NCHRP 20-24(37)F

Establishment of Comparative Performance Measures Program Infrastructure to Support System Performance Data Collection and Analysis

Final Report

Prepared for

National Cooperative Highway Research Program

Prepared by

Spy Pond Partners, LLC

July 15, 2011

The information contained in this report was prepared as part of NCHRP Project 20-24(37)F, National Cooperative Highway Research Program.

Acknowledgements

This study was conducted for the American Association of State Highway and Transportation Officials (AASHTO), with funding provided through the National Cooperative Highway Research Program (NCHRP) Project 20-24. NCHRP is supported by annual voluntary contributions from the state departments of transportation (DOTs). NCHRP Project 20-24 is intended to fund quick response studies to address specific needs of chief executive officers (CEOs) and other top managers of DOTs. The report was prepared by Spy Pond Partners, LLC. The work was guided by a technical working group. The project was managed by Dr. Andrew Lemer, NCHRP Senior Program Officer.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

AADT	Average Annual Daily Traffic
AASHTO	American Association of State Highway and Transportation Officials
ARRA	American Recovery and Reinvestment Act
BTS	Bureau of Transportation Statistics
CPM	Comparative Performance Measures
CSGDM	Content Standard for Digital Geospatial Metadata
DBA	Database Administrator
DOT	Department of Transportation
EIS	Environmental Impact Statement
ETL	Extract, Transform and Load
FARS	Fatality Analysis Reporting System
FGDC	Federal Geographic Data Committee
FHWA	Federal Highway Administration
HPMS	Highway Performance Monitoring System
IRI	International Roughness Index
LTPP	Long Term Pavement Performance
MRI	Mean Roughness Index
NBI	National Bridge Inventory
NCHRP	National Cooperative Highway Research Program
NHS	National Highway System
SaaS	Software as a Service
SCOH	Standing Committee on Highways
SCOQ	Standing Committee on Quality
SD	Structurally Deficient
TCEED	Technical Committee on Electronic Engineering Data
VMT	Vehicle Miles of Travel
XML	eXtensible Markup Language

Executive Summary

Study Background

Six years ago, the AASHTO Standing Committee on Quality (SCOQ) Performance Measures and Benchmarking Subcommittee (QSPMB) initiated a Comparative Performance Measurement (CPM) initiative. This initiative – currently being continued under the auspices of the AASHTO Standing Committee on Performance Management (SCOPM) - has sought to provide information that states can use as they explore ways to improve their performance:

- Where does my state stand relative to the national average and relative to other states with similar characteristics?
- How did the higher performing states achieve their results? What factors contributed to their success?

A series of projects were undertaken under the NCHRP 20-24(37) series to collect data from multiple states, formulate and calculate comparative performance measures, classify states into peer groups for purposes of comparison and collect information on practices of top performing states. Five different performance areas have been covered to date – project delivery (on-time, on budget performance), pavement smoothness, bridge condition, safety (fatalities), and incident response. Results for each project are documented in published reports, and individual data sets have been provided to NCHRP in varying formats.

In 2010, NCHRP Project 20-24(37)F was initiated to consider establishment of a consistent, central infrastructure to store, manage and provide access to the data collected under the CPM series of projects. The scope of this project was to develop a conceptual framework and functional description of such an infrastructure, and develop a strategy for implementing and ongoing stewardship and management. This report presents the results of this project.

Framework for a CPM Infrastructure

A conceptual framework for the CPM infrastructure was developed based on a statement of purpose and need, identification of users and usage scenarios, analysis of the existing CPM data structures and a review of performance data presentation views used in existing web sites. The infrastructure concept was based on the following core assumptions:

- The infrastructure is intended to support the CPM series of projects. It is not intended to be the central repository for national transportation performance measures – though its development and structure can certainly inform the process of establishing a national performance measures web site.
- The infrastructure is intended primarily to serve the needs of state DOTs and other transportation agencies that provide their performance data. While the infrastructure can be made available to others (e.g. transportation researchers), it is not intended to provide a public-facing web site.

- There is value to providing a consistent and centralized way to view multiple comparative performance measures in a single location. That is, it is worth building a single web site where states can view and compare across multiple performance measures. This approach provides substantial advantages over an approach in which each CPM project develops and provides independent views and data sources for particular measures. However, portions of the infrastructure – particularly those dealing with collection and validation of data during the study process – can and should be built independently as part of individual CPM projects.
- The infrastructure should not attempt to duplicate existing performance data collection or measurement initiatives at FHWA or other organizations. However, it should provide a convenient centralized location where data on multiple performance measures can be viewed and analyzed. In cases where these measures are derived from national sources and supported by ongoing data collection programs, the infrastructure should rely on data feeds from those sources and avoid duplication of processes for data compilation and quality assurance.

The infrastructure as envisioned is shown in Figure 2 and would consist of:

- A CPM Database including:
  - Definitions and metadata;
  - Raw performance data and calculated performance results;
  - Data transformation procedures; and
  - National data with information that can be used to construct groups of peer states, with standard processes and protocols for refreshing these tables from external sources (e.g. U.S. DOT, Census Bureau).
- A CPM Web Site including:
  - Reporting and query tools that support data analysis and exploration for executives and practitioners;
  - Access to information gathered on practices identified in states that are in the top tier of performance among their peers;
  - Administrative tools for maintaining and extending the CPM web site and data; and
  - Documents and links of interest.

Specific use scenarios, functions, screens and supporting data structures are detailed in the body of the report.

Implementation and Stewardship

Activities required for initial development and ongoing maintenance and management of the infrastructure were identified, and rough estimates of the level of effort and associated costs for these activities were developed. Development costs were estimated at $200,000-300,000. Continuing costs were estimated at $30,000-$50,000 per year. However, incremental development options and lower cost approaches were identified to begin with core functions

only. In addition, ongoing costs may be reduced through synergies with other efforts or piggybacking on existing technical and administrative infrastructures that are in place at AASHTO and/or FHWA.

Three stewardship options were identified:

- AASHTO lead with FHWA participation - AASHTO would develop, host and maintain the infrastructure, which would include a joint AASHTO/FHWA Comparative Performance web site. FHWA would provide data feeds for those data sets that it collects and maintains.
- FHWA lead with AASHTO participation - FHWA would develop, host and maintain the infrastructure, which would include a joint AASHTO/FHWA Comparative Performance web site. FHWA would implement the CPM database by building upon its current initiative to link data from NBI, HPMS and other sources and develop a data warehouse.
- Pooled fund/consortium approach - AASHTO and FHWA would work with interested states to initiate a pooled fund study for building and sustaining the CPM infrastructure. The work would be accomplished by a contractor, and the web site and database would be hosted either by the contractor or by the lead pooled fund study organization. FHWA and AASHTO would supply data.

Each of these options were assessed based on lead time to implement, flexibility to adapt to changing requirements, likelihood for additional costs to states, and potential opportunities for cost savings or functionality enhancements due to synergies with other initiatives. Note that a detailed analysis of how initial and ongoing costs would vary across these three options was not conducted as part of this research. The activities and level of effort to be performed to build the infrastructure and manage the infrastructure will not vary substantially across the options. What may vary are (1) the extent to which contract resources are used versus in-house staff, and (2) opportunities to piggyback on existing hardware, software and staff resources.

The second stewardship option (FHWA lead with AASHTO participation) was identified as the one with the greatest opportunity for building on existing work and therefore minimizing additional costs to states. This option also offers the greatest opportunity for maximizing functionality and value of the CPM infrastructure through integration with a more extensive, geospatially-enabled data warehouse. However, this option would likely involve a longer implementation timeframe and less flexibility to make modifications in response to changing needs and priorities.

If resource constraints prevent immediate implementation of these components, or if the selected stewardship option involves substantial lead time for implementation, a lower cost ($50,000-$100,000), incremental approach could be considered. This approach would involve compiling the existing data into a database and developing standards for future CPM projects to ensure that any data produced could easily be absorbed into the master database for display on standardized web site views.

One additional issue to be addressed as part of the infrastructure implementation process is that some of the existing CPM performance data were obtained from states on the condition that they remain anonymous. Anonymity was offered in order to maximize state participation in the CPM projects and place the emphasis on identification of effective practices as opposed to distinguishing "best and worst" performers. However, identification of states by name rather than by peer group alone would make the CPM infrastructure a more useful tool. As part of the implementation process, states that provided data can be asked if they would agree to relax the anonymity condition. If a critical mass of states agree, then data can be provided with state names. Remaining states could then choose to have their data only included in aggregate (multi-state) summaries, or included in the state-specific displays but with coded identifiers rather than state names. Going forward, states can be encouraged in future CPM projects to participate in a non-anonymous fashion, but provided the options above (identification by a code or inclusion only in the aggregate results.)

1. Introduction

Study Overview

This report presents work conducted under NCHRP Project 20-24(37)F – "Establishment of Comparative Performance Measures Program Infrastructure to Support System Data Collection and Analysis." The objective of this project was to develop a conceptual framework and functional description of a comparative performance database and analysis infrastructure to support benchmarking and management strategy formulation by DOTs and others. The study also identified an implementation and stewardship approach for the infrastructure to provide ongoing management and maintenance by AASHTO or another suitable organization.

This study consisted of the following tasks:

Task 1: Kickoff Meeting - conduct an initial web/teleconference with the NCHRP project panel to achieve a common understanding of the scope, priorities, issues and options for consideration.

Task 2: Framework - formulate a conceptual framework and functional description of a comparative performance database and analysis infrastructure to support benchmarking and management strategy formulation by DOTs and others.

Task 3: **Workshop** - conduct a workshop via web conference to obtain feedback on the conceptual framework, and document the recommended concept for the database and infrastructure in an Interim Report.

Task 4: **Implementation Strategy** - develop a strategy for implementing and maintaining the database and infrastructure for comparative performance.

Report Organization

Section 2 provides background on the CPM Initiative, and articulates the purpose and need for a CPM infrastructure.

Section 3 presents a conceptual framework for a CPM infrastructure, including a description of functions envisioned, anticipated users and scenarios for how it would be used.

Section 4 identifies potential CPM components and functions and presents a high-level architecture of the CPM infrastructure.

Section 5 presents a high level, logical organization of CPM database.

Section 6 discusses implementation activities, associated level of effort, and stewardship options. It presents a recommended approach for moving forward with establishment of the CPM infrastructure.

2. Background

AASHTO Comparative Performance Measurement Initiative

In 2004, the American Association of State Highway and Transportation Officials (AASHTO) Standing Committee on Quality (SCOQ) Performance Measures and Benchmarking Subcommittee (QSPMB) began to explore establishment of a collaborative initiative to develop "apples to apples" comparisons of performance measures in key strategic focus areas for state departments of transportation (DOT) – for example, project delivery, transportation system safety, preservation, congestion, system operations management, freight and economic development, and environment. An initial project under the NCHRP 20-24(37) series was undertaken to pilot the comparative performance concept. This resulted in publication of a 2006 report that presented the pilot project results and recommended some basic principles and a roadmap for taking the next steps in establishing a comparative performance measurement program.

