, strategies and tools
3 - Defined	Processes, strategies and tools have been developed, agreed-upon and documented
4 – Functioning	Processes, strategies and tools are generally being implemented as defined
5 – Optimizing	Strategies, processes and tools are routinely evaluated and improved

16

Example: Data Quality

Level	Definition
1 – Initial	Data quality is addressed on an ad-hoc basis in response to reported issues.
2 - Developing	There have been some efforts to work with data users to proactively discuss and define data quality requirements. Standard practices are being defined.
3 - Defined	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.
4 – Functioning	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 in to applications.
5 – Sustained	Data quality assurance processes are regularly assessed and improved. Automated error reporting tools are available for data users. Data validation and cleansing tools are used to identify and address missing or invalid values.

17

Benefits of Moving Up the Maturity Scale

- Agency data can be used as intended and can be used to produce reliable information that is valuable for decision making – because:
 - Data quality is addressed proactively, using standard quality control and quality assurance processes
 - Data are validated based on established business rules
 - Data cleansing processes are automated
 - Efficient error reporting and correction processes are in place

18

Data Management – Sample Results

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

19

Data Management – Sample Gaps

Assessment Element	Gaps	Business Impacts
Data Strategy and Governance	Accountability for data hasn't been established.	Data aren't meeting user needs.
Data Architecture and Integration	Coding for districts and jurisdictions hasn't been standardized across data sets.	Takes a lot of manual effort to integrate different data sets to provide value for management decisions
Data Collaboration	Several different districts are independently collecting the same type of data.	Missed opportunity for a more efficient statewide approach
Data Quality	Pavement data are being collected without an established QA process	Districts don't trust the data and are reluctant to use it

20

Data Improvements: Strategy and Governance

- Data Governance Bodies
- Data Governance and Stewardship Policies
- Data Business Plans
- Data Management Roles and Responsibilities
- Data Value Mapping
- Data Communities of Interest

21

Data Improvements: Life Cycle Management

- Standard Operating Procedures
- Data Change Management
- Data Catalogs and Dictionaries
- Data Curation Profiles
- Data Management Plans
- Data Retention Schedules and Archiving
- Data Access Policies
- Data Delivery Platforms

22

Data Improvements: Architecture and Integration

- Common Geospatial Referencing
- Standardized Approach to Temporal Data
- Reference Data Management
- Master Data Management
- Data Architecture Practices and Roles
- Business Glossaries
- Data Integration Tools

23

Data Improvements: Data Collaboration

- Multi-Purpose Data Collection
- Data Clearinghouses/Open Data Platforms
- Data Partnerships
- Data Sharing Agreements
- Data Outsourcing

24

Data Improvements: Data Quality

- Data Quality Metrics
- Data Validation Rules
- Data Cleansing
- Data Collection Quality Management Processes

25

The Self-Assessment Process is Flexible

- Conduct the **data management** assessment at an enterprise level
- Conduct the **data management** assessment for one or more data management areas (e.g. traffic or maintenance)
- Conduct the **data value** assessment 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
- Pursue a comprehensive agency-wide approach using all of the above for priority business areas or data categories

26

The Self-Assessment Process Recognizes Resource Limitations

- Resources for data improvements are limited – staff, expertise, money and time
- All data “wants” and “needs” cannot be met
- It is not necessarily cost-effective to be at the highest maturity level for any given data management element
- *Self-Assessment Process encourages agencies to be selective and prioritize actions based on support for agency priorities and risks of not taking action*

27

Conclusion

The Transportation Data Self-Assessment can help agencies to:

- Understand how well their data is working for them
- Understand what investments in data are not paying off – and why
- Make strategic investments to get data programs in alignment with current and future agency priorities
- Focus and strengthen data management roles, structures, policies, practices and processes to minimize risks and improve efficiencies
- Periodically check on the progress of improvements – and readjust as needed

28

Side 2:

AASHTO Data Sub-Committee Core Transportation Data Principles

The AASHTO Standing Committee on Planning (SCOP) Data Sub-Committee leadership has defined a set of core principles for transportation data.

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, data must be open and usable for diverse applications and open to all.

Principle 3: Reliable

Data quality and extent is fit for a variety of 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 is 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 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 are essential to maximize the utility of data for decision making.

For more information, visit: <http://planning.transportation.org/Pages/Data.aspx>