Project 20-24(37)A, the first full-scale comparative performance measurement (CPM) project in the 20-24(37) series, was completed in April 2007. This project examined on-time, on-budget project delivery performance for 20 states (building on the initial pilot) and highlighted practices from nine of these states. This project was followed by Project 20-24(37)B focusing on pavement smoothness (completed May 2008), Project 20-24(37)C focusing on safety (completed May 2009), and Project 20-24(37)E on bridge condition (completed August 2010). Two additional efforts are currently nearing completion – Project 20-24(37)D on incident response, and an update to Project 20-24(37)A. The intent is to periodically revisit each performance area in order to improve upon prior measurement techniques, build additional trend information and continue to learn about effective practices as they evolve.

Each comparative performance study has involved selection of specific performance measures based on input from participating states; identification of data sources; collection and compilation of raw performance data; establishment of peer groups for comparative analysis; calculation of measures; identification of the top performing agencies with respect to the selected measures; and interviews to identify practices that may be related to performance results. The completed projects have been undertaken based on the principles and process established for comparative performance measurement, but specific approaches to data definition, peer grouping, data collection, data storage, and presentation have been tailored to the specific measures. This flexibility has been appropriate in the startup phase of the program. However, given the experience gained to date with different performance measure areas, and given current plans for the program, there is a need to initiate development of a consistent structure for how data are defined, described, collected, stored and how practice results are compiled and presented.

Purpose and Need for Comparative Performance Support Infrastructure

A well-defined process for comparative performance measurement and a support infrastructure for data description, collection, and presentation of results would be beneficial for several reasons. First, it would provide a reliable platform for data storage and retrieval to facilitate future access for analysis and research. Second, it would support future consistency across comparative performance projects, allowing results of multiple projects to be viewed within a common framework. Third, it would allow for future projects to be accomplished more efficiently by making available tools to facilitate data gathering, compilation, analysis and display. Efficiency could also be enhanced through establishment of standardized data feeds for peer grouping that could be used for multiple performance measures. Finally, it would provide additional visibility for the program which could enhance dissemination of findings and lead to broader understanding of the comparative performance program's value.

3. CPM Infrastructure Conceptual Framework

Priority Elements to Support the CPM Study Process

Development of the CPM Conceptual Framework began with taking a very broad view of how the infrastructure might support the entire CPM study process. While each CPM study to date has varied in its approach and methods, the basic study process involves the following six steps:

- **Startup** – obtaining commitments for participation, working with participating states to gain agreement on specific measures and data elements to be collected. This step typically involves developing an initial questionnaire to ascertain data availability and participant opinions about suitable measures.
- **Data Gathering** – development of a standard template for data submittals, requesting data from participants, receiving and reviewing data submittals, obtaining corrected data as needed.
- **Data Analysis and Ranking** – processing the data to calculate values of the agreed-upon performance measures; providing the calculated measures back to participants for review, identifying any data or calculation issues and making adjustments as needed, calculating final measures and rankings, and selecting states with higher performance for interviews.
- **Best Practice Identification** – developing an interview guide to identify specific practices that may be associated with the identified performance results, conducting interviews, compiling interview results, and finalizing the interview findings based on state review.
- **Documentation of Results** – preparation of draft and final reports, and transmittal of data files.
- **Dissemination of Study Findings** – communication of results at various forums.

Potential elements of a CPM infrastructure to support each of these steps is shown in Figure 1. In the figure, the infrastructure element is shown directly below the process step that it is intended to support.

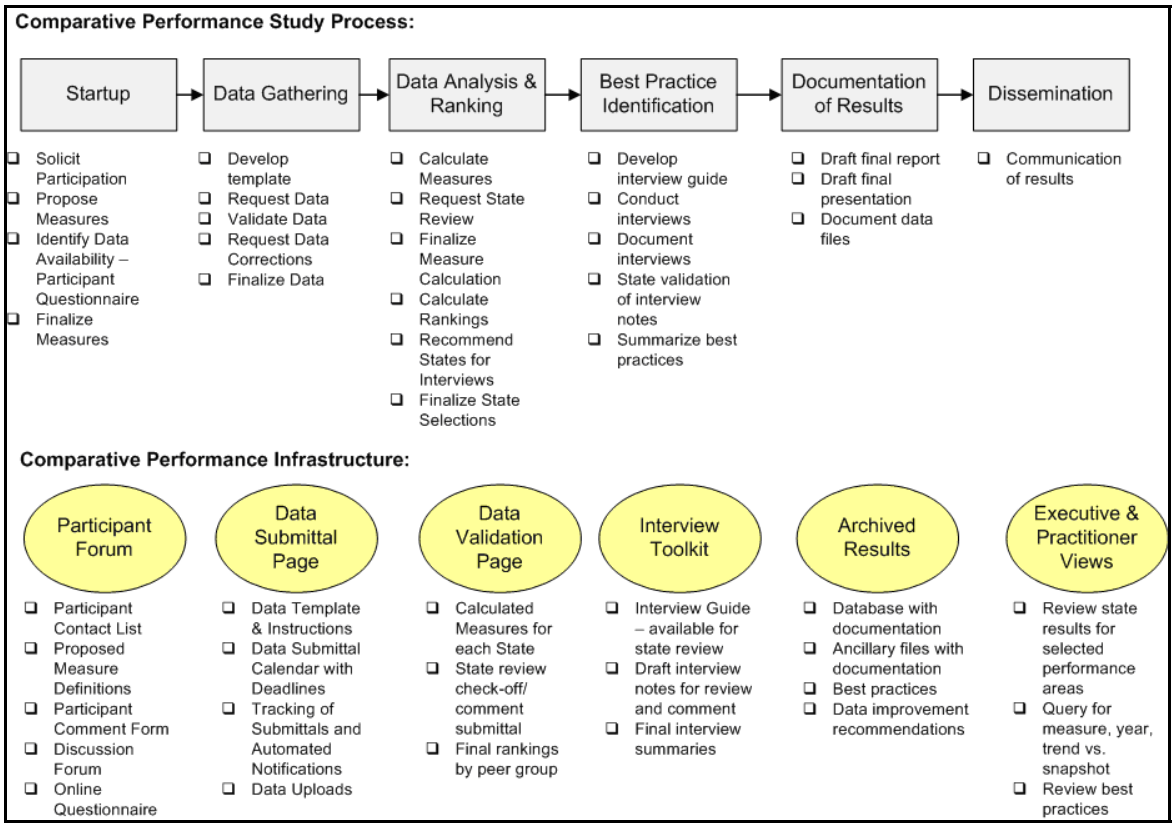

Figure 1. Comparative Performance Infrastructure – Study Process View

These elements are:

- **Participant Forum** – an electronic forum that facilitates the processes of gathering information from participants and reaching consensus on measures and data elements.
- **Data Submittal Page** – a web page where participants could go to download standard templates for data submittal, submit questions on the template, view a calendar of data submittal deadlines, and upload their completed data sets.
- **Data Validation Page** – a web page where participants could go to view the initial calculations of performance measures for their state, make comments, and indicate their approval of the calculations. This page could also be used to post the final rankings for all participating states.
- **Interview Toolkit** – a web page that supports the process of interviewing top performing states to identify potentially effective practices. Features could include posting of the interview guide, posting of draft interview notes, ability for states to transmit comments on the notes, and review the final interview summaries.
- **Archived Results** – database containing raw and processed data collected for the study, as well as web pages providing access to information about best practices and data improvement recommendations.

- **Executive and Practitioner Views** – views of computed performance measures and ranks, both for executives (i.e. simple displays of measures and rankings) and practitioners (i.e. more detailed options for slicing and dicing of the data).

Together, these six elements would provide a comprehensive set of features for supporting the entire CPM study process. The first four of these six elements would provide a standard, centralized infrastructure in support of processes currently accomplished by individual CPM study contractors. The web page set up for the repeat of the On-Time, On Budget CPM project provides a model that includes selected features to support the study process – including background information for states considering participation, description of data submittal requirements, a map showing participating states, and access to the data collected.

The last two elements would provide capabilities that do not currently exist. These elements also represent the most fundamental or core features of the infrastructure. *Therefore, the focus of the conceptual design for the infrastructure is on supporting the documentation and dissemination of results phases of the CPM study process through provision of an archive for results and tools for viewing the information.*

Requirements for the CPM Infrastructure

Based on these priority infrastructure elements, the conceptual framework for the CPM infrastructure is based on the following requirements:

- **Central data storage** – Provide a central repository for storing data collected as part of the CPM studies.
- **Consistent structure** – Provide a common structure for CPM data in order to ensure the standardization needed for storing and viewing data using a common toolset.
- **Standard data feeds** – Define and manage regular data updates from national data sources as needed to support multiple CPM studies.
- **Executive-level views** – Provide simple views of the data that allow executives to obtain a quick picture of their state's performance relative to peer states.
- **Practitioner views** – Provide the ability for practitioners responsible for different performance areas to conduct more detailed exploration of the performance data.
- **Documentation** – Make available detailed information on definitions and derivations of performance measures and associated data.
- **State practices** – Provide convenient access to information about state practices associated with states showing good performance results, as identified by the CPM studies.

The following section presents specific scenarios for how the infrastructure would be used.

CPM Infrastructure Users and Use Scenarios

Given the stated requirements presented above, several use scenarios were developed to guide the CPM infrastructure design. Scenarios cover activities that different classes of end users

might perform as well as activities that would be performed by CPM study contractors or administrators of the CPM web site. Each use scenario is presented in Table 1.

Table 1 – CPM Infrastructure Use Scenarios

Scenario	User Type(s)	Description
1. View CPM Data	Agency executive	• Find out state/agency rank relative to neighboring states for all available measures • Review state/agency performance trend information relative to other states • Pull data on relative performance for incorporation into presentation
2. Analyze CPM Data	Practitioner: Engineer/Planner/ Analyst within a specific functional area	• Analyze performance measure values relative to peer states, drilling down to detailed results • Review measure definitions • Review characteristics of each data set • Download data to spreadsheet for further analysis • Identify peer states based on socio-economic and travel characteristics • Select set of peer states and save for use in future queries
3. Scan CPM Study Findings	Agency executive or Practitioner	• Scan best practices identified across all performance areas
4. Review CPM Study Findings	Practitioner within a specific functional area	• Scan best practices identified for a specific performance area, and follow-up with selected agencies • Review reference list from CPM study

Scenario	User Type(s)	Description
5. Refresh Common Data	CPM Study Team or Site Administrator	• Refresh state characteristics data • Refresh county characteristics data
6. Refresh National Data	CPM Study Team or Site Administrator	• Refresh FARS data (safety) • Refresh VMT data (safety) • Refresh NBI data (bridge)
7. Add New CPM Data	CPM Study Team	• Provide data for new study • Provide data for repeat study • Specify data transformations
8. Configure Data Views	Site Administrator	• Add new performance area (e.g. Congestion) • Add new measures to existing performance area • Configure data viewing options • Implement data transformations/ETL scripts
9. Manage Access	Site Administrator	• Define and manage standard roles and access privileges • Register users and assign roles
10. Get Support	Any	• Request access • Request assistance • Report a bug • Suggest enhancement

4. CPM Components

Based on the use scenarios listed above in Table 1, the following components should be included in the CPM infrastructure:

- A CPM Database including:
  - definitions and metadata;
  - raw performance data and calculated performance results;
  - data transformation procedures; and
  - national data with information that can be used to construct groups of peer states, with standard processes and protocols for refreshing these tables from external sources (e.g. U.S. DOT, Census Bureau).
- A CPM Web Site including:
  - reporting and query tools that support data analysis and exploration for executives and practitioners;
  - access to information gathered on practices identified in states that are in the top tier of performance among their peers;
  - administrative tools for maintaining and extending the CPM web site and data; and
  - documents and links of interest.

An overview of the CPM Infrastructure is shown in Figure 2. Each infrastructure component is discussed below.

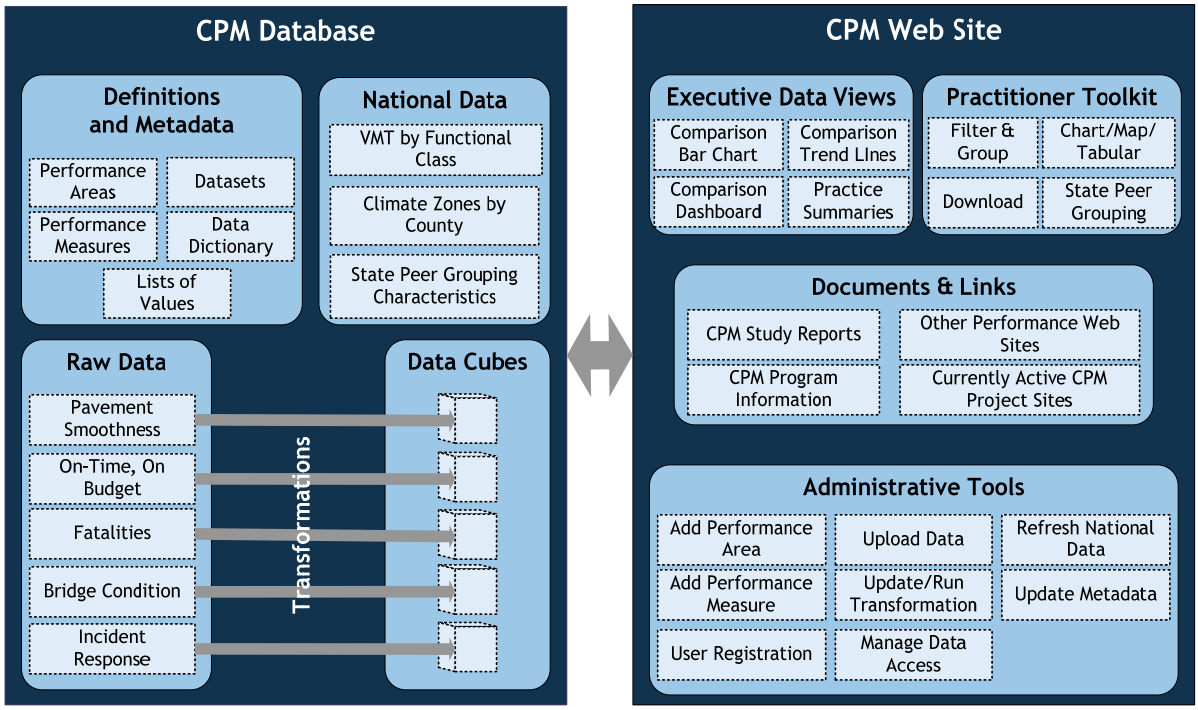

Figure 2. Comparative Performance Infrastructure Components

Performance Data Definitions and Metadata

The foundation for comparing performance across states is a set of shared definitions of performance measures and their derivation. For the pavement smoothness study, for example, these definitions might include those displayed in Table 2.

Table 2 – Example Template for CPM Data Definitions

Type	Description
Measurement Standards	Interstate pavement smoothness is measured by an AASHTO M328-10-compliant inertial profiler. Measurements should be collected in the right-most lane on each direction of every mainline Interstate highway section (no ramps). The MRI statistic should be reported for each 0.1 mile section of road, calculated by averaging IRI values from the left and right wheelpath profiles. Profiler operation and maintenance should be compliant with AASHTO R57-10.
Reporting Standards	Pavement smoothness data sets provided for comparative performance should include the following items in order to construct homogeneous data sets for comparison: state name, county, month and year of measurement, bridge indicator (is section on a structure – Y/N) pavement type, functional classification, section on Turnpike/Toll Road, year of last full depth reconstruction, year of last resurfacing (>1 inch of material added), AADT, Percent Trucks.

Type	Description
	In addition, a separate summary indicating the lane-miles of Interstate pavement treatments by calendar year (broken down by reconstruction, rehabilitation and other), should be provided.
Accuracy Expectation	Expected accuracy for cross-state comparison purposes is in the 10-15% range (based on findings from a 2004 FHWA Profiler Roundup). Variations in IRI measurements across states may be due to differences in environmental conditions, profiler calibration, operator training, and adherence to standards.
Performance Measure Definitions	Average IRI – the sum of IRI time section length for each reported section (in a given section grouping) divided by the total length of all reported sections (in the grouping). Lower values indicate better performance. Percent smooth pavements – cumulative percentage of directional Interstate mileage with IRI less than 94 inches per mile. Higher values indicate better performance.

The Austroads Road Transport performance site[1] provides a good model for structuring extended information about performance measures and their derivation. Each performance measure included in the site has the following associated information:

- **Qualifications** – detailed description of factors impacting data quality and consistency
- **Considerations** – discussion of how performance measures and trends are to be interpreted
- **Methodology** – measure definition, classification, units, aggregations, assumptions and limitations, source, calculation methods and collection frequency

Data definitions and procedures developed within project NCHRP 20-24(37)G can provide baseline information for this component of the CPM infrastructure, though there may be variations in data definitions and reporting standards for measures intended for tracking performance at the national level from those intended for comparison across states.

Descriptive metadata about each data set as well as standard metadata for each database table and column should be maintained and stored as part of the CPM infrastructure. This metadata is integral to the management of historical data sets and will ensure that users can obtain an understanding of data sources and derivations. It will be particularly important to note any variations from established data standards that are present in the data sets, as well as any data updates or adjustments that have been applied to correct for quality issues. This information can be combined with information on performance measure definitions and calculation methodologies for presentation to the user.

[1] http://algin.net/austroads/site/index.asp?id=2

The Federal Geographic Data Committee's (FGDC) Content Standard for Digital Geospatial Metadata[1] (CSGDM) provides a useful template that is applicable to CPM data sets – despite the fact that data may not typically be provided in a geospatial format. Portions of this standard that are most applicable are:

- **Identification** – data set name, description/abstract, source/originator, publication date, time period for content, access and usage constraints, point of contact
- **Quality** – completeness, consistency, processing description, source information
- **Entity and Attribute information** – description of each entity/table; description/type/length/domain for each attribute

Performance Data

The core CPM database would provide a central repository for storing the data collected within each of the CPM program projects. It would also provide the data source that would feed data display, query and reporting tools. Key design considerations are summarized below.

Content. The initial content of the database will consist of performance data for Infrastructure (Smooth Pavements, Bridge Condition), Safety (Fatalities), Traffic Operations (Incident Response), and Project Delivery (On Time, On Budget). The database will need to be designed for future extensibility to accommodate new performance measures within a given performance area (e.g. addition of serious injuries to the Safety area), and new performance areas. In order to provide extensibility, the initial database should include generic tables for storing and managing performance measures, definitions, metadata, and value lists that can be used to drive pick lists. Data cubes can be implemented using generic data structures as well to ensure extensibility and compatibility with business intelligence tools. Raw performance data structures will need to be customized for each data set.

Raw vs. Processed Data. The design of the CPM database needs to distinguish between the raw data from which performance measures are derived, and the processed data used for display and analysis of performance measure results. For example, in the bridge condition performance area, the raw data consists of records representing individual bridges and their conditions for specific years. The processed data consists of aggregated data with measures such as the percent of deck area on structurally deficient bridges. In the safety area, raw data consists of individual fatality records. Processed data consists of fatality rates computed using VMT data from FHWA, averaged across multiple years. (See Chapter 5 below for further details on raw and processed data for CPM measures.) Preserving raw data allows for additional analysis to be conducted, and potentially for new variations of performance measures to be derived. Including both raw and processed data in the infrastructure database allows for the raw data structures to change over time as CPM studies are repeated while minimizing impacts on the processed data that is used to drive the data views.

[1] http://www.fgdc.gov/metadata/documents/workbook_0501_bmk.pdf

Data Cubes. Storing the processed data in "cube format" can facilitate data exploration. Data cubes can be visualized as a multi-dimensional spreadsheet in which the cells contain the performance measures and the rows and columns are dimensions available for "slicing and dicing" the data. For example, in the case of pavement smoothness, dimensions would include state, pavement type (flexible vs. rigid), and climate zone.

Data Access Restrictions. The project delivery and pavement smoothness CPM projects have calculated measures based on data provided by participating states. States were promised anonymity as a condition for providing their data. While permission can be sought to open up access to this data to a select audience, the infrastructure should be designed to support access restrictions based on established agreements. By default, each state can be shown using a number rather than a state name on the displays. Based on login information, a user would be able to identify their state. They would also be able to see their performance relative to a group of peer states meeting selected characteristics. Note that the issue of anonymity needs further consideration. Anonymity allows states with lower performance to avoid any stigma, and for some, this is an important factor in their decision to participate in a CPM study. On the other hand, keeping states anonymous does limit the richness and value provided by the data. Many states want to know where they stand relative to particular states that they view as their peers (as opposed to larger peer groups based on region, size, climate zone, etc.) Anonymity also limits transparency, which may impact perceptions about the credibility of the data.

Storage of Raw Performance Data from National Sources. The bridge and safety CPM projects to date have calculated performance measures from national data sources – the National Bridge Inventory (NBI) and the Fatal Accident Reporting System (FARS) respectively. Where CPM projects derive data from national sources, there is a decision to be made as to whether the CPM infrastructure should store snapshots of the raw data used to calculate measures, or just the computed measures. Storage of a snapshot of the data from national sources would provide greater flexibility to add new computed measures. It would also provide greater stability in the measures themselves, insulating the calculated measures from updates to the raw data in the live sources. However, it would also substantially increase the size of the database.

Geospatial Linkages. Data collected to date within the CPM program has not been consistently linked to specific facilities or geographic units below the state level. This is because the emphasis of the program has been on aggregated statewide performance data rather than performance on specific facility locations. An assumption is that the CPM would not be required to support detailed mapping of performance at the facility level. Any mapping features would display summary data at the state level only.

Data Transformation Procedures and Tools

Business rules for extracting data from national sources to populate "raw" data structures, and for transforming raw data to processed, cubed data that can be used to drive the data exploration tools need to be clearly documented to ensure consistent calculations of performance measures over time. At a minimum, these rules should be documented in text form and archived with each data set. Ideally, the rules should be formally codified in the form of scripts that can be applied to perform the needed data transformations. A business

intelligence software toolset should be considered for management of the raw data and application of transformation scripts. Such a toolset would also provide user interface components for data exploration.

National Data for Peer Grouping and Common Groupings

Another important component of the CPM infrastructure is a compilation of characteristics about each state that can provide the basis for identification of peer states for purposes of performance comparisons. Relevant characteristics may vary by performance measure. For example, for on-time, on-budget project delivery the total program size or average project size may be of most relevance. For safety, the degree of urbanization or population density would be of particular interest. For pavement and bridge condition, climatic factors are relevant. The following table lists information (and sources) that can be compiled and maintained as part of the CPM infrastructure, in support of multiple performance areas. Note that this information should be maintained in time series form, with periodic updates (based on frequency of source data updates). This would allow performance trends to be viewed within the context of socio-economic or agency financial trends. Items listed below provide a starting point, which could be augmented over time as additional needs are identified.

Note that several of the indicators listed below (and many others) that can be used for state peer grouping are compiled and published as part of FHWA's Highway Statistics series[1]. AASHTO's Center for Excellence in Project Finance web site[2] also presents a number of state-by-state demographic, highway network and financial statistics for purposes of benchmarking that the infrastructure could draw upon.

Table 3 – CPM Peer Grouping Data Elements

Peer Grouping Element	Description and Source(s)
Climate Zone	Climate zones used for the Long Term Pavement Performance (LTPP) program – Wet Freeze, Dry Freeze, Wet No-Freeze, and Dry No-Freeze. Linked to county codes. The LTPP climate zones (as of 2010) are being used by HPMS. FHWA is maintaining climate zone boundaries and climate zone assignments for each county.
Region	Pre-defined regions: AASHTO Regions (I-IV), Census Regions (West, North Central, Northeast, South Atlantic, South Gulf)

[1] http://www.fhwa.dot.gov/policyinformation/statistics/2008/ps1.cfm

[2] http://www.transportation-finance.org/tools/state_by_state/

Peer Grouping Element	Description and Source(s)
Socio-economic	Total state population, population density, percent elderly (>65) percent urbanized population, median household income. Data available from US Census Bureau.
Travel	Vehicle miles of travel by functional class. Data available from FHWA based on HPMS.
System Extent	Centerline miles and lane miles by functional class.
AADT/Lane on Structures	General classification of bridge traffic loading. Source is National Bridge Inventory
Bridge Age Distribution	General classification of bridge network age. Source is National Bridge Inventory

Peer grouping data would be provided in a downloadable tabular display, to be used for identification of a group of peer states.

Performance Result Exploration and Reporting Tools

The first two usage scenarios displayed in Table 1, View CPM Data and Analyze CPM Data, require provision of a data query and reporting component.

The *View Data* function is targeted at agency executives, and would include simple graphics for identifying the state's current performance relative to other states, and viewing trend information. The ability to copy information for inclusion in reports or presentations would be provided in addition to standard charts viewable on screen.

The executive view could be implemented as a basic set of bar graphs (with one bar per state showing a selected performance indicator) – as shown in Figure 3:

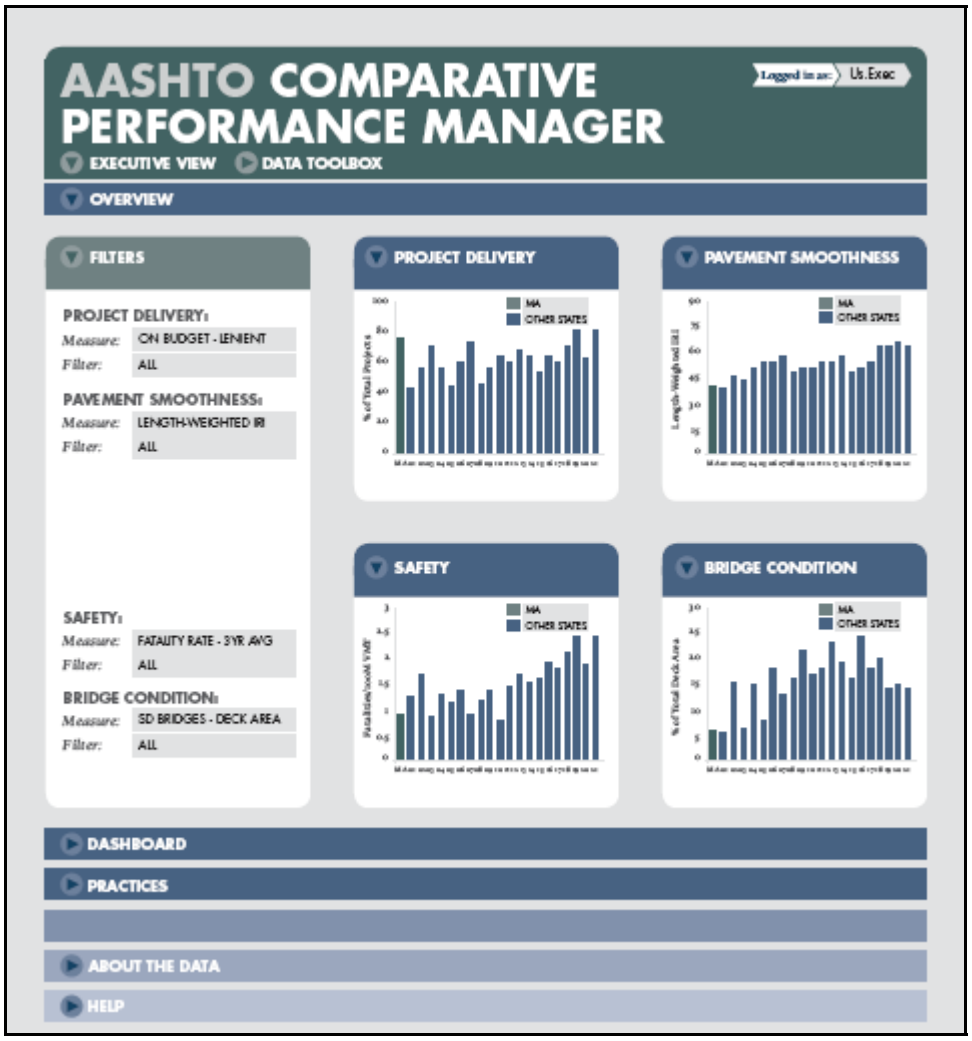

Figure 3 – CPM Executive View – Comparison Bar Chart

Options would be provided to select a measure within each available performance area, as well as a filter condition (e.g. show IRI for only concrete pavements).

The bar chart view could be augmented by the dashboard view shown in Figure 4 providing a display of a selected state's performance together with the mean or median value for all other states (or a set of peer states).

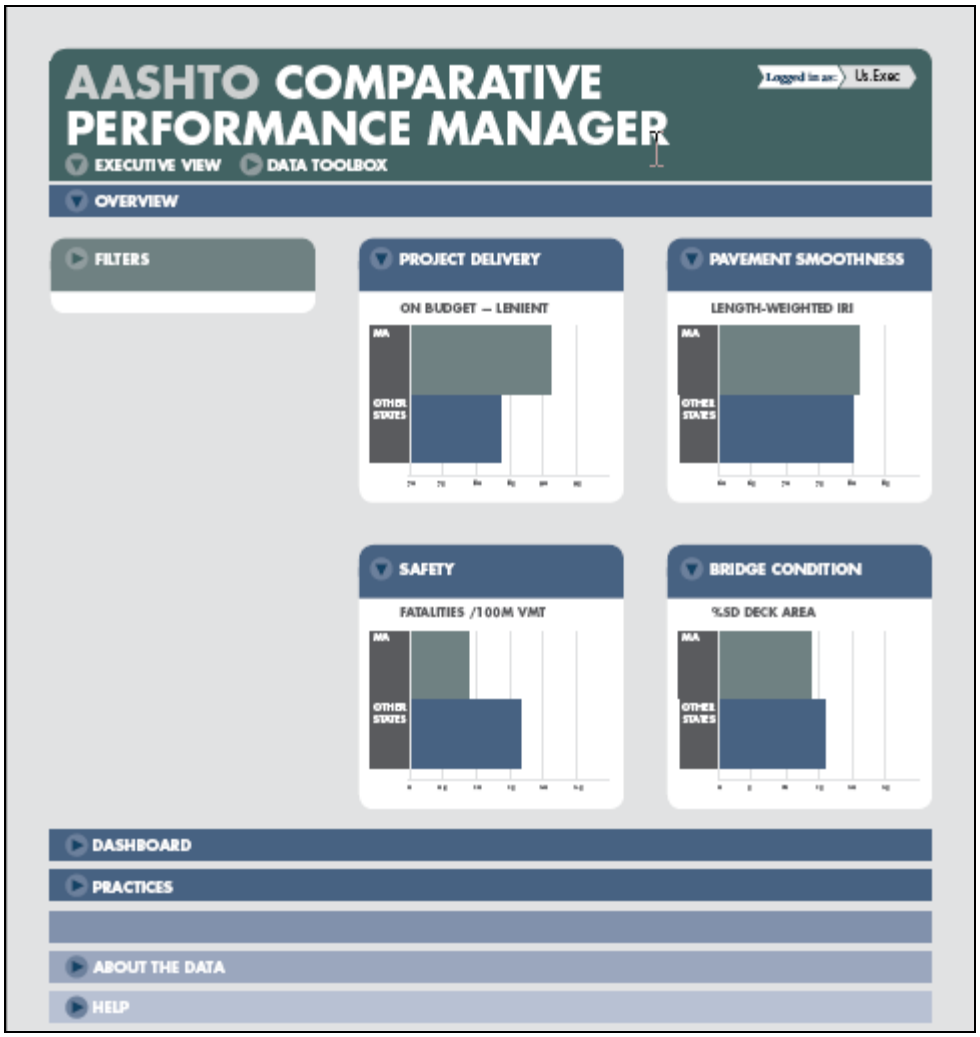

Figure 4 – CPM Executive View - Dashboard

This approach is used in the Transportation For America performance web site, as shown in Figure 5.

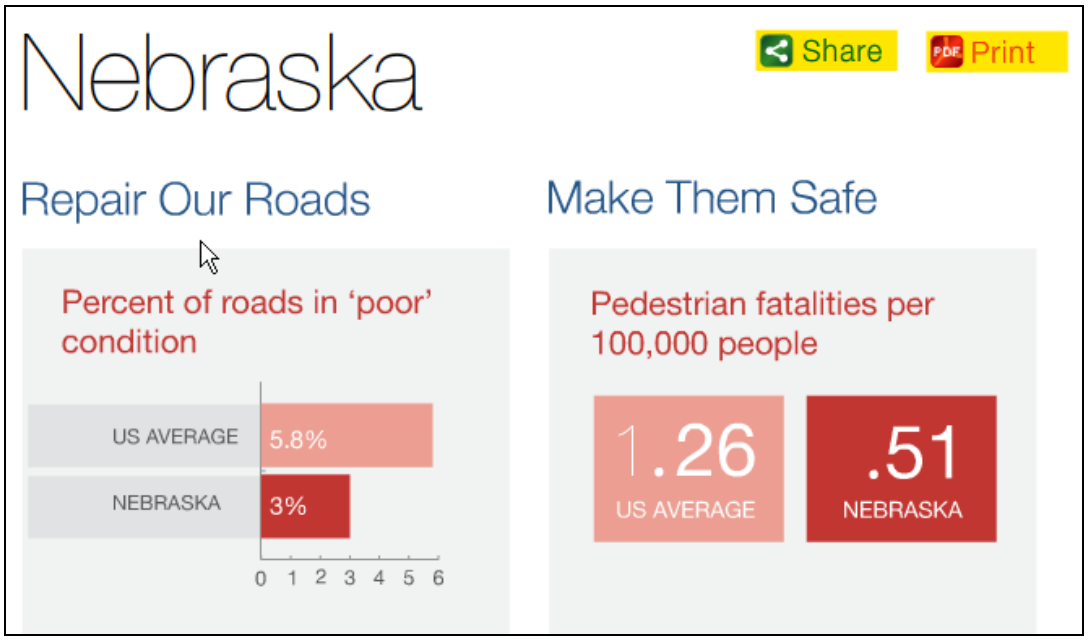

Figure 5 – State Performance Versus National Average View from Transportation For America

As shown in Figure 6, trend information could be displayed using a line graph, with time on the horizontal access and the performance measure of interest on the vertical axis. In order to reduce clutter, a set of checkboxes could be provided for selection of states to be included in the view.

Figure 6– CPM Executive View - Trend Comparison

The *Analyze Data* function (see Table 1) is targeted at a planner/analyst/engineer within a state DOT or research agency. This function would provide a set of data toolbox views to allow for more detailed exploration of selected measures. Whereas the executive view displays all of the available measures on a single screen, the practitioner view would have separate tabs for each performance area. As shown in Figure 7, a quick view could be provided to allow for display of different measures within an area – e.g. IRI for rigid and flexible pavement, or trends in motorcycle versus all fatalities. Options for basic tabular, chart and map views could be provided.

Figure 7 – CPM Practitioner Quick View

A separate screen for downloading a data set of interest could also be included, as illustrated in Figure 8. This would allow those wishing to do more detailed analysis with the data to use their own toolsets. Figure 9 shows an example of an interactive trend report for exploration of query results. This is based on a similar view provided by Google Public Data Explorer[1].

[1]http://www.google.com/publicdata/home

Figure 8 – CPM Practitioner Query Screen

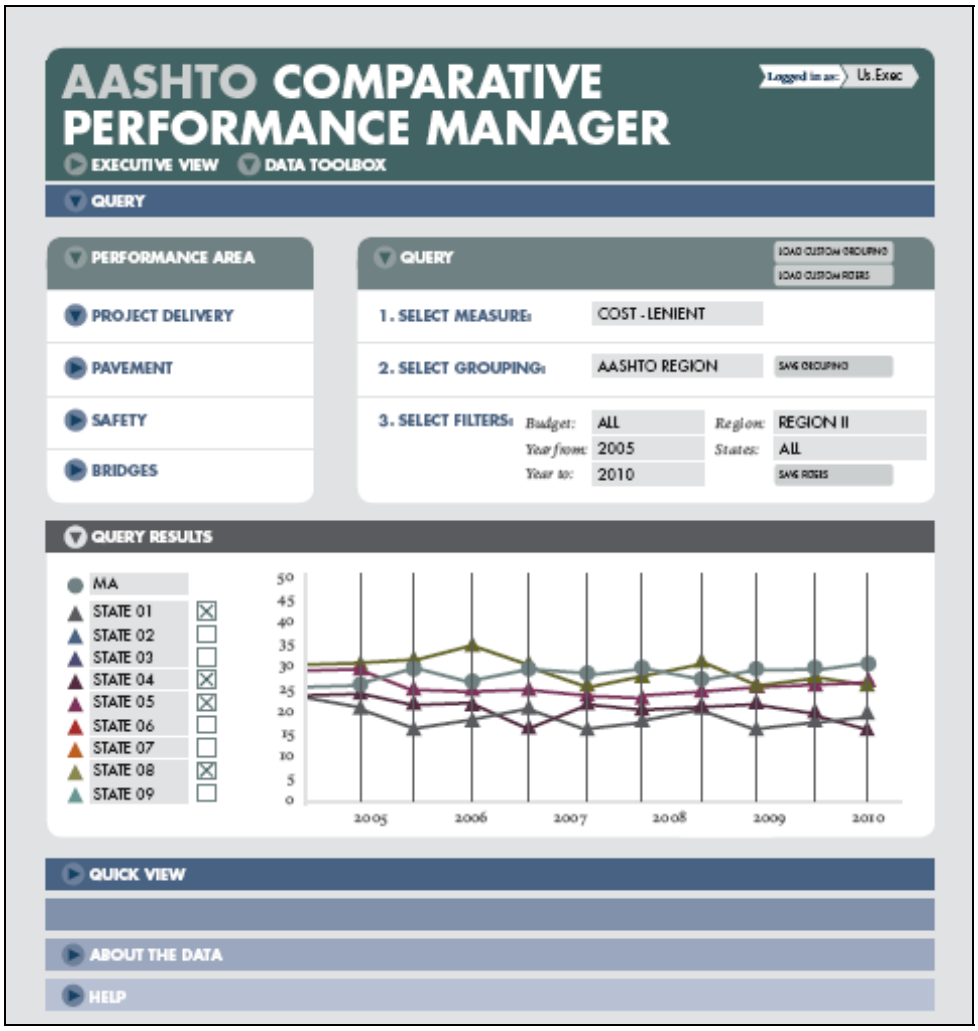

Figure 9 – CPM Practitioner Interactive Query Results

State-Reported Practices Supporting Performance Improvement

The purpose of this component is to provide an easily accessible repository of practices that have been linked to achievement of good performance results as part of the CPM studies. While each study will continue to produce a report (which can also be posted to the CPM web site), highlighting practices within each performance area in conjunction with views of performance data would provide additional visibility for these important research results.

CPM Administration Tools

This component of the infrastructure would be used by the site administrator to add performance measures to existing areas, set up a new performance area, populate meta data, refresh or modify data feeds from public data sources, load new performance data sets, and create/update and run data extraction and transformation (ETL) scripts. If resources are

constrained, administrative tools could be rudimentary at first, and upgraded over time to facilitate direct data uploads by contractors engaged to perform CPM studies.

Documents and Links of Interest

The CPM web site would include general information about the Comparative Performance Measurement program, links to completed CPM study reports, notices of current studies – highlighting those currently soliciting participants, links to web sites that have been established for administration of those studies, and links to other performance data sites.

5. CPM Data Structure

Database Contents

Table 4 identifies the contents of the CPM database at a conceptual level.

Table 4 – CPM Database Contents

Element	Description
Performance Area Definition Template	A table including one record per performance area (e.g. safety, pavement smoothness, bridge condition), with measurement standards, reporting standards, accuracy expectations, and other descriptive information about the performance area.
Performance Measure Definition Template	A table including one record per performance measure (e.g. 3 year moving average fatality rate), with definitions and calculation methodologies.
Data Set Documentation Template	A table including descriptive information about each distinct data set within the CPM infrastructure – including date collected, number of states included, etc. This table (or an associated table) could also store access restrictions based on agreements for anonymity with states.
Data Dictionary	A table with descriptive metadata for each table and column in the database.
Lists of Values	A set of tables with lists of values for the filtering/classification variables in the system (e.g. pavement type, functional class, etc.)
Raw Performance Dataset Tables	A set of tables with raw data assembled for each CPM study. See below for detailed documentation by performance area.
Processed Performance Data (Cubes)	Multidimensional data sets – one for each performance area. See below for detailed documentation by performance area.
State Characteristics	A table with peer grouping characteristics by state, including a year field to allow for time series data.
Climate Zones	A table linking climate zones to counties.
VMT by Functional Class	A table used for calculation of fatality rates for urban and rural roadways within each state.

Data Elements for Specific Performance Areas

The tables below identify data elements for both the raw and processed data sets for pavement, bridge, safety and project delivery performance areas. Information can be added for the incident response area when it becomes available.

Pavement Smoothness – Raw Data

One record per pavement section per year.

Data Element	**How Used:** ID – identification L- link to other data M – performance measurement W – weighting factor G – primary grouping variable F - additional grouping/filtering variable
State	ID
Year	ID
Section ID	ID
County	L (to Climate Zone)
IRI	M
Length	W
Pavement Type	G
AADT	W
% Trucks	E
Last Resurface Year	F
Last Reconstruct Year	F
Bridge	F
Toll Road/Turnpike	F

Pavement Smoothness – Processed Data

One record per unique combination of each dimension value.

Data Element	**How Used:** Fact/Dimension
Length-Weighted IRI	Fact
AADT-Weighted IRI	Fact
Length Very Smooth	Fact
Length Smooth	Fact
Length Acceptable	Fact

Data Element	How Used: Fact/Dimension
Length	Fact
State	Dimension
Year	Dimension
Pavement Type	Dimension
Climate Zone	Dimension
Urban/Rural	Dimension
On Toll Road	Dimension
On Bridge	Dimension

Bridge Condition – Raw Data

One record per NBI bridge per year

Data Element	How Used: ID – identification L- link to other data M – performance measurement W – weighting factor G – primary grouping variable F – filtering variable
State	ID, L (to census region, AADT category, Bridge Age category)
NBI Reporting Year	ID
Structure Number	ID
County	L (to climate zone)
Deck Rating	M
Superstructure Rating	M
Substructure Rating	M
Sufficiency Rating	M
Open/Posted/Closed	M
Status (Structurally Deficient)	M
Deck Width out-to-out	W
Structure Length	W
Record Type (on/under bridge)	F
Type of Service	F
Maintenance Responsibility	F
Design Type	F

Data Element	How Used: ID – identification L- link to other data M – performance measurement W – weighting factor G – primary grouping variable F – filtering variable
Functional Class - Interstate	F
Functional Class – Urban/Rural	G
Inventory Route System (NHS)	F
National Truck Network	F
AADT (coded into groups)	G
Year Built Constructed	G
Year Last Replaced	G

Bridge Condition – Processed Data

One record per unique combination of each dimension value.

Data Element	How Used: Fact/Dimension
Deck Area on SD Bridges	Fact
Deck Area on Bridges with Deck, Superstructure and Substructure Ratings 7 or higher	Fact
Deck Area on bridges with SR>=50	Fact
Deck Area on posted bridges	Fact
NBI Year	Dimension
Climate Zone	Dimension
Urban/Rural	Dimension
Interstate	Dimension
National Network	Dimension
AADT Distribution Class	Dimension (at state level)
Age Distribution Class	Dimension (at state level)

Safety – Raw Data

One record per state per year (Note: raw data is aggregated up from data in FARS)

Data Element	**How Used:** ID – identification L- link to other data M – performance measurement W – weighting factor G – primary grouping variable F – filtering variable
State	ID
Year	ID, L (to VMT data)
Crash - Number of Fatalities in Crash	M
Crash – Roadway Function Class	G, L (to VMT data)
Person – Person Type	G
Person - Age	G
Person – Injury Severity	M
Person – Restraint System Use	F
Vehicle – Body Type	F
Driver – Driver-Related Factors	F

Safety – Processed Data

One record per unique combination of each dimension value.

Data Element	**How Used:** Fact/Dimension
Fatalities-Moving average	Fact
VMT-Moving average	Fact
Speeding-related fatalities-Moving average	Fact
Alcohol-related fatalities -Moving average	Fact
Young driver fatalities-Moving average	Fact
Unrestrained occupants in fatal crashes-Moving average	Fact
Pedestrian Fatalities-Moving average	Fact

Data Element	How Used: Fact/Dimension
Motorcycle fatalities-Moving average	Fact
Unhelmeted Motorcyclist fatalities-Moving average	Fact
State	Dimension
Year (end year for moving averages)	Dimension
Urban/Rural	Dimension

Project Delivery – Raw Data

One record per contract.

Data Element	How Used: ID – identification L- link to other data M – performance measurement W – weighting factor G – primary grouping variable F – filtering variable
State	ID
Contract ID	ID
Original Bid Award Amount	M
Actual Project Cost	M
Final cost adjustment for performance bonus or penalty, if any	M
Final cost adjustment for inflation in materials costs, if any	M
Originally Specified Completion Date	M
Current Specified Completion Date	M
Actual Substantial Completion Date	M
Original Workdays Allowed	M
Current Workdays Allowed	M
Actual Workdays Charged	M

Data Element	How Used: ID – identification L- link to other data M – performance measurement W – weighting factor G – primary grouping variable F – filtering variable
Final Voucher Date	M
Notice to Proceed Date	M
ARRA Funded?	G
Project Type	G
Contract Type	G
Contract Value	G

Project Delivery – Processed Data

One record per unique combination of each dimension value.

Data Element	How Used: Fact/Dimension
# Contracts completed on budget (strict)	Fact
# Contracts completed on budget (lenient)	Fact
# Contracts completed on time (strict)	Fact
# Contracts completed on budget (lenient)	Fact
Total Value of Contracts completed on budget (strict)	Fact
Total Value of Contracts completed on budget (lenient)	Fact
Total Value of Contracts completed on time (strict)	Fact
Total Value of Contracts completed on budget (lenient)	Fact
Time overrun (days)	Fact
Cost overrun (dollars)	Fact
State	Dimension
Completion Year	Dimension
Project Type (ARRA, Major/EIS)	Dimension

Data Element	How Used: Fact/Dimension
Contract Type	Dimension
Contract Value Category	Dimension

6. Implementation and Stewardship Approach

Overview

Work to define implementation and stewardship options was based on the following objectives:

- *Build* an initial infrastructure that compiles and provides access to comparative performance measure (CPM) data collected to date.
- *Establish* a sustainable funding and stewardship model so that the database and web site can be maintained over time to support future CPM efforts and ongoing data access.
- *Leverage* existing data collection and performance measurement and reporting initiatives through partnerships to minimize duplication of effort and implementation cost.
- Provide an *incremental* or phased approach to allow for implementation and operation to begin with limited resources.

Options were developed with consideration of three distinct sets of implementation activities:

- *Development* – how will development of the database and CPM web site be funded and carried out?
- *Hosting and Management* – what organization will have primary responsibility for database and web site operation, and how will these activities be funded?
- *Data Management* – what organization(s) will be responsible for compiling, quality-checking, and loading data into the system?

The following section of this chapter discusses key implementation assumptions and considerations. Then, specific activities required for implementation and ongoing management of the infrastructure are identified and the associated level of effort and rough cost for these activities is estimated. Finally, three stewardship options are presented with an assessment of their implications on overall cost, lead time, flexibility and user experience.

Key Implementation Considerations

Audience

Implementation and stewardship options were developed with the assumption that the target audience for the infrastructure would be transportation agency executives and practitioners seeking to understand their performance relative to their peers and identify opportunities for improvement. This assumption was based on input from the project panel in Task 1 of this study. A public-facing web site for national transportation performance can be envisioned to include some of the same measures, but the views and content would likely need to be tailored for a general audience that has less background and context for the data being presented, and a

different set of uses for the information. Given current national interest in public accountability and open data, other independent platforms for data display could be pursued for a public facing site. For example, State Of The USA (stateoftheusa.org) has developed an architecture for a key national indicator system and has a web site in beta testing for viewing different indicators at national and state levels. The CPM infrastructure could potentially be a feeder to public facing sites in the future, while maintaining its focus on serving the internal strategic planning and decision making needs of transportation agencies. Any information fed by the CPM infrastructure to public facing sites would be by agreement with the participating states that provide CPM information.

System Life Cycle

The CPM infrastructure described in this report was designed for the purpose of supporting AASHTO's CPM program. Therefore, the life cycle of the database and web site would track with the CPM program at AASHTO. After the initial infrastructure was rolled out, data for existing measures would be refreshed on a cyclical basis, and data and views for new measures would be added as new comparative measures came on line. If at some point the CPM program is discontinued, the database and site would be kept open for a limited time (e.g. 6 months to one year) and then archived.

AASHTO's CPM program has been in place for five years, and can continue to play an important role in serving as a catalyst for maturation of national transportation performance measures. Project 20-24(37)G – *Technical Guidance for Deploying National Level Performance Measures* has identified only 5 measures classified as ready for deployment; and 12+ others that require further work before they can be implemented. The projects in the 20-24(37) series have provided a means to learn about practical issues surrounding performance measurement for different goal areas. They have also produced data sets that are valuable for understanding not only high-level performance comparisons across states, but additional nuances of comparative performance measurement. For example, exploration of multiple measures in the bridge condition project showed how selection of "top performing states" depended on which measure of bridge condition was used. When data for multiple performance areas are combined in a central location, there will be additional opportunities to gain insight about factors contributing to variations in performance across states.

The CPM infrastructure would contain more measures with more detail and finer breakdowns than would be appropriate for a general purpose national transportation measurement web site. Since the CPM program is exploratory and evolutionary in nature, only a portion of the data collected for the CPM program is suitable for public display and presentation as national measures. For this reason, as noted above, the initial CPM infrastructure is not envisioned as a public facing site. While it is possible that the infrastructure could evolve in this direction and could certainly be a feeder (for selected measures) to public facing sites, decisions about costs and benefits of implementation should be based on the assumption that the infrastructure will exist primarily to support the AASHTO CPM program.

Data Access and Confidentiality

As designed, the infrastructure will not include sensitive data. Data made available from the web site would either be in aggregated form (i.e. without identification of specific roadway locations or individuals) or would be data that is already available from other public sources (e.g. the NBI, FARS, HPMS).

The main issue to be addressed with respect to confidentiality is the fact that existing data for selected measures was compiled from states with the agreement that state identities would not be disclosed. The guarantee of anonymity was an important consideration for some states in agreeing to provide data. Keeping state names anonymous does limit the value of the comparative performance information. However, prior commitments to data providers need to be honored.

As part of the infrastructure implementation process, states that provided data can be asked if they would agree to relax the anonymity condition. If a critical mass of states agree, then data can be provided with state names. Remaining states could then choose to have their data only included in aggregate (multi-state) summaries, or included in the state-specific displays but with coded identifiers rather than state names.

Going forward, states can be encouraged in future CPM projects to participate in a non-anonymous fashion, but provided the options above (identification by a code or inclusion only in the aggregate results.)

Implementation Activities and Costs

Initial Development

The initial development of the infrastructure will involve:

- **Database development** – design, build and populate a CPM database with data gathered to date from prior CPM studies
- **Web site development** – design and develop the CPM web site (including data display and analysis tools.)

Database Development. Designing and building the CPM database is the first logical step in the infrastructure implementation effort. This will ensure that the data, currently in a variety of formats and stored with NCHRP project files, are brought together, documented consistently, compiled into a single source, and preserved for future access. The fundamental database design and core data loading can be accomplished prior to the CPM web site design, though additional modifications to the database should be anticipated in conjunction with the web site implementation. Specific database development activities are:

- Obtain Existing CPM Data Sets and Documentation:
  - On-time, on budget (2007 and 2010)

  - Smooth pavements – Interstate IRI (2009)
  - Fatalities - FARS (2010)
  - Bridge condition - NBI (2009)
  - Incident response (2009)
- Analyze the existing data and finalize target CPM Database Design - use the logical database structure provided in Chapter 5 above as a starting point.
- Develop and document data transformation specifications from source to target tables – these transformations would filter and aggregate data into groupings that support "apples to apples" comparisons across states, and calculate performance measure values.
  - FARS data
  - NBI data
  - State-provided IRI data
  - HPMS IRI data
  - On-time, on budget data
  - Incident response data
- Build and load database; conduct tests to ensure that the data are loaded correctly.
- Develop proposed data standards to facilitate future data exchanges, to include CPM meta data, state peer grouping characteristics, raw performance tables, and processed performance data.
- Prepare technical documentation for database administration.

Web site Development. The web site would include the views and features described in the Interim Report. The Interim Report includes screen shots that can guide the initial development. As the site is built, it would be desirable to identify a group of potential users to provide feedback on initial iterations. Web site development steps are:

- Identify user group to guide design process
- Develop initial web site design based on prototype developed under NCHRP 20-24(37)F
  - Menus and navigation
  - Data query and display tools
  - Data loading and validation tools
- Review with user group and refine
- Build site, allowing for 3-4 iterations of review and revision

Estimated effort and costs for developing the infrastructure (database and web site) are between 10-16 months of effort, equating to $200-300,000 (at a flat $120/hour rate). However, the initial infrastructure could be developed incrementally. For example:

- Phase 1: Define standard data structures, load existing data, publish data formats and make data available for use by state DOTs and others (e.g. post to a public web site at AASHTO, USDOT or DATA.GOV)
- Phase 2: Develop a central database and a basic web site with a set of simple executive views
- Phase 3: Add practitioner views with additional query and analysis features
- Phase 4: Automate data loading and transformations from source systems

Infrastructure Hosting and Management

Hosting and management of the database and web site will involve the following *technical* activities:

- Deployment on an internally managed server, external hosting service or cloud provider
- Monitoring to ensure availability and acceptable performance and troubleshooting problems
- Administering security/granting user privileges and permissions
- Backup and data recovery
- Tracking and reporting usage patterns
- Providing user support

It will also involve *administrative and liaison* activities to:

- Publicize the initial site launch
- Compile and respond to user and stakeholder feedback
- Plan and implement site enhancements
- Coordinate with new Comparative Performance studies so that new measures can be incorporated into the database and web site

The Interim Report estimated effort and costs for managing the infrastructure (database and web site) at between 3-5 person months of effort per year, equating to $30-50,000 (at a flat $120/hour rate). This is a high -level estimate that is dependent on the level of web site use (and associated support requirements), and the extent of enhancements made each year.

This estimate will also require refinement to include costs for disk space, bandwidth consumption and backup/recovery based on the selected hosting arrangement (e.g. in-house, hosted, cloud.) These latter costs will vary based on the size of the database and the volume

and nature of web site traffic. While a definitive database sizing analysis is dependent on a database design, a "back of the envelope" calculation was conducted to determine the likely order of magnitude database size. Based on the table structures in Chapter 5 , if the CPM database were to store processed data for the current five performance areas, and raw data for the incident management and project delivery measures – for all 50 states and 20 years, it would contain between 1 and 2 million records. Given that most of the data consist of numeric values, a 10 GB database would be more than sufficient – even accounting for additional staging tables and allowing for indexes, etc. If potential new performance measures are taken into account, a conservative assumption would be that the maximum database size would be 100 GB. Commercial hosting packages for relatively low transaction volume web sites with storage requirements on this order can be obtained for under $200 per month.

Data Management

Data management refers to procedures for data compilation, quality assurance, and loading (into the CPM database) – both to periodically refresh information for existing measures and to add data for new measures. Data management procedures will need to recognize variations across measures with respect to source and level of maturity. Standardized and automated data loading procedures can be established where national sources are available and measures are unlikely to change.

Data compilation, quality assurance and loading for new or emerging measures without established national data sources will need to be handled on a more flexible, ad-hoc basis in the context of CPM studies. Future CPM studies that repeat collection of existing measures or add new measures can be scoped to supply data in standard formats consistent with the database design. For repeat of existing measures a study task can be included to review and recommend modifications to existing data formats, insuring backwards compatibility so that trend information can be maintained. For new measures a study task can be included to recommend new standard formats.

Data without Established National Sources

Two of the five established CPM areas (incident response and project delivery) do not have established national data sources. The incident response CPM study involved only ten states and measures for this area are at a relatively early stage of maturity. A continuing strategy to refresh this data source has not yet been established. The current project delivery (on-time, on-budget) contractor (for Project 20-24(37)A) has developed procedures and tools for agencies to upload and validate data, as well as analytics to calculate measures from raw data. This contractor has also developed options for continued improvements to measures and annual collection of these data. The existing tools for project delivery can be incorporated into the master CPM web site. Alternatively, as these measures continue to evolve, the project delivery site could be maintained separately, with final data loaded into the master CPM site.

Data with Established National Sources

Data for bridge condition and safety have been derived from established FHWA and NHTSA data programs and then transformed (filtered, normalized, displayed for selected peer groups) for comparative analysis purposes. Bridge data used for CPM are collected as part of the FHWA National Bridge Inventory (NBI). Fatality data are part of the NHTSA Fatal Accident Reporting System (FARS). While the CPM project on pavement smoothness collected IRI data directly from states (based on the recommendations of participating states), these data are collected via the FHWA HPMS program, and recent enhancements should make the HPMS IRI data more suitable for CPM.

Going forward, where data can be derived from national sources, automated data transfer procedures should be sought to avoid duplication of effort. Prior CPM efforts have obtained extracts from the NBI (via file downloads) and FARS (via the online query tool) to provide the raw data for performance measure calculation. This manual data loading method is time-consuming, results in duplicate storage of the raw data and creates the potential for confusion when different measures are obtained from seemingly the same source. A more efficient approach would involve working with FHWA and NHTSA to provide either a standard data feed, or to arrange for application of direct database queries that pull out filtered and aggregated data for the CPM database. The former approach would be preferable given that CPM queries would need to be adjusted over time as FHWA and NHTSA make changes in their source data formats and definitions.

A strong partnership between AASHTO and FHWA and NHTSA would be beneficial for establishing and maintaining efficient data sharing mechanisms. Such a partnership would seek to establish a shared and clearly documented understanding of the source data and the requirements of the data consumer (represented by AASHTO SCOPM.) It would also provide a solid foundation for future expansion of available CPM data as FHWA and NHTSA data programs are extended.

Costs for data compilation, quality assurance and loading for measures without a national source will vary based on the measure and can be built in to CPM study efforts. Costs for refreshing annual data from national sources (HPMS, IRI, and FARS) will depend on methods established for data transfer. If manual methods are used, approximately 80 hours of effort each year (roughly $10,000) would be required to refresh the data. If fully automated methods are used, this could be reduced to 8 hours of effort for coordination and spot checking.

Stewardship Options

Three basic stewardship approaches were identified, all of which involve collaboration between AASHTO and FHWA.

AASHTO- Hosted with FHWA Participation

AASHTO would develop, host and maintain the infrastructure, which would include a joint AASHTO/FHWA Comparative Performance web site. FHWA would provide an annual data set from the NBI and HPMS to refresh pavement and bridge condition measures, as well as

additional peer grouping characteristics. This data feed would be automated following implementation of FHWA's data warehouse. The effort would be funded through a combination of AASHTO membership dues, FHWA funds and additional assessments for participating states as needed. This option is illustrated in Figure 10.

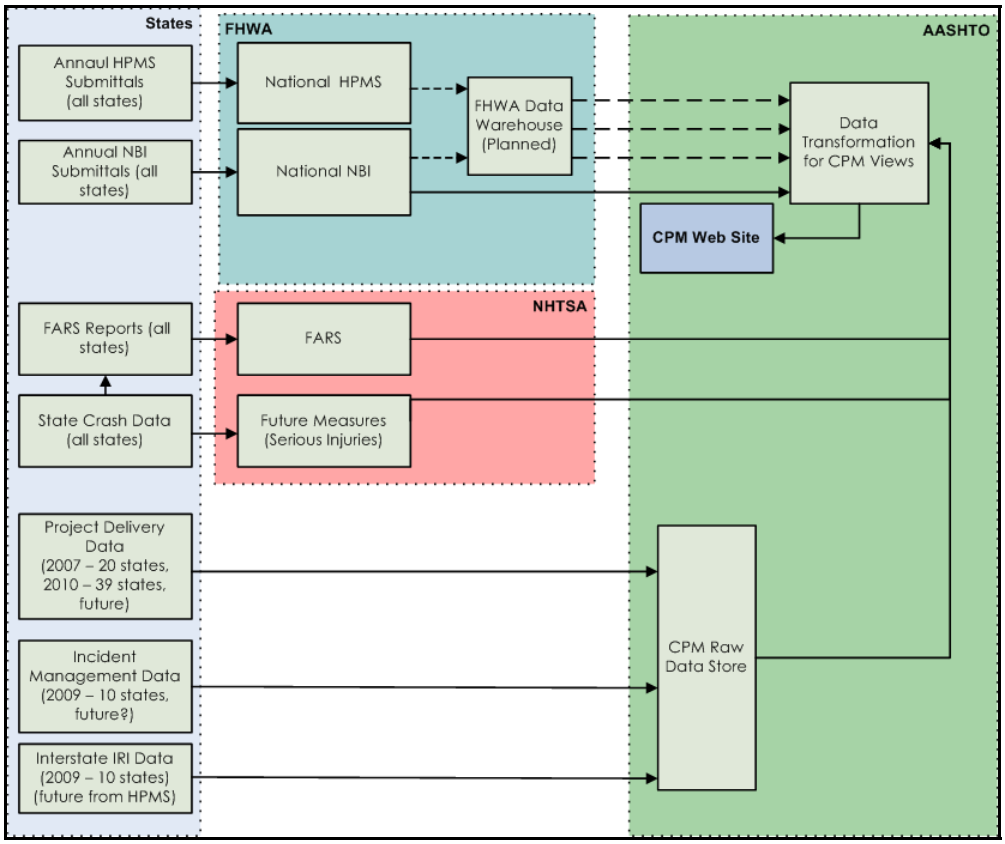

Figure 10 – AASHTO Lead

FHWA-Hosted with AASHTO Participation

FHWA would develop, host and maintain the infrastructure, which would include a joint AASHTO/FHWA Comparative Performance web site. FHWA would implement the CPM database by building upon its current initiative to link data from NBI, HPMS and other sources and develop a data warehouse. AASHTO would work closely with FHWA to design an outward facing web site that meets state DOT needs for comparative performance analysis. AASHTO would also work with FHWA to integrate data collected as part of CPM projects (for incident response and project delivery) into the FHWA data repository. This effort would be primarily funded internally by FHWA, with AASHTO participation as needed. This option is illustrated in Figure 11.

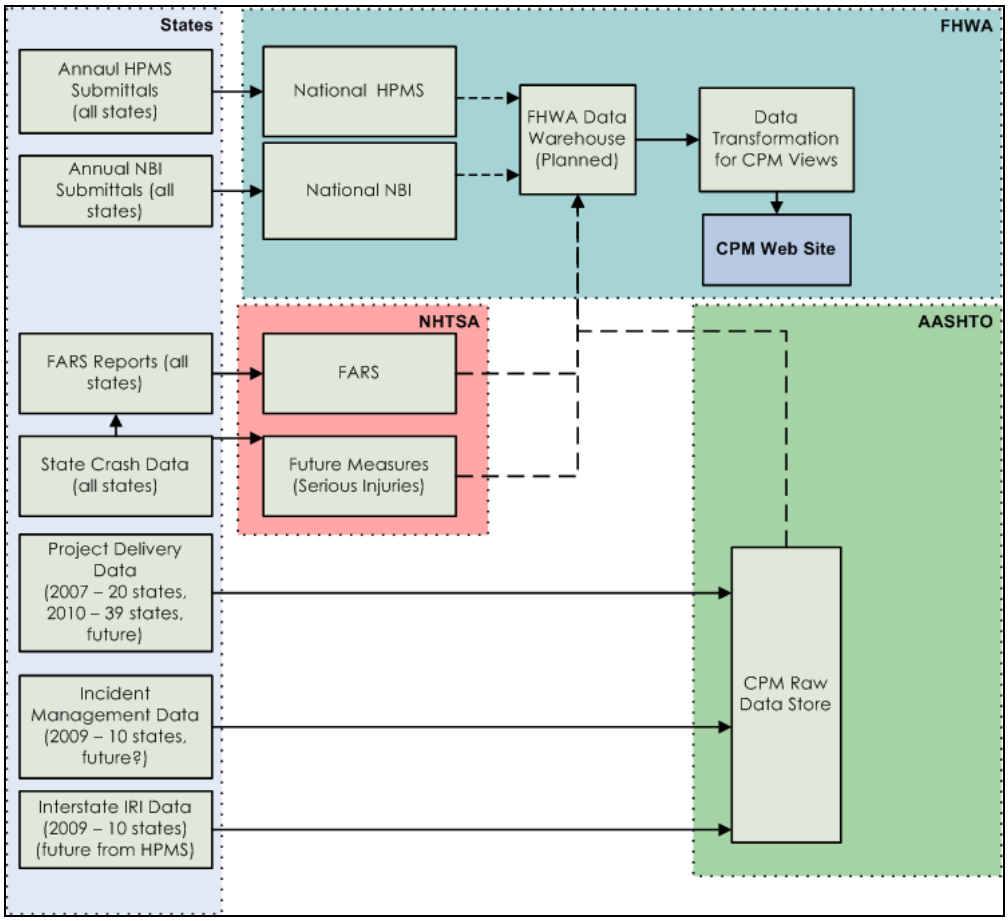

Figure 11 – FHWA Lead

Consortium/Pooled Fund

AASHTO and FHWA would work with interested states to initiate a pooled fund study for building and sustaining the CPM infrastructure. The work would be accomplished by a contractor, and the web site and database would be hosted either by the contractor or by the lead pooled fund study organization. FHWA and AASHTO would supply data. This option is illustrated in Figure 12.

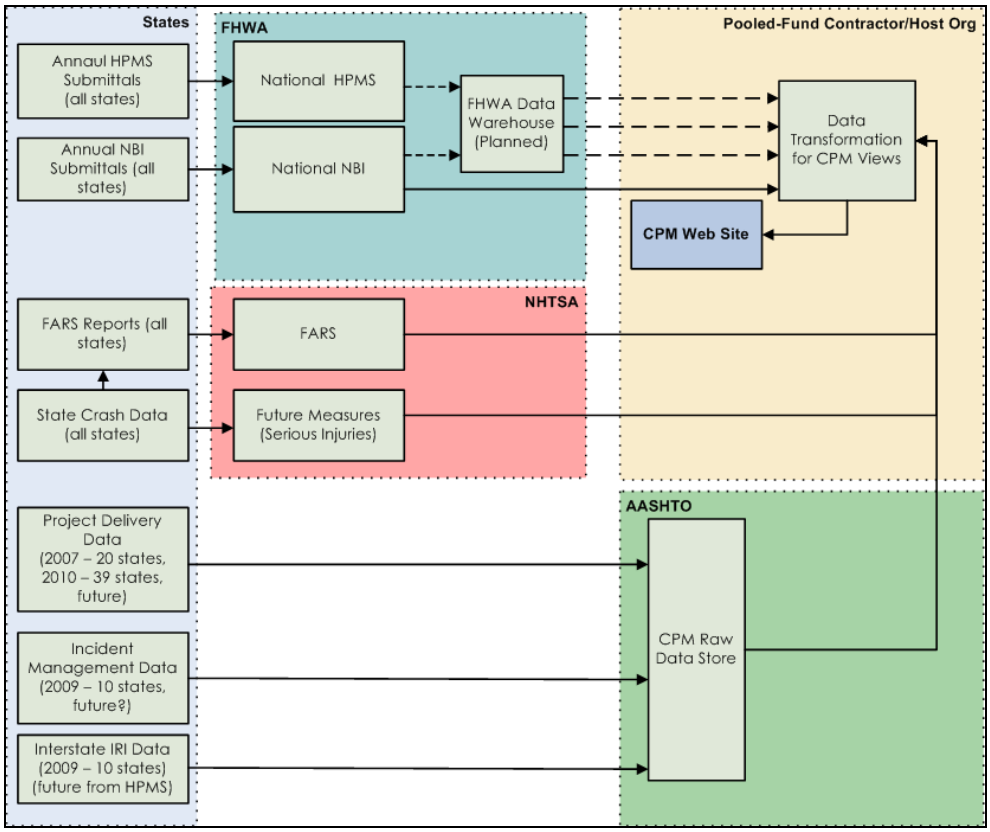

Figure 12 – Consortium/Pooled Fund

Evaluation of Stewardship Options

Table 5 displays approaches to Infrastructure development, maintenance and hosting, and data management under these three basic stewardship approaches.

Table 5 – CPM Infrastructure Stewardship Options: Description

	1. AASHTO Lead	2. FHWA Lead	3. Consortium/ Pooled Fund
Infrastructure Development	AASHTO via contractor	FHWA via contractor	Pooled Fund Study via contractor
Hosting and Maintenance	AASHTO – internal staff or contract	FHWA – internal staff or contract	Pooled Fund Contractor
Data Management	Annual data refresh (automated if possible) from FHWA and NHTSA sources; new data loaded by CPM study contractors	NBI, HPMS, Highway Statistics, Fatality data drawn from internal data warehouse to refresh CPM site; other CPM data provided by AASHTO (from CPM study contractors)	Annual data refresh (automated if possible) from FHWA and NHTSA sources; new data loaded by CPM study contractors

Each of these three stewardship options is designed to avoid any duplicative data reporting requirements for states, and to leverage existing data and performance management activities at FHWA. The advantage of Option 1 (AASHTO Lead) is that implementation can begin immediately in an incremental fashion (depending on resource availability), based on the priorities of AASHTO SCOPM. The advantage of Option 2 (FHWA Lead) is that it provides the potential for a more full featured and integrated set of tools for data exploration, with capabilities for mapping and linkages to additional data. In addition, by incorporating CPM displays into an existing FHWA initiative, incremental costs to meet CPM requirements would likely be substantially lower then those estimated above for an independent effort. However, Option 2 would likely take longer to implement given that the CPM component is just one element of a larger and more complex project with a multi-year timeframe. In addition, Option 2 would require that AASHTO SCOPM negotiate specific requirements in advance with FHWA to ensure that the CPM infrastructure objectives were addressed. There would also likely be less flexibility to make modifications in response to changing needs and priorities. Option 3 (Consortium/Pooled Fund) is similar to Option 1 with respect to cost and flexibility for incremental implementation. Lead time for this option would be in between Options 1 and 2 given the need to solicit participation. Option 3 can be considered if available funds from FHWA and AASHTO are not sufficient to cover infrastructure costs without additional contributions from states. Because a pooled-fund approach is by nature temporary and would lack the stature and commitment of a true national effort, this approach should be viewed as a stopgap measure only to be pursued if Option 1 or 2 cannot be moved forward.

Table 6 summarizes considerations for each option.

Table 6 – CPM Infrastructure Stewardship Options: Assessment

	1. AASHTO Lead	2. FHWA Lead	3. Consortium/ Pooled Fund
Lead Time to Implement	Can begin immediately	Uncertain – likely 2+ years	Uncertain, but could begin within 6 months given sufficient interest
Potential end-user features (e.g. integration with other data, mapping, etc.)	Limited to CPM-specific needs	High	Limited to CPM-specific needs
Flexibility to develop and modify based on current priorities	High	Low	Medium
Opportunities for minimizing total costs to develop and operate the infrastructure	Low-Medium – some synergies possible with existing web sites and database management functions	High – given opportunity to piggyback on existing technical and administrative infrastructure for performance measurement and data linking initiatives	Minimal
Additional cost to states	Possibly – Modest	None	Yes
Other considerations	May be limited support for AASHTO to take on substantial new information infrastructure and maintenance responsibilities	Strong partnership with AASHTO would be required to ensure that objectives of the CPM infrastructure were addressed	Lacks permanence and national stature

Of the three options presented, Option 2 (FHWA lead) offers the greatest opportunity for minimizing additional costs to states and maximizing functionality and value of the CPM infrastructure.

Lower Cost/Incremental Approach

The stewardship options presented above each presume that there would be both a core CPM database and a web site for data presentation. If resource constraints prevent immediate implementation of these components, or if the selected stewardship option involves substantial lead time for implementation, a lower cost, incremental approach could be considered. This approach would involve simply compiling the existing data into a consistent (and well documented) format, and making it available on a public web site (FHWA, AASHTO, DATA.GOV, etc.) for state DOTs, researchers and others to access.

By standardizing the format and publishing the data, individual organizations could develop (or customize existing) applications for data query, display and reporting to consume the data. For example, the Idaho DOT has developed its own web site with performance indicators for neighboring states (see http://itd.idaho.gov/dashboard/bridge_chart_states.htm).

Preparation of this initial data set could be accomplished via a project funded under the NCHRP 20-24(37) program. It would likely require on the order of $50,000-$100,000.
This approach would require a modest level of management oversight for data updates and inquiries/communication support. Note also that while it would satisfy the objective of preserving the CPM data, it would *not* provide for easy access to this data to an executive audience. Without a central web site, there would be little or no visibility for the data, so its use would be limited to those willing to make an investment to download and analyze it.

In order to ensure that any new data collected would be compatible with the existing CPM database, standard language for future CPM projects could be developed to provide a consistent approach to data structures, lists of values for classification variables, data documentation (metadata) and data provisioning. Data gathering, validation and presentation tools for participants for individual CPM projects could be developed under the auspices of each project, following the model provided by Project 20-24(37)A. As long as the underlying data structures for these individual efforts are compatible with the standards established for the CPM program as a whole, compilation of new data into the master database would be a straightforward task.